The Mathematics Education into the 21st Century Project

Proceedings of the International Conference
The Decidable and the Undecidable in Mathematics Education
Brno, Czech Republic, September 2003

Does a proof render the truth of a claim decidable?

Nella Yollesh and Tommy Dreyfus *

Tel Aviv University, Israel

Introduction

Assume that you are given a theorem, that is, you are given a mathematical claim together with a proof. Assume that you read, understood and checked the proof and consider it to be correct. Do you think it might be possible to find a special case in which the theorem does not hold? And is it conceivable that a statement contradicting the claim is correct?

Many mathematicians might wonder why someone could or should ask these questions. To them, the very meaning of a theorem having been proved is that it is correct and therefore holds in all cases. This is so, as long as the meaning of "having been proved" is not examined too closely. A closer examination, however, may raise questions. One such question concerns the checkability of proofs. This has become a public issue most recently with respect to the proof of the Poincaré conjecture (Milnor 2003a, 2003b), but also previously in other cases including the following:

- Occasionally gaps or mistakes are discovered in proofs of theorems after their publication (Zahler, 1976). A case in point is Wiles' proof of Fermat's last theorem; it took Wiles and a colleague over a year to bridge the gap discovered by colleagues in late 1993 (van der Poorten, 1993/1995).
- Some proofs are so long that it is not realistic for a single human being to check them. For example, the proof of the classification of simple groups said to be about 15000 pages long.
- Other proofs are based on complex and extensive computer programs. Lam's proof of the non-existence of a projective plane of order 10 is an example. Lam wondered how to deal with the non-zero probability that the computers had made a mistake when carrying out those computations (Lam, 1990).
- A similar but even more acute problem arises for proofs, which make use of biological experiments. While such proofs are not common (yet?), the possibility and rationale for their use has been discussed in the literature (Fallis, 1996).

What then leads practicing mathematicians to believe in the truth of theorems and to use them in their own research? Is it, indeed, the proof?

An even more fundamental concern than checkability has been raised by Velleman (1997): He asks whether the existence of a correct proof for, say, Fermat's Last Theorem implies that Fermat's Last Theorem is correct? Or does it only imply that the theorem is correct modulo some assumption on the consistency of the underlying axiomatic system, which according to Gödel cannot be proved. While Velleman's interest lies with issues in the foundations of mathematics, ours lies with issues in students' conception of the role and status of proof. But the basic question remains the same: Does proof constitute a warrant for the truth of a mathematical statement (Rodd, 2000)?

Background

We leave it up to mathematicians and philosophers of mathematics to deal with this issue as far as it concerns mathematical practice, and focus on the question a formal proof does or does not constitute a warrant for the universal truth of a mathematical statement whether for students of mathematics. Davis and Hersh (1983) and Hanna (1983) first drew attention to the fact that there are no absolute standards for proof. As a consequence, social conventions among the community of mathematicians play an important role for the convincing power of a proof (Thurston, 1994).

In school, such social conventions are usually introduced by the teacher and proofs are most often presented and supported by the teacher. One might thus expect students to have full confidence in the power of these proofs. Research has shown that this is far from obvious. Even when fully aware of a statement and its proof, many of high school and college students

- expressed the need to check special cases of the statement (Fischbein, 1982; Fischbein & Kedem, 1982; Gazit, 1998);
- chose to redo the proof in a special case rather than rely on the general statement (Vinner, 1983);
- ignored, when solving a problem, the main idea of a proof they had themselves proposed immediately prior to solving the problem, and instead reasoned in a manner that contradicted the proof idea (Schoenfeld, 1987).

* Corresponding author: Tommy Dreyfus; e-mail address: tommyd@post.tau.ac.il

Many of the students who participated in the above research studies were not aware of the implications of the universality of formally proved mathematical statements, or at least they did not make optimal use of this universality, and this in spite of the fact that they seem to have understood and been convinced by the proofs. In other words, many students appear to think that the answers to the questions in the opening paragraph of this paper are affirmative, or at least they act in a way compatible with affirmative answers to these questions:

a. Do you think it might be possible to find a specific case in which the theorem does not hold?

b. Is it conceivable that a statement contradicting the claim is correct?

We assume that the foundations of mathematics are not at issue for the students. The length or complexity of a proof may well be but there are methodological tools to overcome this concern, namely to check whether the students understand and accept the claim and proof and to eliminate students who don't. Concerning the remaining students, a range of research questions arises:

1. To what extent do they give affirmative answers to a and/or b?
2. For what reasons do they give affirmative answers to a and/or b?
3. Are they consistent, i.e. if they give an affirmative answer to a do they also give an affirmative answer to b and vice versa?
4. If they are not consistent, how can their inconsistency be explained?
5. How are the answers to the above questions influenced by
   a. The students' age (or mathematical maturity as measured by age)?
   b. The students' achievements in mathematics?
   c. The students' familiarity with the statement?
   d. The content domain (e.g., algebra versus geometry)?
   e. The type of proof (e.g., by contradiction, mathematical induction)?

Methodology

In the present paper, we report on an investigation designed to partially answer some of the above research questions. Specifically, we worked with high school students taking intermediate level mathematics; we focused on the content domain of Euclidean geometry, and addressed questions 1, 2, 5a and 5c and, to a lesser extent, question 3. We designed a questionnaire with the aim of answering the more quantitative aspects of these questions. However, since we expected that it would be difficult to obtain differentiated answers to Question 2 on the basis of the questionnaire, we decided to also interview some of the students who had answered the questionnaire.

Population In view of question 5a, we were interested in comparable student populations from two different age groups. In Israel, two such populations studying geometry at a similar level are available: High school students take Euclidean geometry in grade 9 and return to the topic in grade 12, in preparation of the matriculation examination. In grade 9, school attendance is compulsory and mathematics is a compulsory topic. Students are streamed, with about 40% of the population attending the high-level A-stream. Students studying in the B- and C-streams are not expected to be very familiar with the notion of proof; moreover, many of them do not go on to attend grade 12 classes preparing for matriculation. We thus chose A-stream students as the grade 9 population. In grade 12, about 60% of the population attend classes preparing for matriculation. All students in these schools take mathematics, at either the 3 (about 35%), 4 (about 15%) or 5 units level (about 10%). We have chosen 4-unit students as the grade 12 population that is most closely comparable to the grade 9 A-stream population. The population to whom the questionnaire was administered consisted of 69 grade 9 A-stream students and 67 grade 12 four-unit students. Nine of these students (four grade 9 students and five grade 12 students) were later interviewed. They were chosen on the basis of their different answering patterns to the questionnaire questions.

Questionnaire A questionnaire was prepared with the purpose of answering research questions 1, 5a, and 5c and, to a lesser extent, questions 2 and 3. The questionnaire dealt with a familiar as well as an unfamiliar geometry theorem at each grade level. The familiar theorem was the same at both grade levels, namely that an external angle of a triangle equals the sum of the two non-adjacent interior angles. This is a standard theorem that was discussed, proved and repeatedly used in class during the geometry course. The unfamiliar theorem was one that could have been included in the curriculum but was not. For example, in grade 9, it was that the quadrilateral formed by the midpoints of the edges of an arbitrary quadrilateral is a parallelogram. The questionnaire contained two parts. Part I dealt with the familiar theorem and Part II with the unfamiliar one. Each part contained two sections. In Section 1, after a brief set of instructions, the theorem was presented, accompanied by a figure, a proof and three questions:

1a. To what extent do you agree with the statement "the theorem is true"?

1b.To what extent do you agree with the statement "the proof of the theorem is convincing"?

1c. Dan believes that one needs to check a variety of triangles in order to be sure that the theorem is true. Dana claims that there is no need for further checks because the proof is convincing. With whom do you agree?

There were four levels of agreement for Questions 1a and 1b (agree / tend to agree / tend not to agree / don't agree / don't know) and four possible answers to Question 1c (Dan is right / Dana is right / both are right / both are wrong). Furthermore, students were given seven empty lines in which to explain their choice of answer to Question 1c.

Section 2 included a statement clearly contradicting the theorem given in Section 1. For example, following the familiar theorem, the contradictory statement was: "In an obtuse triangle, one of the external angles is acute and smaller than the sum of the two non-adjacent interior angles." This statement was accompanied by a figure showing the external angle α and the internal angles β and γ, and the inequality $\alpha < \beta + \gamma$. In Question 2, the students were asked to express their level of agreement with the contradictory statement (the same answer options as above were offered) and to explain their choice. Half an empty page was provided for this explanation. No lines were used since they might have prevented students from making drawings. Part II of the questionnaire was formulated in precisely the same manner as part I, except that it dealt with the unfamiliar theorem.

Interview

The goal of the interview was to make progress toward answering research question 2 and, to a lesser extent, 5c. One aim of the interviews was thus to find out for what reasons students who had accepted the statement and the proof of one of the given theorems tended to agree with Dan that a variety of cases needed to be checked in order to be sure that the theorem was true. Similarly, we aimed to find out for what reasons students who had accepted the statement and the proof of one of the given theorems tended to accept that a statement contradicting the theorem was true. These phenomena can be considered as an expression of the lack of a proper universality conception for mathematical theorems on the part of the students. We designed the interview so as to identify components of this lack of a proper universality conception. In addition to explicit statements by the students pointing to the lack of awareness that a statement and its negation cannot hold at the same time, or a lack of awareness that a convincing proof confers universal validity on a statement, we expected to possibly find signs of the influence of experience from daily life or from the less exact sciences, as well as signs of a psychological need on the part of the students to intuitively confirm any general claim or to check everything several times. Moreover, the nature of mathematics as it transpires in mathematics classrooms, could conceivably have an influence; for example, mathematics instruction stressing standard answers to standard questions could lead students to believe that every statement needs it own explicit technical proof, even if its validity can be inferred from the proof of a previous statement.

The interview was thus carefully designed to include opportunities for the students to express opinions, conceptions, needs and feelings about the role and meaning of proof. It included question such as

- "Can you explain the internal logic of what you wrote?"
- "What does it mean that the proof convinced you, if you say you need to check more?"
- "Since you feel you need to check more, maybe you are not convinced by the proof?"
- "Did you notice the relationship between the previous statement and the present one?"
- "Do you think there could be a case where two contradictory statements could both be true?"

Some results

At the time this paper is being written, the analysis of the data is still in progress. We can therefore present only some of the quantitative data in this written version. More complete results including an analysis of the students' explanations in the questionnaire as well in the interviews will be presented at the conference.

After eliminating students who answered Questions 1a and/or 1b negatively, because they were not convinced either by the statement or by the proof, the remaining population consisted of 62 grade 9 students and 62 grade 12 students. The following tables present the percentages of these student groups who answered the main questions according to a proper universality conception; in other words, those who answered that there is no need to check specific cases in order to be sure that the theorem is true (Question 1c) and those who disagreed with the claim that the contradictory statement was true (Question 2). The results are presented separately for the familiar and for the unfamiliar statements.

Q1c	Familiar	Unfamiliar
Grade 9	52%	27%
Grade 12	68%	52%

Q2	Familiar	Unfamiliar
Grade 9	57%	24%
Grade 12	61%	40%

Tables: Percentages of students answering Questions 1c and 2 according to universality conception

Not surprisingly, consistently more students behaved according to a proper universality conception on the familiar statement than on the unfamiliar one (research question 5c) and more grade 12 students than grade 9 students behaved according to a proper universality conception (research question 5a). It thus appears that the students' general mathematical maturity is related to their conception of the universality conferred upon a statement by a proof, even though the topics taught in grades 10 and 11 are less explicitly proof oriented than Euclidean geometry. This apparent relationship is qualified, however, by the observation that the number of students who were fully consistent in applying universality (that is answered according to universality on all four questions) remains very low (9 students, 15%) even in grade 12. There were 5 such students (8%) in grade 9 (research question 3).

With one exception (Q2, Grade 9, unfamiliar statement) fewer than 5% of the students chose the evasive answer ("both are right" in Q1c; "I don't know" in Q2). Thus, the data show that, depending on question and grade level, at least 25%, and up to 70% of the students did not behave according to a proper universality conception (research question 1). Even these numbers may be too low, though, because they do not include those students (and there were many) who started by trying to find a proof for the contradictory statement and only marked that statement wrong after failing to find a proof.

Conclusion

Clearly, many high school students, even among those learning mathematics at an advanced level, do not have a proper conception of the universality conferred upon a mathematical statement by a proof. We hope that the data from the students' explanations and interviews will show whether they "simply fail to understand the meaning of proof" (Fischbein, 1982), or what other reasons they have for refraining from taking proof as warrant for a mathematical statement.

In terms of the conference theme, our investigation can be interpreted as follows: Does the existence of a correct proof render the truth of a claim decidable? And does the truth of a theorem imply that the truth of a statement contradicting the theorem is decidable? In other words, does a correct proof make a theorem universally valid? In the eyes of many students, apparently not.

References

Davis P. J. & Hersh R. (1983). *The Mathematical Experience*. Boston, MA, USA: Houghton Mifflin.

Fallis D. (1996). Mathematical Proof and the Reliability of DNA Evidence. *Am. Math. Monthly* 103 (6), 491-497.

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

Fischbein E. & Kedem I. (1982). Proof and certitude in the development of mathematical thinking. In A. Vermandel (ed.), *Proceedings of the Sixth International Conference on the Psychology of Mathematics Education* (pp. 128-131). Antwerp, Belgium: Universitaire Instelling.

Gazit A. (1998). *Consistency in the conception of the property of universal validity that formal proof confers upon a mathematical statement*. Unpublished MA thesis. Tel Aviv, Israel: Tel Aviv University.

Hanna G. (1983). *Rigorous Proof in Mathematics Education*. Toronto, Canada: OISE Press.

Lam C. W. H. (1990). How reliable is a computer-based proof? *The Mathematical Intelligencer* 12 (1), 8-12.

Milnor J. W. (2003a). The Poincaré conjecture 99 years later: A progress report (February 2003). Available online http://www.math.sunysb.edu/~jack/PREPR/poiproof.pdf

Milnor J. W. (2003b). Towards the Poincaré conjecture and the classification of 3-manifolds (version of 6-14-03). Available online http://www.math.sunysb.edu/~jack/PREPR/tpc.pdf

Rodd M. M. (2000). On mathematical warrants: Proof does not always warrant, and a warrant may be other than proof. *Mathematical Thinking and Learning* 2 (3), 221-244.

Schoenfeld A. H. (1987). What's all the fuss about metacognition? In A. H. Schoenfeld (ed.), *Cognitive Science and Mathematics Education* (pp. 189-216). Hillsdale, NJ, USA: Lawrence Erlbaum Associates.

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

Van der Poorten A. (1993/1995). Remarks on Fermat's last theorem. Available online http://www-centre.mpce.mq.edu.au/alfpapers/a129.pdf

Velleman D. (1997). Fermat's last theorem and Hilbert's program, *The Mathematical Intelligencer* 19 (1), 64-67.

Vinner S. (1983). The notion of proof – some aspects of students' views at the senior high level. In R. Hershkowitz (ed.), *Proceedings of the Seventh International Conference on the Psychology of Mathematics Education* (pp. 289-294). Rehovot, Israel: The Weizmann Institute of Science.
Zahler R. (1976). Errors in Mathematical Proofs, *Science* 193, 98.