PROSPECTIVE TEACHERS' UNDERSTANDING OF PROOF: WHAT IF THE TRUTH SET OF AN OPEN SENTENCE IS BROADER THAN THAT COVERED BY THE PROOF?

Andreas J. Stylianides*, Gabriel J. Stylianides*, & George N. Philippou**

*University of Michigan, U.S.A., **University of Cyprus, Cyprus

This paper investigates prospective elementary and secondary school teachers' understanding of proof in a case where the truth set of an open sentence is broader than the set covered by a valid proof by mathematical induction. This case breaks the boundaries of students' usual experience with proving tasks. The most important finding is that a significant number of students from both groups who recognized correctly the validity of the purported proof thought that it was not possible for the truth set of the open sentence to include any number outside its domain of discourse covered by the proof. The discussion of student difficulties provides insights into the development of instructional practices in teacher preparation programs aiming to uncover these aspects of students' knowledge fragility and address them accordingly.

Proof is a defining feature of mathematics and, in current school reform recommendations in various countries, is considered a fundamental aspect of instructional programs in all grade levels. However, to have success in the goal to make proof central to all students' mathematical experiences, prospective teachers need to have solid understanding of this mathematical concept. If teacher preparation programs are to develop effective instructional practices that will help prospective teachers cultivate proof in their classrooms, it is essential that these practices be informed by research that illuminates prospective teachers' understanding of proof. Despite the importance of this kind of research, only few studies have investigated in-service or preservice teachers' knowledge of proof (Knuth, 2002; Martin & Harel, 1989; Movshovitz-Hadar, 1993; Simon & Blume, 1996; Stylianides, Stylianides, & Philippou, 2004). Also, these studies have focused more on the logical components of different proof methods than on other important features of the proving process, such as the relationship among the domain of discourse D and truth set U of an open sentence, and a proof that purports to show that the sentence is true in D.

This paper contributes to this research area, focusing on the proof method of mathematical induction. Specifically, we examine what might be some common difficulties that prospective teachers have in dealing with a proof by mathematical induction that is not as encompassing as it could be (D is a proper subset of U). Based on anecdotal evidence that students' normal experience is of being given opportunities to engage in 'universal' proofs (D = U), this study aims to advance the field's understanding of possible issues of knowledge fragility by exposing prospective teachers to a case that falls outside the boundaries of what appears to constitute 'standard practice' for them.

2005. In Chick, H. L. & Vincent, J. L. (Eds.). *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education*, Vol. 4, pp. 241-248. Melbourne: PME.

METHOD

The data for this report are derived from a larger study that aimed to examine the understandings of proof held by the undergraduate seniors of the departments of Education and Mathematics at the University of Cyprus. The participants were 70 education majors (EMs) and 25 mathematics majors (MMs). The EMs, prospective elementary teachers, constituted the 50% of the seniors of the Department of Education during the academic year 2000-01. All of them were taking one particular class to which they were allocated randomly. The MMs, prospective secondary school mathematics teachers, were all the seniors of the Department of Mathematics.

Consider the following statement:

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

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

<u>Proof:</u>

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

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

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

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

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

$> 2^k \bullet 2$ (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.

Figure 1: The test item.

The program of study at the Department of Education includes several mathematics courses that emphasize logical thinking. These courses provide EMs with a fair amount of mathematical knowledge about different types of proof including proof by mathematical induction. The preparation of the MMs focuses primarily on mathematical content and abstract thinking.

All 95 participants responded to a specially designed test that included items on different methods of proof: empirical/inductive proof, proof by counterexample, proof by contradiction, proof by contraposition, proof by the use of computer, and proof by mathematical induction. The data from the test were supplemented by semi-structured interviews with a purposeful sample (Patton, 1990) of 11 subjects (eight EMs and three MMs). The interviews were used to investigate further students' thinking and illuminate patterns that arose from the analysis of the tests.

In this paper we focus only on the test item that appears in Figure 1. This item included a statement and a proposed proof for that statement, and was asking the participants to evaluate the validity of the proof and explain their thinking. The subjects were additionally asked to state whether the sentence (inequality marked with *) in the statement to be proved is true or false in four particular cases: n = 3, 4, 6, and 10. The purported proof is valid; the best response to Part A is choice '2.' The truth set **U** of the inequality is $\{n \mid n \in \mathbf{N}, n \geq 4\}$), but the domain of discourse is taken as $\mathbf{D} = \{n \mid n \in \mathbf{N}, n \geq 5\}$ that does not include n=4. This set up of the test item created a rich context within which we were able to advance our primary goals.

RESULTS

Table 1 summarizes the student responses to Part A of the item. The values represent percentages (rounded to the nearest integer) within major. The vast majority of MMs (92%) said that the proof showed that the statement is always true, while the rest (8%) noted that the proof is invalid. Unfortunately, the students who selected the latter option did not explain their thinking; therefore, we cannot examine further their reasoning. Regarding the responses of the EMs to the same question, approximately half of them (54%) said that the proof showed that the statement is always true, almost one out of three (29%) noted that the proof showed that the statement is true in some cases, 13% considered the proof as invalid, and 4% expressed no opinion.

Of particular interest is the way in which the students justified their responses. 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 in addition raise the issue of whether the students' belief about the validity of the proof was well grounded on reason or not.

> 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 is not the best possible, because it has not proved that the statement is also true for n=6 (since 6 is greater than 5) before proceeding with n=k.

On the other hand, the majority of MMs could justify their choice, even though their arguments were mostly limited to saying that the proposed proof followed correctly the steps of the induction method. The argument of the student MM10 is indicative.

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

Table 1: Percentages for each response option in Part A of the test item by major.

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 $n \geq 5$.

The last comment of MM10 about the interpretation of the word 'always' with respect to the domain of discourse of the statement in the test item lies at the heart of the concept we wanted to investigate and marks a point that caused considerable trouble to students. Specifically, the data suggest that many EMs considered that the proposed proof showed the statement to be true in some cases, because they thought that when we say that a statement is 'always' true we mean that (a) the sentence in the statement is true for all natural numbers (the most commonly met case in high school and even college mathematics), or (b) the sentence is true for all natural numbers that belong to its truth set (in this particular case, $\{n \mid n \in \mathrm{N}, n \geq 4\}$).

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 is 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 proof as invalid.

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 $n=4$.

The student EM20 seems to believe that the validity of a proof by mathematical induction depends on whether or not the proof establishes the truth of the mathematical relationship under consideration on the entire set of natural numbers rather than on the specific set to which the relationship refers. The students EM49 and EM52 rejected the validity of the proposed proof based on opposite reasons. EM49 rejected the proof because he found a value for n (=4) outside the domain of discourse for which the inequality *was* satisfied. He seemed to believe that the proof was invalid because it was not as encompassing as it could be, that is, it did not cover the largest subset of natural numbers for which the inequality was true. EM52 failed to see that the inequality was 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 the set of all natural numbers.

	$n = 3$	$n = 4$	$n = 6$	$n = 10$
Education	83	56	87	91
Mathematics	96	64	100	96

Table 2: Percentages of correct responses to Part C of the test item by major.

Part C of the test item helps investigate further students' understanding of the relation between the domain of discourse of the statement to be proved and the truth set of the inequality. Table 2 summarizes the percentages of correct responses to each of the special cases in the test item by student major (the values are rounded to the nearest integer). The highlight of the table is the failure of many students from both majors to realize that the inequality is true for n=4; the percentages of success were 56% and 64% for EMs and MMs, respectively. Given the simplicity of the calculations required to check the inequality for n=4, it is plausible to assume that the students reached this conclusion based on an erroneous reasoning. This reasoning was most probably associated with the fact that number 4 was not included in the domain of discourse of the statement. The difference in the percentages of success between the first two 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 the domain of discourse of the proved statement would accidentally get the first right and the second wrong. The higher percentages of success for the other special cases, n=6 and n=10, were expected given that many students accepted the validity of the proposed proof and these cases belonged to the domain of discourse of the proved statement. Also, the calculations were not difficult for the students who chose to carry them out.

Table 3 presents a detailed analysis of the results obtained from parts A and C of the test item. In particular, the table provides five response types, each of which corresponds to a different combination of student responses to the two parts.

A significant number of students from both majors, 38 EMs and 23 MMs, recognized the validity of the purported proof, thus responding correctly to Part A of the test item. From these students, only 15 EMs and 13 MMs responded correctly to all four special cases of Part C (Response Type 0). From the same group of students, three

EMs and one MM said that the inequality is true for all special cases (Response Type 1). This response suggests that the students believed 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 the test item (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 of the inequality cannot include natural numbers outside its domain of discourse in the proved statement.

Response Type	Description of Response Types and Frequencies by Major	
	Part A of the test item[a]	Part C of the test item[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. (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. (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)

Table 3: Frequencies of selected student responses to parts A and C of the test item.

[a] The first number in each parenthesis in this column represents the number of EMs who responded the specified way in Part A, and 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 the item that appear in the same row of the table. The first represents the number of EMs and the second the number of MMs.

The remaining two response types are associated only with EMs. Specifically, from the nine EMs who considered the purported proof as invalid, two said that the inequality is false for all four special cases (Response Type 3). These students most likely 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 purported 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 likely thought that the domain of discourse of the inequality in the statement to be proved should be the same with its truth set.

Some of the interviews shed further light on students' thinking regarding the investigation of the special cases. For example, student EM38, whose response in the test belonged to Response Type 0, had difficulty understanding the 'mismatch' between the domain of discourse of the statement to be proved in the test item and the truth set of the inequality. However, after some probing from the interviewers (the first two authors), EM38 appeared to have grasped the relation between these two sets in the context of the given proof ('*I*' denotes the interviewers).

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: Oh… Perhaps we have indeed… Seeing what happens for $n=4$ together with the fact that I considered the statement to be true, 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 be the reason I said 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 the test item] 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: Do you mean to say that proving the inequality for $n \geq 5$ excludes the possibility of the inequality to also hold for smaller values of n?

EM38: Oh… yes. The statement doesn't say '*only* for $n \geq 5$'! Therefore, it leaves open the possibility for other values. Consequently the statement is true.

The student EM38 constantly refers to the correctness or not of the statement rather than to the validity of the purported proof as asked in Part A of the test item. The student seems to believe that there is a strong link between the truth of the statement and the validity of the proof, namely, that the two go together.

DISCUSSION

In this study, we examined prospective teachers' understanding of proof in a case where the truth set of the open sentence in the statement that was to be proved by mathematical induction was broader than the set covered by a valid proof. The analysis of student responses in the test item suggests that a significant number of students of both majors who recognized the validity of the purported proof thought that it was not possible for the truth set **U** of the inequality to include any number outside its domain of discourse **D**. These students incorrectly considered the inequality to be false for $n=4$, and said that the inequality was true only for the two values that belonged to **D** ($n=6$ and $n=10$). The response of a considerable number of EMs that the purported proof shows that the statement is true in some cases seemed to have been influenced by the belief that the domain of discourse of the sentence in the statement to be proved should always coincide with its truth set. These students' observation that the inequality was true for $n=4$ (i.e., $4 \in \mathbf{U}$), coupled with their knowledge of the fact that $4 \notin \mathbf{D}$, seemed to have interfered with their ability to evaluate appropriately the validity of the purported proof.

The results of our analysis highlight difficulties that prospective teachers seem to have in dealing with a proof that is not as encompassing as it could be, thereby uncovering possible aspects of knowledge fragility. The investigation of what might have caused this fragile knowledge requires further research. One possibility is that knowledge fragility has its roots in didactic contracts that possibly prevail in high school and even college mathematics and that promote inappropriately the conception that proofs are always as encompassing as possible.

A related important direction for future research concerns how mathematics teacher educators can organize instruction so that prospective teachers' difficulties in proof surface and become the objects of reflection. Movshovitz-Hadar (1993) suggests one possible way to help prospective teachers reconsider their knowledge, see its problematic aspects, and realize the need for developing a deeper understanding:

> [A]ctivities designed for student teachers should be aimed at accelerating the process of crystallization of their knowledge of particular mathematics notions, such as mathematical induction, by putting them in problem-solving situations which will make them confront their present knowledge and examine it carefully through social interaction with their peers. This process is supposed to reduce the fragility of knowledge. (p. 266)

The test item used in this study has the potential to support the development of learning opportunities that can facilitate the crystallization of prospective teachers' knowledge, as it sets up a situation that breaks the boundaries of what seems to constitute students' normal experience. For example, mathematics teacher educators can use this test item to engage prospective teachers in thinking about whether it is necessary for a valid proof to cover the truth set of an open sentence in its entirety. To manage successfully discussions around issues of this kind, mathematics teacher educators need to be able to anticipate prospective teachers' common conceptual difficulties. A research-based knowledge about these difficulties can support the design of instructional practices aiming to help prospective teachers improve their understanding of proof. Prospective teachers' written and oral responses discussed in this paper can contribute toward this direction.

Author Note

This paper is based on the senior thesis of the first two authors conducted at the University of Cyprus, under the supervision of the third author; the order of authorship of the first two authors is alphabetical.

References

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

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*(3), 253-268.

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

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.

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.