Preservice teachers' knowledge of proof by mathematical induction

Gabriel J. Stylianides · Andreas J. Stylianides · George N. Philippou

Abstract   There is a growing effort to make *proof* central to all students' mathematical experiences across all grades. Success in this goal depends highly on teachers' knowledge of proof, but limited research has examined this knowledge. This paper contributes to this domain of research by investigating preservice elementary and secondary school mathematics teachers' knowledge of proof by mathematical induction. This research can inform the *knowledge about preservice teachers* that mathematics teacher educators need in order to effectively teach proof to preservice teachers. Our analysis is based on written responses of 95 participants to specially developed tasks and on semi-structured interviews with 11 of them. The findings show that preservice teachers from both groups have difficulties that center around: (1) the essence of the base step of the induction method; (2) the meaning associated with the inductive step in proving the implication $P(k) \Rightarrow P(k + 1)$ for an arbitrary k in the domain of discourse of $P(n)$; and (3) the possibility of the truth set of a sentence in a statement proved by mathematical induction to include values outside its domain of discourse. The difficulties about the base and inductive steps are more salient among preservice elementary than secondary school teachers, but the difficulties about whether proofs by induction should be as encompassing as they could be are equally important for both groups. Implications for mathematics teacher education and future research are discussed in light of these findings.

G. J. Stylianides (✉)
University of Pittsburgh, Posvar Hall, Rm# 5517, Pittsburgh, PA 15260, USA
e-mail: gstylian@pitt.edu

A. J. Stylianides
University of Oxford, Oxford, UK
e-mail: andreas.stylianides@edstud.ox.ac.uk

G. N. Philippou
University of Cyprus, Nicosia, Cyprus
e-mail: edphilip@ucy.ac.cy

🖄 Springer

Introduction

There is a growing appreciation of the idea that *proof* should become central to all students' mathematical experiences as early as the elementary grades (e.g., NCTM, 2000; Schoenfeld, 1994). There are three main reasons for this increased emphasis on proof. First, proof is essential for deep learning in mathematics (e.g., Hanna, 2000) and within the conceptual reach of even young children in supporting classroom environments (e.g., Maher and Martino, 1996). Second, students' proficiency in proof can improve their mathematical proficiency more broadly, because proof is ''involved in all situations where conclusions are to be reached and decisions to be made'' (Fawcett, 1938, p. 120). Third, the difficulties that many high school and university students face with proof have been attributed, at least in part, to students' abrupt introduction to proof in high school (e.g., Marrades and Gutiérrez, 2000; Sowder and Harel, 1998), so it is important that students be offered appropriate experiences with proof earlier in their schooling.

A prerequisite in the effort to make proof central to school mathematics is that teachers of all levels have *solid* knowledge of proof, that is, sturdy knowledge that withstands attempts to inject contradictions into it. If teachers' knowledge of proof is *fragile*, that is, it is shaky and yields to attempts to inject contradictions into it (Steiner, 1989; see also Brousseau and Otte, 1991; Movshovitz-Hadar, 1993), it is likely that teachers will teach proof poorly or will not teach proof at all. As a result, the current vision in mathematics education to provide all students with rich opportunities to develop proficiency in proof will be undermined and the difficulties that many students face with proof (e.g., Fischbein, 1982; Healy and Hoyles, 2000) will probably persist.

The above remarks highlight the necessity for studies that will illuminate common strengths and weaknesses in preservice teachers' knowledge of proof, thus creating a research base for use by mathematics teacher educators in teacher preparation programs. This research base can inform the *knowledge about preservice teachers* that mathematics teacher educators need in order to teach proof effectively to preservice teachers. The knowledge about preservice teachers can be seen as an important component of the *pedagogical content knowledge* (Shulman, 1986) that is useful for mathematics teacher educators. Although many parallels have been drawn between the work of mathematics teacher educators in teaching preservice teachers and the work of mathematics teachers in teaching students (e.g., Cooney, 1994), the idea that mathematics teacher educators need to develop kinds of knowledge that are isomorphic to those that mathematics teachers need to develop has received significantly less attention. In science education, however, the notion of 'pedagogical content knowledge for science teacher educators' has been widely elaborated (e.g., Magnusson, 1996; Smith, 2000).

Despite the need for a research base of preservice teachers' knowledge of proof, little work has been done in this domain (Harel, 2002; Knuth, 2002; Martin and Harel, 1989; Movshovitz-Hadar, 1993; Simon and Blume, 1996; Stylianides, Stylianides, and Philippou, 2004). Most of the available studies have focused on secondary teachers and few have focused on specific methods of proof, such as proof by mathematical induction. This paper contributes to this research area by reporting on an empirical inquiry into preservice elementary and secondary mathematics teachers' knowledge of proof by mathematical induction. Our analysis also provides useful insight into preservice teachers' knowledge of proof more generally.

As background for the paper, we present the mathematical structure of proof by mathematical induction, elaborate on why this proof method warrants attention in teacher

preparation programs, and summarize research findings about adults' knowledge of it. Then, we define the specific scope of our paper, stating clearly our research goals.

Background

Mathematical structure of proof by mathematical induction

The proof method of mathematical induction is important for proving propositions of the form '$\forall n \in \mathbf{D}$, $P(n)$', where $P(n)$ is an open sentence asserting a relation among the elements of the set $\mathbf{D} = \{n | n \in \mathbf{N}, n \geq n_0\}$, the *domain of discourse*.[1] The method proceeds in two steps: the *base step*, which establishes $P(n)$ for an initial value $n = n_0$, and the *inductive step*, which proves the implication $P(k) \Rightarrow P(k + 1)$ for an arbitrary k in $\mathbf{D}$. Finally, by appeal to the Principle of Mathematical Induction (Peano's fifth postulate for the foundation of natural numbers) one can conclude that $P(n)$ holds for all n in $\mathbf{D}$. The Principle of Mathematical Induction permits the application of *modus ponens* to establish the truth of an infinite set of propositions $P(n)$, one for each n in $\mathbf{D}$. Below is a formal representation of proof by mathematical induction:

$$[P(n_0) \wedge \forall k \in \mathbf{D}, P(k) \Rightarrow P(k + 1)] \Rightarrow \forall n \in \mathbf{D}, P(n).$$

Why proof by mathematical induction warrants attention in teacher preparation programs

Apart from its mathematical significance, mathematical induction warrants attention in teacher education programs for at least four other reasons. The first two reasons relate to both elementary and secondary teachers, the third relates to secondary teachers, and the fourth relates to elementary teachers.

First, preservice teachers' engagement with any proof method offers them insight into what it means to *do* mathematics. It is part of the role of teacher education programs to help preservice teachers to appreciate the process of establishing new knowledge as well as other processes that are central in the discipline. Sullivan (2003) criticizes the fact that preservice mathematics and science teachers have limited knowledge of the breadth of their disciplines and the nature of disciplinary thinking, because he considers this limitation in preservice teachers' knowledge to be a key influence on the way they will teach. Also, he argues that this limitation is most likely a product of the way university courses are taught and assessed.

Second, mathematical induction can provide a context within which preservice teachers can improve their conceptions of proof and enhance their ability for logical thinking. Harel (2002) examined the effect of a novel instructional approach on students' conceptions of proof with a group of preservice secondary school mathematics teachers. He found that an approach based on the premise that students should develop the principle of mathematical induction through problems they can understand and appreciate improved students' ability to think logically.

---

[1] Mathematical induction can be used more broadly to prove statements about the elements of sets that can be placed in one-to-one correspondence with $\mathbf{N}$ (the set of positive integers).

Third, in many countries, secondary school mathematics teachers are expected to teach mathematical induction (e.g., NCTM, 2000, p. 345), so they need to have solid knowledge of this proof method in order to avoid fragile teaching of proof (Movshovitz-Hadar, 1993). One could argue further that, because secondary school students are expected to understand mathematical induction, elementary school teachers need to know this proof method in order to build the foundations for their students' future learning.

Fourth, although elementary school teachers are not expected to teach explicitly mathematical induction, solid knowledge of mathematical induction can enhance their ability to: (1) handle certain mathematical issues that may arise in the classroom, and (2) recognize rudimentary versions of mathematical induction in their students' arguments and promote students' understanding of these arguments. Regarding (1), the report 'Goals for mathematical education of elementary school teachers' (Cambridge Conference on Teacher Training, 1967) offers specific examples of how mathematical induction can be useful for elementary teachers:

> With [mathematical] induction we can prove theorems such as the fundamental theorem of arithmetic, the division algorithm, and the Euclidean algorithm; then we can point out for what kinds of integral domain the proofs hold. But there are other reasons for introducing a prospective K-6 teacher to mathematical induction. A teacher is very likely to be asked two specific questions by children 'Is there a biggest number?' and 'Is 1/0 equal to infinity?' (p. 64).

Solid knowledge of mathematical induction will help elementary teachers to reason about questions like those in the quotation and to figure out appropriate ways to communicate this reasoning to their students. Regarding (2), existing evidence from elementary classrooms shows that young children can produce rudimentary arguments by mathematical induction and engage in relevant activities (Maher and Martino, 1996; Reid, 2002). For example, Maher and Martino (1996) report an episode with a 9-year-old child who extended an earlier argument by cases to an argument by mathematical induction. Solid knowledge of mathematical induction would allow the teacher of the class to appreciate the mathematical value of this argument and design instruction so that the other students in the class would make sense of it.

Summary of research findings on adults' knowledge of proof by mathematical induction

Although proof by mathematical induction is central to the university mathematics curriculum, many undergraduate students lack understanding of this proof method, showing an indication of fragile knowledge. A number of studies (Dubinsky, 1986, 1990; Dubinsky and Lewin, 1986; Harel, 2002; Knuth, 2002; Movshovitz-Hadar, 1993) offer insights into university students' and other adults' knowledge of proof by mathematical induction.[2]

Movshovitz-Hadar (1993) presents evidence of knowledge fragility in connection with mathematical induction, as observed during her work with preservice secondary school teachers on an invalid proof that involved a statement with a complicated set of hypotheses. The cognitively demanding task, coupled with students' limited understanding of the

---

[2] The main goal of many of these studies has been the development of novel approaches to teaching proof by mathematical induction and not so much the examination of adults' knowledge of this proof method.

underlying principles of mathematical induction, made apparent the fragility of their knowledge about whether the proving process is circular. In a study with 22 mathematics majors, Dubinsky and Lewin (1986) found that students had difficulties in encapsulating modus ponens and in coordinating it with the structure of implication-valued function. In a different study with sophomore students (mainly computer science and engineering majors), Dubinsky (1986, 1990) identified two main student difficulties. The first relates to the character of the implication as a total entity: instead of trying to prove the inductive step $P(k) \Rightarrow P(k + 1)$, several students tried to prove $P(k + 1)$. The second relates to the essence of the base step, which was viewed by some students as a procedure without any real meaning.

Harel (2002), in his analysis of undergraduate students' proof schemes of mathematical induction, identified similar difficulties around the essence of the base step and also three other difficulties: (1) Students consider mathematical induction as a case of circular reasoning because they believe that the proof assumes that $P(n)$ is true for all positive integers; (2) Students believe that mathematical induction is a technique in which the drawing of a general argument is derived from a number of particular cases; and (3) Students follow the rule prescribed by the principle of mathematical induction without understanding what they are doing.

Finally, Knuth (2002) examined 16 in-service secondary school teachers' conceptions of proof, including proof by mathematical induction. Some of the subjects accepted an argument by mathematical induction as a proof to a given statement not because they understood the method, but because they heard of the method before and they knew that it offered a valid way to prove mathematical statements.

The scope of the paper

This paper aims to inform the *knowledge about preservice teachers* that mathematics teacher educators need in order to teach proof effectively in teacher preparation programs, by testing and extending what is currently known about preservice teachers' knowledge of proof by mathematical induction. We focus on three ideas important for the development of solid, as opposed to fragile, knowledge of proof by mathematical induction, and we explore the extent to and ways in which these ideas cause preservice teachers difficulties. Our operational definitions for 'solid' and 'fragile' knowledge are that these terms determine the endpoints of a continuum; the more difficulties preservice teachers have in these three ideas the more fragile their knowledge is.

1. The necessity of the base step in applying the induction method to prove a proposition of the form '$\forall n \in \mathbf{D}$, $P(n)$', where $\mathbf{D} = \{n | n \in \mathbf{N}, \ n \geq n_0\}$ (focal idea 1);
2. The meaning associated with the inductive step in proving the implication $P(k) \Rightarrow P(k + 1)$ for an arbitrary k in the domain of discourse $\mathbf{D}$ (focal idea 2); and
3. The relation between the domain of discourse $\mathbf{D}$ of the open sentence $P(n)$ and its truth set $\mathbf{U}$ when the former is a proper subset of the latter, and how this relation is mediated by a proof that purports to show that the sentence is true in $\mathbf{D}$ (focal idea 3).

The first two focal ideas have been identified by existing studies as two primary sources of undergraduate students' difficulties in proof by mathematical induction. However, as we noted earlier, few of these studies have included preservice secondary teachers in their samples, and none, as far as we know, has included preservice elementary teachers. Do

preservice elementary and secondary school teachers have similar difficulties in these focal ideas? In this paper, we will address this question. The third focal idea seems to have attracted no research attention thus far. Anecdotal evidence suggests that students' normal experiences with mathematical induction are in the context of proofs that are 'universal' (i.e., $\mathbf{D} = \mathbf{U}$) or proofs where the base step is always verifiable. In this paper, we will study possible issues of knowledge fragility by exposing preservice teachers to proofs that fall outside these normative patterns.

Method

Participants

The participants were senior undergraduate students of the University of Cyprus: 70 education majors (EMs) and 25 mathematics majors (MMs). These students participated in a larger study that examined the knowledge of different methods of proof held by the seniors of the respective Departments. The EMs constituted the 50% of the 2000-01 seniors of the Department of Education; all of them were taking a particular class during the Fall 2000 semester to which they were allocated randomly. The MMs were all the 2000–01 seniors of the Department of Mathematics.

With only few exceptions, those graduating from the Department of Education become elementary school teachers (grades K-6) and most of the graduates of the Department of Mathematics become secondary school mathematics teachers (grades 7–12). Therefore, EMs can be considered as preservice elementary school teachers and MMs as preservice secondary school mathematics teachers.

The program of study at the Department of Education includes several mathematics courses that emphasize logical thinking and ask for proofs. Specifically, it requires that all EMs take the following courses relevant to mathematics: (1) Two courses on foundations and fundamental concepts of mathematics;[3] (2) One introductory course on statistical methods and probability; and (3) One course on mathematics teaching and learning. These courses provide a fair amount of knowledge about different methods of proof, including proof by mathematical induction. Proof has a prominent place in the program of study at the Department of Mathematics that focuses on advanced mathematical content and logical thinking.

It is also worth noting that, contrary to what happens in some other countries, entry to the Department of Education has been highly competitive; guaranteed employment, high salary, and fringe benefits have influenced students to want to major in education.

---

[3] These two courses constitute an introduction to the development of fundamental mathematical ideas from ancient to modern times. The aim is to help preservice teachers see elementary mathematics from an advanced standpoint, but also to get a sense of contemporary mathematical thinking. The following topics are discussed: number systems; proofs of the Pythagorean theorem and properties of schematic numbers; the axiomatic foundation of geometry by Euclid, irrational numbers and Archimedes's method to find π; Ptolemy's theorem and the birth of trigonometry; a selection of Pappus's, Heron's, and Diofantus's works; the Fibonacci sequence; the discovery of calculus and some applications; Euclid's fifth postulate and elements of non-Euclidean geometries; non-commutative algebras (including elements of matrix algebra); set theory (including Cartesian products, binary relations, and functions), logical propositions (elements of algebra of propositions); Boolean algebra (logic gates and circuits).

Consider the following statement:

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

Study carefully the following proof for the above statement and answer the questions.

Proof:

I assume that (*) is true for $n = k$:

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

I check whether (*) is true for $n = k + 1$:

$1 + 3 + 5 + ... + (2k - 1) + (2k + 1) = (k^2 + 3) + (2k + 1) = (k^2 + 2k + 1) + 3 = (k + 1)^2 + 3$
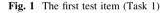
True.

Therefore (*) is true for every $n \in \mathbf{N}$.

(A) Choose the best response for the above proof:

1. The proof is invalid.

2. The proof shows that the statement is always true.

3. The proof shows that the statement is true in some cases.

4. I have no opinion.

(B) Use the space below to explain your thinking.

....

....

Fig. 1 The first test item (Task 1)

Data and procedure

All 95 participants responded to a specially designed test that included items on different methods of proof, such as proof by counterexample, contradiction, contraposition, and mathematical induction. The data from the test were supplemented by interviews with eight EMs and three MMs.

Figures 1 and 2 show the two test items on which we report in this paper.[4] The first item (Task 1) proposes a proof for a given statement that has some of the characteristics of proof by mathematical induction: the inductive step is applied correctly but the base step is missing. The students are asked to evaluate the validity of the proof and explain their thinking. The proposed proof is invalid; the best response to Part A of the task is choice 1. The truth set of the equation marked with (*) in the statement to be proved is the empty set.

The purpose for including this item was threefold. First, we wanted to examine whether students would realize that the base step of the induction method is missing (assuming that they would have already noticed the attempt to prove the statement by using this proof method). Second, we aimed to see whether students who would notice the omission of the base step would be in a position to explain why this step was so important (focal idea 1). Third, we intended to investigate the students' ability to assign meaning to the correct

---

[4] The original test items and the student responses we present in the next section were in Greek.

Consider the following statement:

For every natural number $n \geq 5$ the following is true: $1 \cdot 2 \cdot \ldots \cdot (n-1) \cdot n > 2^n$ (**)

Study carefully the following proof for the above statement and answer the questions.

Proof:

I check whether (**) is true for $n = 5$:

$1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120 > 2^5 = 32.$ True.

I assume that (**) is true for $n = k$: $1 \cdot 2 \cdot \ldots \cdot (k-1) \cdot k > 2^k$ (***)

I check whether (**) is true for $n = k + 1$:

$1 \cdot 2 \cdot \ldots \cdot (k-1) \cdot k \cdot (k+1) > 2^k \cdot (k+1)$ (using (***))

$> 2^k \cdot 2$ (since $k + 1 > 2$)

$= 2^{k+1}$ True.

Therefore (**) is true for every $n \geq 5$.

(A) Choose the best response for the above proof:

1.   The proof is invalid.

2.   The proof shows that the statement is always true.

3.   The proof shows that the statement is true in some cases.

4.   I have no opinion.

(B) Use the space below to explain your thinking.

………

………

(C) State what happens in the special cases where:

$n = 3$	A. The inequality is true.	B. The inequality is not true.
$n = 4$	A. The inequality is true.	B. The inequality is not true.
$n = 6$	A. The inequality is true.	B. The inequality is not true.
$n = 10$	A. The inequality is true.	B. The inequality is not true.

Fig. 2 The second test item (Task 2)

application of the inductive step in the absence of the base step (focal idea 2). What does the proposed proof show? For example, does it show that there is a natural number after which the equation (*) holds?

The second item (Task 2) has, in part, similar structure to the first: it includes a statement and a proposed proof for that statement, and is asking the students to evaluate the validity of the proof and explain their thinking. The proposed proof is valid (the best response to Part A of the task is choice 2), thus offering us the opportunity to examine the students' ability to recognize a correct application of proof by mathematical induction. The students are additionally asked to state whether the inequality marked with (**) in the statement to be proved is true or false in four particular cases: $n = 3, 4, 6,$ and 10 (Part C).

Table 1 Possible student responses to the two multiple-choice parts of Task 2 and possible student difficulties in relation to focal idea 3

Possible Response	Description of Possible Student Response		Description of Possible Student Difficulty Associated with Each Response
	Part A of Task 2	Part C of Task 2	
1	The proof shows that the statement is always true.	The inequality is *true* for all special cases ($n = 3$, 4, 6, and 10).	The proof showed the truth of the inequality outside its domain of discourse, possibly all natural numbers.
2	The proof shows that the statement is always true.	The inequality is *false* for $n = 3$ and 4, and *true* for $n = 6$ and 10.	The truth set of the inequality cannot include natural numbers outside its domain of discourse.
3	The proof is invalid.	The inequality is *false* for all special cases ($n = 3$, 4, 6, and 10).	Because the proof 'failed' to prove the statement, the truth set of the inequality is the empty set.
4	The proof shows that the statement is true in some cases.	The inequality is *false* for $n = 3$, and *true* for $n = 4$, 6, and 10.	The domain of discourse of the inequality should coincide with its truth set.

The truth set **U** of the inequality is $\{n | n \in \mathbf{N}, \ n \geq 4\}$, but the domain of discourse **D** does not include $n = 4$.[5]

We deliberately set up Task 2 in this way in order to create a rich context for examining further students' knowledge of proof by mathematical induction. In light of different possible responses to the two multiple-choice parts of Task 2 (Parts A and C), we could draw inferences about students' knowledge. In this paper, we focus on four possible student responses (summarized in Table 1) that help us examine student difficulties related to focal idea 3.

Interviews

The interviews were conducted with a purposeful sample (Patton, 1990) of 11 students whose responses in the test were of special interest to the research questions of the larger study. Special interest was defined in terms of exploring further common student responses, eliciting explanations to items where students did not provide one, and clari-

---

[5] The statements to be proved in Tasks 1 and 2 are closed sentences and, as such, are either true or false. Therefore, from a strictly mathematical standpoint, the adverb 'always' in choice 2 of Part A in each task is redundant. However, had we eliminated the word 'always' from the phrasing of choice 2, this would possibly make unclear to students the distinction between choices 2 and 3. Choice 3 was clearly a wrong choice for both tasks, but we considered it important to include it because of its potential to uncover knowledge fragility. Also, despite our effort to state clearly what we intended to be taken as the 'statement' under consideration in Task 2, some students might have missed the difference between the instruction, "Consider the following statement: For every natural number $n \geq 5$ the following is true: $1 \bullet 2 \bullet ... \bullet (n - 1) \bullet n > 2^n$ (**)," and the sentence "$1 \bullet 2 \bullet ... \bullet (n - 1) \bullet n > 2^n$ (**)" within that statement. If they took the latter to be the statement under consideration, then they might consider it true for only some values, which it is. Yet, the actual statement includes the quantifier "For every natural number $n \geq 5$" that makes it a closed sentence. The examination of the different ways students interpreted the various choices was part of our goals in this paper. Therefore, we not only asked students to select the *best* choice but also to *explain* their thinking. The results we obtained provide a lot of insight into how students interpreted each choice. For example, students who selected choice 3 in Task 2 seemed to face difficulties related to focal idea 3. Students who selected choice 2 seemed to have seen the choice as equivalent to saying 'the proof shows that the statement is true,' as we intended.

fying infrequent student responses. The interviews, which were conducted jointly by the first two authors, lasted approximately 35 minutes each and were audio taped and fully transcribed.

Before each interview session, the interviewers would note aspects of students' written responses that were of special interest and prepare two to five questions with relevant probes for use as necessary. Issues of interest to our research goals that emerged from the discussion were examined further. To preserve space for such examinations, the interviews had a semi-structured character. A typical interview session would begin with the interviewers probing the interviewee to elaborate on a particular response in his or her test.

Analysis

We obtained frequencies and percentages of students' responses to the multiple-choice questions in each task by whether they were to major in education or mathematics. These frequencies and percentages offered a general picture of students' thinking about proof by mathematical induction. We did not pursue any inferential statistics to explore possible relationships between the students' major and their performance in the tasks, for such analysis would not serve our research goal to gain insight into students' thinking. The descriptive statistics were supplemented by qualitative analysis of both the interview data and the students' written explanations in the test. The selection of student responses to be presented in the following section was guided by two criteria: (1) illumination of student thinking, and (2) illustration of a variety of student thinking strategies. To avoid improper generalizations, we do not report percentages of types of student explanations that exhibit particular characteristics; a number of students did not provide written explanations or provided cryptic explanations. Nevertheless, we note patterns (or the absence of patterns) in students' available explanations wherever this is informative.

Limitations

We acknowledge four limitations of this study that readers may take into account in appraising its outcomes and in designing future studies. The first concerns the small number of mathematics majors who participated in the study, even though they were all the seniors in the program that specific academic year. The second concerns the relatively small number of interviews we conducted, though they allow significant insight into the questions raised. The third, which might not allow generalization, relates to the characteristics of the participants and their programs of study (especially that of EMs). Nevertheless, because the two subgroups in our sample had different experiences with mathematical induction (in terms of breadth and depth), the common difficulties we found are likely to overlap with difficulties of students from the same or different programs of study in other countries. In this sense, our findings can be useful for mathematics teacher educators in other countries and also university instructors of mathematics more generally. The fourth limitation relates to our omission in Task 2 of the test to ask students to explain their responses in the four particular cases. Although we investigated this issue to some extent in the interviews, it would be useful to know, for example, how many of the students who recognized the validity of the proposed proof did the calculations to test the two cases that were within the domain of discourse ($n = 6$ and $n = 10$). Research shows that students often do not understand that a valid proof makes further checks superfluous (Fischbein, 1982).

Results

The results are organized in three parts. In the first two, we report separately findings related to Tasks 1 and 2. In the third, we focus on how our results inform the issue of student difficulties with regard to the three focal ideas.

Task 1

Table 2 summarizes the student responses to Part A of Task 1. The vast majority of the MMs and almost a quarter of the EMs considered correctly the proof to be invalid. The written explanations of these students were mostly limited to saying that the proof did not check for $n = 1$. For example, students EM23 (education major #23) and EM38 wrote:

EM23: The proof is invalid because its first step is missing.
EM38: The proof is invalid because it begins by checking for $n = k$, whereas it should check for $n = 1$.

We interviewed both of these students and none of them took the extra step to check the base step; this suggests a focus on the form (i.e., the appearance) of the presented argument. Some of the other students who considered that the proof was invalid also checked the base step. However, even in these cases the students had difficulty explaining the essence of the base step. The interview excerpt with student MM19 (mathematics major #19) is indicative.

MM19: Here we use the induction method but we don't check for $n = 1$. When we check for $n = 1$ we see that the equation doesn't hold.
I (interviewers): Why is it necessary to check for $n = 1$?
MM19: We are supposed to check for $n = 1$ because the induction method states that a sentence is true for all natural numbers, from 1 up to ... for all n. If we don't check for $n = 1$ and we begin by setting $n = k$ then we begin from something arbitrary that we don't know whether it is true. Therefore, the induction hypothesis that the equation holds for $n = k$ is false and we cannot proceed any further to prove the equation for $n = k + 1$. Had the equation been true for $n = 1$, one would be able to make the induction hypothesis and say that, if it is true for $n = 1$ and $n = k$, and is proved to be true for $n = k + 1$, then the equation is true in general.

Table 2 Percentages of student responses to Part A of Task 1 by major

Response Option	Major	
	Education[a]	Mathematics[b]
The proof is invalid[c]	24	88
The proof shows that the statement is always true	53	0
The proof shows that the statement is true in some cases	17	12
I have no opinion	6	0
Total	100	100

Note. The values represent percentages (rounded to the nearest integer) within major.
[a]$n = 70$. [b]$n = 25$. [c] Best response option for Part A of Task 1.

From the above excerpt, it seems that the student MM19 does not have a solid knowledge of the inductive step either. He considers invalid the proof of the implication 'if it holds for k then it holds for $k + 1$' because the truth of the equation has not been checked for $n = 1$. As it is illustrated further in the following excerpt, he does not appear to understand the meaning of the hypothesis '*if* it holds for k':

I: Is there a natural number for which the equation is true?
MM19: There may be such a natural number, but the method is supposed to show the truth of the equation for all the natural numbers. Therefore it is false.
I: The fact that if it holds for k then we can prove that it also holds for $k + 1$...
MM19: This is also false.
I: What is false?
MM19: It is false to say that 'if it holds for k we can prove that it also holds for $k + 1$' because the equation is not true for $k = 1$.

The student's fragile knowledge of the induction method was further revealed in the subsequent exchange:

I: If the equation were verified for a natural number $x \neq 1$, would you say that it holds for all natural numbers greater than or equal to x?
MM19: I don't know. This may or may not happen. I know that for the method to be correct and for the equation to be true for all n, one needs to begin the induction as follows: check for $n = 1$ and then prove that if it is true for $n = k$ it is true for $n = k + 1$. One needs to check cases [to be sure]. I haven't checked any cases.

The inductive step caused a different difficulty to student EM23. As we mentioned above, in her test, EM23 (like MM19) considered the proposed proof as invalid. In the interview session, she said that the proof of the implication $P(k) \Rightarrow P(k + 1)$ guaranteed the existence of a natural number x for which the statement is true. She then used the 'existence' of x and modus ponens to explain that one could deduce that the statement is true for all natural numbers greater than or equal to x.

I: Are you saying that it is possible to find a number for which the equation is true?
EM23: Yes, I believe that one such number can be found.... I believe that there is a number for which the equation is true.
I: What makes you so certain about the existence of a number for which the equation is true?
EM23: Basically I believe that because I can prove that it holds for $k + 1$ by assuming that it holds for k, where k can be any number. This means that if this number is 1 then the equation also holds for 2 and we continue this way by adding 1 each time to get 3 and any other number we want.

Some other students who considered invalid the presented proof faced a different difficulty. Although they showed solid knowledge of the role of the base step in the induction method, they did not appear to be clear about the idea that the inductive step did not actually prove the equation for $n = k + 1$. For example, student EM56 seems to believe that the proposed proof showed $P(k + 1)$ rather than $P(k) \Rightarrow P(k + 1)$, but it is unclear what she actually means by saying that the equation was proved to hold for $n = k + 1$.

EM56: The equation holds neither for $n = 1$ nor for $n = 2$. We may have proved that the equation holds for $n = k + 1$, but this doesn't mean anything because we didn't find a specific number for which the equation is true. The equation is definitely not

true for every n that belongs to $\mathbf{N}$. It may be true for some number and beyond, assuming that we manage to find such a number.

Turning now to the EMs who said that the proposed proof shows that the statement is always true, we observe that many of them thought that the proof of the inductive step implies the truth of the equation for all natural numbers. These students did not seem to understand the function of the base step.

EM1: Since the equation holds for both $n = k$ and $n = k + 1$ then it is always true.
EM32: The proposed proof uses the method of mathematical induction. According to this method, when an equation holds for $n = k$ and we prove that it also holds for $n = k + 1$, then we conclude that the equation is always true.
EM54: By checking the equation for $n = k + 1$ we check the truth of the equation for all numbers.

Some other students, who also supported the validity of the proof, based their responses on the 'observed pattern' (EM51) or on the 'examination of some numbers' (EM14). Yet, it is unclear what these students took as a pattern or which numbers they tested, given that there is no natural number that satisfies the equality in the task. It is possible, though, that they were referring to the inductive step.

EM51: The statement is always true as it represents a pattern. Because of the pattern the statement holds for every $n \in \mathbf{N}$.
EM14: The proposed proof shows that the statement is always true, because the equation was tested for some natural numbers.

A considerable number of students, both EM and MM majors, claimed that the proposed proof showed that the statement is true in some cases. Analyzing their written responses to the task, we observed several variations of the idea that the equality is true for all natural numbers greater than or equal to a natural number $x \neq 1$. Some students said that the statement holds for every $n \in \mathbf{N} - \{1\}$ (EM4), others for all $n \geq k$ (MM4) or for all $n \geq k + 1$ (EM16), and others for all natural numbers greater than a specific number that needs to be found (MM9).

EM4: We assumed that the statement holds for a random number $n = k$. If we manage to prove that the statement also holds for the subsequent number, then we can conclude that the statement holds for all $n \in \mathbf{N}$. I check for $n = 1$. The statement doesn't hold for $n = 1$. Therefore, the statement holds for every $n \in \mathbf{N} - \{1\}$.
MM4: The proposed proof shows that the statement holds for some cases, because it holds for all $n \geq k$ and not for every $n \in \mathbf{N}$.
EM16: The proposed proof didn't check the statement for $n = k-1$ or for values $< k-1$ ... The proposed proof shows that the statement holds only for $n \geq k + 1$.
MM9: The proposed proof shows that the statement holds for some cases. I assume that the statement holds for $n = k$ and, based on this assumption, I prove that the statement also holds for $n = k + 1$. For example, if the statement holds for $n = 2$ then it holds for $n = 3$. If it holds for $n = 3$ then it holds for $n = 4$. If we prove that the statement holds, let's say for $n = 5$, then it will hold $\forall n \geq 5$ but not necessarily for $n = 1, ..., 4$.

Moreover, some students (EM58) thought that for the proof to be complete it would also need to 'check' for $n = k + 2$, $n = k + 3$, etc.

*EM*58: The fact the statement was checked for $n = k + 1$ doesn't mean that it's true for all cases; it just means that it holds for some cases. To make sure that the statement is always true we need to also test it for $n = k + 2$, $n = k + 3$, ..., $n = k + n$.

Task 2

Table 3 summarizes the student responses to Part A of Task 2 by major. The vast majority of MMs and more than half of the EMs said that the proposed proof showed that the statement is always true. Some EMs who supported the validity of the proof faced difficulties in formulating a mathematically accurate explanation. The responses of the students EM24 and EM50 illustrate these difficulties and raise additionally the issue of whether the students' belief about the validity of the proof was grounded on reason.

EM24: The proof shows that the statement is always true. Most of the possible cases have been checked and, therefore, we can conclude that the statement is true in general.

EM50: The statement holds. However, the way mathematical induction is applied isn't the best possible, because it doesn't prove that the statement is also true for $n = 6$ (since 6 is greater than 5) before proceeding with $n = k$.

On the other hand, many MMs could explain their choice in Part A of the task in a sufficient way. The explanation of student MM10 is indicative.

MM10: You used the method of mathematical induction. You checked all the steps of the method and you concluded that they are applied correctly. Therefore, we can conclude that the statement always holds. By saying 'always' we mean 'always' as it is indicated in the context of the statement, that is, for $n \geq 5$.

This comment about the meaning of the word 'always' with regard to the domain of discourse of the statement in Task 2 is at the heart of focal idea 3 we aimed to investigate. The data suggest that many EMs considered that the proposed proof showed the statement to be true in some cases, because they thought that a valid proof should show that: (1) the sentence in the statement is true for all natural numbers, or (2) the sentence is true for all natural numbers that belong to its truth set (in this particular case, $\{n | n \in \mathbf{N}, \ n \geq 4\}$).

Table 3 Percentages of student responses to Part A of Task 2 by major

Response Option	Major	
	Education[a]	Mathematics[b]
The proof is invalid	13	8
The proof shows that the statement is always true[c]	54	92
The proof shows that the statement is true in some cases	29	0
I have no opinion	4	0
Total	100	100

Note. The values represent percentages (rounded to the nearest integer) within major.

[a] $n = 70$. [b] $n = 25$. [c] Best response option for Part A of Task 2.

Table 4 Percentages of correct student responses to Part C of Task 2 by major

Major	Part C of Task 2			
	$n = 3$	$n = 4$	$n = 6$	$n = 10$
Education[a]	83	56	87	91
Mathematics[b]	96	64	100	96

Note. The values represent percentages (rounded to the nearest integer) within major.
[a] $n = 70$. [b] $n = 25$.

> *EM51*: The proof shows that the statement is true in some cases, because if we check some other numbers, e.g., 3, the statement is false.
> *EM9*: The proof shows that the statement is true in some cases. The statement is always true for $n \geq 5$. I don't know whether it's true for $n < 5$.

The same thinking that led some students to conclude that the proposed proof showed that the statement is true in some cases, led others to consider the proposed proof as invalid.[6]

> *EM20*: The proof is invalid. The testing of cases should begin from the first natural numbers: 1, 2, 3, 4. The statement is also true for $n = 4$.
> *EM49*: The proof is invalid because the statement is true for $n \geq 4$.
> *EM52*: The proof is invalid because the statement is false for $n = 4$.

Student EM20 seems to believe that the validity of a proof by mathematical induction depends on whether the proof establishes the truth of the mathematical relationship under consideration on the entire set of natural numbers rather than on its domain of discourse. Students EM49 and EM52 rejected the validity of the proposed proof for opposite reasons. EM49 rejected the proof because he found a value for n (= 4) outside the domain of discourse for which the inequality is satisfied. He seemed to believe that the proof is invalid because it is not as encompassing as it could be. EM52 failed to see that the inequality is satisfied for $n = 4$ and considered that this violated the assertion 'the proof shows that the statement is always true.' He therefore appears to think that a valid proof would show the truth of the inequality over a broader set than its domain of discourse, possibly all natural numbers.

Data from Part C of Task 2 offer further insight into students' understanding of the relation between the domain of discourse of the inequality that was proved and its truth set. Table 4 summarizes the percentages of correct student responses to each of the special cases by major. The highlight of the table is the failure of many students from both majors to realize that the inequality is true for $n = 4$. Given the simplicity of the calculations required to check for $n = 4$, it is plausible to assume that the students reached their conclusion based on an erroneous thinking. This thinking was probably associated with the fact that number 4 is not included in the domain of discourse of the inequality. The difference in the percentages of success between the special cases $n = 3$ and $n = 4$ may be attributed to the fact that the latter belongs to the truth set of the inequality whereas the former does not. A student who believed that the inequality could not hold for values outside its domain of discourse would inadvertently get $n = 3$ right and $n = 4$ wrong. The

---

[6] The MMs who considered invalid the proposed proof in Task 2 did not provide a written explanation for their selection and, therefore, we cannot examine further their thinking.

Table 5 Frequencies of selected student responses to Parts A and C of Task 2 by major

Response Type	Description of Response Types and Frequencies by Major	
	Part A of Task 2[a]	Part C of Task 2[b]
0	The proof shows that the statement is always true. (38, 23)	The inequality is *false* for $n = 3$ and *true* for $n = 4, 6,$ and 10. (15, 13)
1	The proof shows that the statement is always true. (38, 23)	The inequality is *true* for all special cases ($n = 3, 4, 6,$ and 10). (3, 1)
2	The proof shows that the statement is always true. (38, 23)	The inequality is *false* for $n = 3$ and 4, and *true* for $n = 6$ and 10. (18, 8)
3	The proof is invalid. (9, 2)	The inequality is *false* for all special cases ($n = 3, 4, 6,$ and 10). (2, 0)
4	The proof shows that the statement is true in some cases. (20, 0)	The inequality is *false* for $n = 3$, and *true* for $n = 4, 6,$ and 10. (8, 0)

[a] The first number in each parenthesis in this column represents the number of EMs who responded the specified way in Part A of Task 2; the second represents the corresponding number of MMs.

[b] The numbers in each parenthesis in this column represent the numbers of students who responded the specified way in Parts A *and* C of Task 2. The first represents the number of EMs and the second the number of MMs.

higher percentages of success for the cases $n = 6$ and $n = 10$ were expected, given that these cases belong to the domain of discourse and that the majority of the students accepted the validity of the proposed proof. Also, the calculations were easy for the students who chose to do them.

Table 5 presents a detailed analysis of the results obtained from Parts A and C of Task 2; it provides five response types, each corresponding to a different combination of responses to the two parts. Response Type 0 represents the correct responses to both parts of the task. Response Types 1 through 4 correspond, respectively, to Possible Responses 1 through 4 described in Table 1.

A significant number of students from both majors, 38 EMs and 23 MMs, recognized the validity of the presented proof, thus responding correctly to Part A of Task 2. From these students, only 15 EMs and 13 MMs responded correctly to all four special cases of Part C (Response Type 0), while three EMs and one MM said that the inequality is true for all special cases (Response Type 1). The latter response type suggests a belief that the proof showed the truth of the inequality for values outside its domain of discourse (possibly all natural numbers). Almost all other students who responded correctly to Part A of Task 2 (18 EMs and eight MMs) said that the inequality is false for $n = 3$ and $n = 4$, and true for the other two special cases (Response Type 2). This response type is most likely associated with the misconception that the truth set cannot include elements outside the domain of discourse.

The remaining two response types are associated only with EMs. From the nine EMs who considered the presented proof as invalid, two said that the inequality is false for all special cases (Response Type 3). These students probably believed that, because the proof 'failed' to prove the statement, the truth set of the inequality is the empty set. Finally, from the 20 EMs who said that the presented proof shows that the statement is true in some cases, eight said that the inequality is false for $n = 3$ and true for the other three special cases (Response Type 4). These students most probably thought that the domain of discourse of the inequality should coincide with its truth set.

Some of the interviews shed further light on students' thinking with regard to the special cases. For example, student EM38, whose written responses fit in Response Type 0, had difficulty understanding the 'mismatch' between the domain of discourse of the inequality and its truth set. However, after some probing from the interviewers, EM38 appeared to have grasped the relation between these two sets in the context of the given proof.

I: Do you find problematic the fact that the statement says that the inequality holds for all $n \geq 5$, but you said here [pointing to his test] that the inequality also holds for $n = 4$?
EM38: ... I believe that there is a problem here. Perhaps the source of the problem is that the proof doesn't specify the value of k. I assumed that k is greater than or equal to 5, and this might have been the reason I said that the statement is always true.
I: Now that you have the opportunity to think about this problem again, which of the multiple-choice options [referring to Part A of Task 2] would you choose?
EM38: I wouldn't choose this option [he refers to choice 1 of Part A] because the statement holds for $n \geq 5$. The issue here is whether the statement also holds for some values smaller than 5.
I: Are you saying that proving the inequality for $n \geq 5$ excludes the possibility that the inequality also holds for smaller values of n?
EM38: Oh! The statement doesn't say '*only* for $n \geq 5$'! Therefore, it leaves open the possibility for other values. Therefore, the statement is true.

Three Focal Ideas

Below we summarize how the results inform the issue of student difficulties (indicative of knowledge fragility) with regard to the three focal ideas. Because focal ideas 1 and 2 (FI 1 and FI 2) were both investigated in the context of Task 1, we discuss them together. The discussion of focal idea 3 (FI 3) is based on findings from Task 2.

The necessity of the base step in applying the induction method (FI 1), and the meaning associated with the inductive step in proving the implication $P(k) \Rightarrow P(k + 1)$ for an arbitrary k in the domain of discourse of the open sentence $P(n)$ (FI 2). A considerable percentage of EMs did not realize the essence of the base step in the application of the induction method in Task 1, thus saying that the proposed proof shows that the statement is true for all n in $\mathbf{N}$. It seems that many of these students applied intuitively the false implication rule: $[\forall k \in \mathbf{N}, P(k) \Rightarrow P(k+1)] \Rightarrow \forall n \in \mathbf{N}, P(n)$. Although the vast majority of MMs recognized that the omission of the base step constitutes an incomplete application of the induction method in Task 1, some of them could not articulate clearly the importance of this step. A notable percentage of students of both majors believed that the proposed proof shows that the statement is true in some cases. The inference rule that seemed to have guided their thinking was the following: $[\forall k \in \mathbf{N}, P(k) \Rightarrow P(k+1)] \Rightarrow \forall n \in \mathbf{S}, P(n)$, where $\mathbf{S}$ is a proper subset of $\mathbf{N}$ of the form $\{n | n \geq m > 1, n, m \in \mathbf{N}\}$. In other words, they seemed to think that the proof of the implication $P(k) \Rightarrow P(k + 1)$ for an arbitrary k in $\mathbf{N}$ guarantees the existence of a natural number $m \neq 1$ such that $P(m)$ holds. In turn, the existence of such a number enables application of modus ponens, thereby establishing the truth of the proposition $P(n)$ for all $n \geq m$. Finally, a considerable number of EMs thought that the inductive step shows that the equality is true for $n = k + 1$, which suggests again their difficulty to grasp the meaning of the implication $P(k) \Rightarrow P(k + 1)$.

The relation between the domain of discourse **D** *of the open sentence P(n) and its truth set* **U** *when the former is a proper subset of the latter, and how this relation is mediated by a proof that purports to show that the sentence is true in* **D** *(FI 3).* The analysis of student responses in Task 2, where $\mathbf{D} = \{n|n \in \mathbf{N}, n \geq 5\}$ and $\mathbf{U} = \{n|n \in \mathbf{N}, n \geq 4\}$, suggests that a significant percentage of students of both majors who recognized correctly the validity of the proposed proof thought that it is impossible for the truth set of the inequality to include a number outside its domain of discourse. These students considered that the inequality is false for $n = 3$ and $n = 4$, and said that the inequality is true only for the two values that belong to **D**. The responses of a considerable number of EMs that the presented proof shows that the statement is true in some cases seemed to have been influenced by the belief that the domain of discourse **D** of the inequality in the statement to be proved should coincide with its truth set **U** in order for the statement to be always true. These students' observation that $4 \in \mathbf{U}$ and their knowledge of the fact that $4 \notin \mathbf{D}$ presumably affected the way they evaluated the validity of the proposed proof. The participants did not particularly show any other difficulties associated with FI 3.

Discussion

In this paper, we have investigated preservice teachers' knowledge of proof by mathematical induction in order to: (1) extend the existing research base of preservice teachers' knowledge of methods of proof, and (2) inform the knowledge about preservice teachers that mathematics teacher educators need to teach proof effectively in teacher preparation programs. Our findings showed that preservice secondary teachers were more successful in obtaining correct answers than preservice elementary teachers. However, students from both groups faced three common difficulties in explaining the induction method. The difficulties centered around: the essence of the base step (Difficulty 1 [D1]); the meaning associated with the inductive step in proving the implication $P(k) \Rightarrow P(k + 1)$ and the inferences that can be drawn from this proof about the existence of a natural number m such that $P(m)$ holds (D2); and the possibility of the truth set of a sentence to include values outside its domain of discourse (D3). D1 and D2 were more salient among preservice elementary than secondary teachers, but D3 seemed to be equally important in the two groups.

The findings regarding D1 and D2 are aligned with findings of earlier studies with preservice secondary mathematics teachers or other undergraduate students. Dubinsky and Lewin (1986) report that mathematics majors faced difficulties in encapsulating modus ponens and in coordinating it with the structure of implication-valued function. In our study, students did not understand that it takes both the base and the inductive steps to deduce by modus ponens the truth of an infinite set of propositions. In addition, some students thought that the inductive step in itself ensures the truth of a statement for all natural numbers greater than a specific number that needs to be found. This result extends the field's understanding of students' knowledge of the meaning of the inductive step. Dubinsky (1986, 1990) reports two difficulties that are also consonant with D1 and D2. The first relates to the character of the implication as an object; similar to what we found, many sophomores in Dubinsky's study tried to prove $P(k + 1)$ rather than $P(k) \Rightarrow P(k + 1)$. The second relates to the essence of the base step; Harel (2002) reports that undergraduate students tend to view the base step as meaningless or non-essential.

In sum, our findings regarding D1 and D2 support and extend existing findings about undergraduate students' difficulties in proof by mathematical induction, and provide

evidence in support of the hypothesis that preservice elementary and secondary school teachers face similar difficulties when they engage with this proof method.

The findings associated with D3 deserve special attention, for the investigation of possible student difficulties in proofs that are not as encompassing as they could be seems to have attracted no research attention thus far. Our investigation in this area complements Fischbein's (1982) investigation, which showed that many students have difficulties understanding the logical equivalence of the following two affirmations in the context of a proposed (valid) proof for a theorem: (1) to agree with the validity of the proof, and (2) to agree that the (accepted) proof guarantees the truth of the theorem for all elements in its domain of discourse. In our investigation of D3, we essentially examined whether students understood that the following two affirmations are *not* logically equivalent: (1) to agree with the validity of the proof, and (2) to agree that the (accepted) proof guarantees that the theorem cannot be more general. A significant number of preservice teachers from both groups who recognized the validity of the proof in Task 2 thought that it was impossible for the truth set of the open sentence to include a number outside the set covered by the proof. Future research needs to triangulate these findings with other student populations and in the context of test items that will examine: (1) students' understanding of the logical equivalence or non-equivalence of all the different affirmations mentioned earlier, and (2) manifold proof methods. For example, a test item can focus on the following open sentence whose truth set is the real numbers (De Moivre's theorem):

$$(\cos A + i\ \sin A)^n = \cos(nA) + i\ \sin(nA).$$

The item can begin by proposing a valid proof by mathematical induction that demonstrates the truth of this sentence for all natural numbers and then ask the students who will recognize the validity of the proof to explain: (1) whether the sentence is true for specific natural numbers (Does one need to check them to be sure?), and (2) whether the sentence is true for specific non-integer real numbers that are relatively easy to verify with the use of trigonometric identities (Does one need to check them to be able to answer the question, or does one know for sure that the sentence is true/false?). This item is interesting because, given the use of mathematical induction, the mismatch between the truth set and the set covered by the proof is unavoidable: mathematical induction cannot be used to prove sentences over the entire set of real numbers. Accordingly, students will be exposed to a situation where a specific proof method cannot cover the entire truth set of an open sentence.

We turn now to what might have caused preservice teachers' fragile knowledge in D1, D2, and D3. This issue requires further research, but one plausible hypothesis is that the fragile knowledge is rooted in students' prior experiences with didactic contracts that prevail in high school and even university mathematics and that promote the following misconceptions: (1) the base step in proofs by mathematical induction is always verifiable and thus one only needs to worry about the inductive step, and (2) proofs are always as encompassing as they can be.

Mathematics teacher educators can use the findings of this paper related to difficulties D1–D3 to design instructional conditions that will first elicit and then address preservice teachers' difficulties in mathematical induction. If preservice teachers' difficulties remain tacit and pass unchallenged through mathematics teacher education, they are likely to become sources of misconceptions or reasons underlying fragile instruction of proof in school mathematics.

There are at least four ways in which the findings of this paper can be useful to mathematics teacher educators. First, the three difficulties we identified draw attention to a set of ideas that mathematics teacher educators need to emphasize in their instruction for preservice teachers to have opportunities to develop solid knowledge of mathematical induction.

Second, a precise presentation of mathematical induction in itself is not enough to advance preservice teachers' knowledge of this proof method. Our analysis showed, for example, that, although many students knew that the base step is a necessary component of the induction method, they could not explain why this step is essential or what can be the consequences from its omission. Also, some of them focused on the form (i.e., the appearance) of the proposed proof in Task 1 to judge its validity rather than on the reasoning it was based on; they considered the proposed proof to be invalid because it omitted the base step. Thus, they demonstrated what Sowder and Harel (1998) call 'ritual proof scheme.'

Third, our findings can inform mathematics teacher educators' knowledge of preservice teachers' understanding of proof more broadly. For example, preservice teachers' difficulties related to the encompassing character of proofs are presumably not specific to proof by mathematical induction.

Fourth, our results support mathematics teacher educators' efforts to organize instructional situations in which preservice teachers' difficulties in proof by mathematical induction (or proof more broadly) surface and become the objects of reflection. Specifically, the two tasks we discussed in this paper can be used to engage preservice teachers in activities that will aim to reduce the fragility of their knowledge and accelerate the process of its crystallization (Movshovitz-Hadar, 1993, p. 266). By breaking the boundaries of students' normal experience, the two tasks set up situations where procedural knowledge of the induction method is not enough for successful performance. Task 1 presents a situation where the base step is omitted and the truth set of the sentence of interest is the empty set. In implementing this task, mathematics teacher educators may ask questions like, 'Why is the base step essential?' and 'What does the inductive step prove?', in order to bring to the fore issues of meaning. Task 2 presents a situation where the domain of discourse and the truth set of a sentence that is proved by a valid proof are different. In implementing this task, mathematics teacher educators may ask questions like, 'Is there a relation between the truth set of the sentence and the domain of discourse?' and 'Do the two sets have to be the same?', in order to help preservice teachers think about important ideas related to proof. To manage successfully their discussions with preservice teachers, mathematics teacher educators need to be able to anticipate common responses by preservice teachers. The interview segments and the written responses reported in this paper can be useful in this respect.

The four ways we described above in which the results of this paper can be useful to mathematics teacher educators apply also for other mathematics instructors who teach proof to university students. As we mentioned earlier, the two subgroups in our sample had different experiences with mathematical induction (in terms of breadth and depth), and so the common difficulties we identified (D1–D3) are likely to be faced by university students from other programs of study.

To conclude, teachers' knowledge of mathematics is a central issue in policy and research discussions about the improvement of students' mathematical education. Given the increased emphasis on offering students of all grades rich opportunities to develop proficiency in proof, it is important that mathematics teacher educators develop a good sense of how preservice teachers understand proof. This way, mathematics teacher

educators will be well positioned to design and implement appropriate instructional interventions to help preservice teachers develop a solid knowledge of proof.

References

Brousseau, G., & Otte, M. (1991). The fragility of knowledge. In A. J. Bishop, S. Mellin-Olsen, & J. van Dormolen (Eds.), *Mathematical knowledge: Its growth through teaching* (pp. 13–36). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Cambridge Conference on Teacher Training. (1967). *Goals for mathematical education of elementary school teachers*. Boston, MA: Houghton Mifflin Company.

Cooney, T. J. (1994). Teacher education as an exercise in adaptation. In D. B. Aichele, & A. F. Coxford (Eds.), *Professional Development for Teachers of Mathematics (1994 Yearbook of the National Council of Teachers of Mathematics)*. Reston, VA: National Council of Teachers of Mathematics.

Dubinsky, E. (1986). Teaching mathematical induction I. *Journal of Mathematical Behavior, 5*, 305–317.

Dubinsky, E. (1990). Teaching mathematical induction II. *Journal of Mathematical Behavior, 8*, 285–304.

Dubinsky, E., & Lewin, P. (1986). Reflective abstraction and mathematics education: The generic decomposition of induction and compactness. *Journal of Mathematical Behavior, 5*, 55–92.

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

Fischbein, E. (1982). Intuition and proof. *For the Learning of Mathematics, 3*(2), 9–18, 24.

Hanna, G. (2000). Proof, explanation and exploration: an overview. *Educational Studies in Mathematics, 44*, 5–23.

Harel, G. (2002). The development of mathematical induction as a proof scheme: A model for DNR-based instruction. In S. Campbell, & R. Zaskis (Eds.), *Learning and teaching number theory: Research in cognition and instruction* (pp. 185–212). New Jersey: Ablex Publishing Corporation.

Healy, L., & Hoyles, C. (2000). Proof conceptions in algebra. *Journal for Research in Mathematics Education, 31*, 396–428.

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

Magnusson, S. (1996). *Building a pedagogical content knowledge base for elementary science teacher education*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, St. Louis, MO.

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, 27*, 194–214.

Marrades, R., & Gutiérrez, A. (2000). Proofs produced by secondary school students learning geometry in a dynamic computer environment. *Educational Studies in Mathematics, 44*, 87–125.

Martin, W. G., & Harel, G. (1989). Proof frames of preservice elementary teachers. *Journal for Research in Mathematics Education, 20*, 41–51.

Movshovitz-Hadar, N. (1993). The false coin problem, mathematical induction and knowledge fragility. *Journal of Mathematical Behavior, 12*, 253–268.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Patton, M. Q. (1990). *Qualitative evaluation and research methods*. Newbury Park, CA: Sage.

Reid, D. A. (2002). Elements in accepting an explanation. *Journal of Mathematical Behavior, 20*, 527–547.

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

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

Simon, M. A., & Blume, G. W. (1996). Justification in the mathematics classroom: A study of prospective elementary teachers. *Journal of Mathematical Behavior, 15*, 3–31.

Smith, D. C. (2000). Content and pedagogical content knowledge for elementary science teacher educators: Knowing our students. *Journal of Science Teacher Education, 11*(1), 27–46.

Sowder, L., & Harel, G. (1998). Types of students' justifications. *The Mathematics Teacher, 91*, 670–675.

Steiner, H. (1989). The nature of theoretical concepts in physics and mathematics-implications for the fragility of knowledge in the educational context. In S. Vinner (Ed.), *Proceedings of the Second International Jerusalem Convention on Education: Science and Mathematics Education – Interaction between Research and Practice* (pp. 387–396). Israel: University of Jerusalem.

Stylianides, A. J., Stylianides, G. J., & Philippou, G. N. (2004). Undergraduate students' understanding of the contraposition equivalence rule in symbolic and verbal contexts. *Educational Studies in Mathematics, 55*, 133–162.

Sullivan, P. (2003). Editorial: Incorporating knowledge of, and beliefs about, mathematics into teacher education. *Journal of Mathematics Teacher Education, 6*, 293–296.