How major mathematics students evaluate the mathematical proof

Abstract: The article presents results of the study on the ability to evaluate mathematical proofs by students being prepared for the profession of mathematics teacher (up to, and including, the high school level). A group of 74 students (from Poland and Slovakia) were presented two proofs of the same theorem to assess their validity (proofs were obtained from other mathematics students). Analysis of the materials obtained in the study was aimed at identifying and characterizing of difficulties in understanding mathematical proof and of symptoms. Symptoms of the subjects’ comprehension of the mathematical proof essence were also being considered. One of the particular objective of our analysis refers to the question on understanding the link between a theorem and its proof.

Introduction and background

Over the last 25 years, numerous studies were conducted, both theoretical and empirical ones, related to explanation, justification, proof in learning and teaching mathematics at various stages of education. Increased interest in this issue is associated with the fact that the terms explanation, justification, proof are explicit in the documents determining objectives and content of mathematics taught at schools in various countries around the world (as demonstrated in: Balacheff, 1991; Hanna, 2000; Weber, 2010; Žeromska, 2013). The result of theoretical studies in this respect is, among others, identification of the

Key words: teachers training, mathematical proof, evaluation of proof, mathematical activity, mathematical false beliefs.
aspects and functions of proof in mathematics (e.g. Balacheff, 1987, 1991; de Villiers, 1990; Hanna, 2000; Raman, 2003; Yackel, 2004; Hanna, de Villiers, 2008). Other studies refer to mathematical proof in the wider context of mathematical methodology (e.g., Krygowska, 1977; Hersh, 1993; Davies, Hersh, 1994; Rav, 1999; Turnau, 2001; Tall, 2002; Avigad, 2006; Solomon, 2006). Terminology framework for the creation and evaluation of proofs is established. This framework is created from two different perspectives. One of these perspectives emphasizes mathematical aspect of the problem (e.g. Tulmin, 1958; Stylianides, Stylianides, 2009), the other one – characterizes individual subjective view of proof (e.g. Martin, Harel, 1989; Harel, Sowder, 1998, 2007; Raman, 2003).

Empirical research are largely focused on understanding of the concept of proof by students of middle schools and high schools, university students of mathematics and in-service teachers of mathematics. Most generally speaking, the results of these studies demonstrate that students’ and teachers’ views of mathematical proof are frequently different from contemporary mathematicians’ views (Weber, 2010). Let us refresh how deductive proof is characterized in the didactic literature related to the subject matter. For example, let us quote two such definitions. Krygowska (1959, 1977) states: “Deductive proof involves drawing logical conclusions from assumptions, axioms and previously proven assertions. Each step of proof consists of formulating the necessary condition of the preceding premise” (Krygowska, 1959, p. 100). Similarly, in the USA Principles and Standards for School Mathematics (NCTM, 1989, p. 144), it is said that mathematical proof consists of “careful sequence of steps with each step following logically from an assumed or previously proved statement and from previous steps”. The difference in understanding of proof between mathematics learners and mathematicians is analyzed e.g. by Harel, Sowder (1998, 2007) by introducing the term proof schemes with psychological meaning, individual-centered. Proof scheme is explained as the way in which a person attempts to convince themselves or persuade others about the truth of a mathematical theorem. The taxonomy of proof schemes by Harel, Sowder consists of three classes: [a] external conviction proof schemes class (e.g. depends on the authority: teacher, book), [b] empirical proof schemes class (evidence from examples or perceptions) and [c] deductive proof schemes class (essential characteristics: generality, operational thought, logical inference). The research revealed that students often held empirical proof schemes. Also, they frequently did not know what features of the argument should have been acceptable in mathematics.

The studies which aim to understand how the concept of proof is understood by university students of mathematics – potential school teachers, and also
How major mathematics students evaluate in-service teachers of mathematics, is of particular interest to the authors of this paper (e.g. Martin, Harel, 1989; Harel, Sowder, 1998, 2007; Dreyfus, 1999; Gardiner, 1999; Jones, 2000; Recio, Godino, 2001; Knuth, 2002; Raman, 2003; Selden, Selden, 2003; Inglis et al., 2007; Pieprzyk, Żeromska, 2009; Tsamir et al., 2009, 2010; Weber, 2010; Weber et al., 2014; Boyle et al., 2015). It is clear that correct understanding of proof by mathematics teachers is crucial to the proper development of this concept at school. Empirical studies conducted so far provide certain insight into the question of how future and currently practicing teachers understand the concept of mathematical proof. These studies are diversified according to the type of activity of the participants induced by a specific research tool which was used. Several research studies on this subject were focused on the ability of mathematics majors students and also mathematics teachers to evaluate given statements: true or false (e.g. Martin, Harel, 1989; Gardiner, 1999; Selden, Selden, 2003; Tsamir et al., 2009; Weber, 2010). Another type of a research task involved autonomous construction of a proof of a given theorem by the subjects under examination (Recio, Godino, 2001; Sowder, Harel, 2003; Furinghetti, Morselli, 2009). In addition to the research focused on one of the aforementioned skills, there were also studies where the research task included a combination of two skills: evaluating a statement and evaluating an argument (e.g. Stylianides, Stylianides, 2009).

The results of the research conducted on the subject under consideration are analyzed from different perspectives. Inglis et al. (2007) distinguish two types of analyses: those that concentrate on the argument’s content and those that concentrate on the argument’s structure. In the analyses focused on the argument’s structure, Toulmin’s model of general argument is used (Inglis et al., 2007; Pedemonte, 2007), taking into account basic types of statements, each playing a different role in an argument.

Studies of the above-mentioned authors have proven that pre-service and in-service mathematics teachers have various difficulties in understanding the essence of mathematical proof. They are demonstrated, among others, by:

1. Doubts when making decisions whether a given argument can be regarded as mathematical proof (Gardiner, 1999).
2. Recognizing the empirical argument as sufficient to justify the truth of the general statement (Martin, Harel, 1989; Vinner, 2012).
3. Accepting invalid general argument as a proof and acknowledging correct explanation as erroneous one (e.g. Stylianides, Stylianides, 2009).
4. Failure to understand the relationship: the truth of the theorem results from the validity of its proof (Knuth, 2002; Pieprzyk, Żeromska, 2009).
5. Lack of awareness of the adopted reasoning basis in deductive reasoning (e.g. Gardiner, 1999).

6. Recognizing the validity of a proof and simultaneous verification of the truth of the theorem on examples – failure to understand generality of the theorem and the role of deduction in justification of mathematical statements (Knuth, 2002).

7. Seeking a counter-example despite acknowledging the truth of the theorem (Knuth, 2002).

8. Assessing the validity of a proof through a prism of familiar rituals related to the record, form, but not to the mathematical content (Martin, Harel, 1989).

It is worth mentioning one more difficulty associated with proving in mathematics: failure of a learner to distinguish between two different aspects of this activity:

(a) inner conviction of an individual as to the correctness of a given statement (which may be, for example, the result of verification based on numerous examples);

(b) conducting deductive reasoning as confirmation of the truth based on the previously acquired knowledge.

Long-term observations of the learning process of mathematics, as well as the results of some studies (e.g. Raman, 2003; Ciosek, Turnau, 2004, 2015; Ciosek, 2005; Harel, Sowder, 2007; Weber, 2010), prove that such a difficulty really exists.

This paper presents results of the authors’ own didactic research study on understanding of mathematical proof by prospective teachers of mathematics. The following terms used in this article: reasoning, deductive reasoning, proof, and adopted deduction base, shall be understood as defined below. “Reasoning is the line of thought, the way of thinking, adopted to produce assertions and reach conclusions. It is not necessarily based on formal deductive logic, and my even be incorrect as long as there are some kind of sensible (to the reasoner) reasons behind the thinking” (Lithner, 2006, p. 4). The author of the above definition uses the term reasoning in a much broader sense than deductive reasoning. The latter type of reasoning is characterized by the authors of the paper (Weber et al., 2014) and they recognize that mathematical proof is a kind of deductive reasoning: When we refer to deductive reasoning, we are referring to arguments where each new assertion is a necessary logical consequence of previous assertions, not where each statement is derived by the
mechanical application of explicit logical rule” (Weber et al., 2014, p. 3) “(...) A proof is a deductive argument that claims to show a conclusion is a logically necessary consequence from agreed upon assumption” (Weber, 2014, p. 4). According to Krygowska (1977), the adopted deduction base is understood as those elements of the previously asserted mathematical knowledge (axioms, logical laws, definitions, theorems, properties, etc.), which apply to justifying the various inferences in a proof.

1 Methodology of the Research

1.1 Objective

The research objective – in general terms – was to examine understanding of mathematical proof by participants of the research with particular emphasis on insight into the nature of students’ justifications (adopted deduction base) of inferences occurring in a given argument. In addition, the task for students was selected so as to provide at least a partial answer to the question: How do participants understand the link between the theorem and its proof?

The analysis of the material provided by our research focused on identification and characterization of the symptoms of understanding of mathematical proof and also on the difficulties in this respect. Detailed study objectives include:

• Identifying behaviors and providing description of mental activities performed by students to answer the question: correct reasoning or incorrect reasoning.
• Presenting symptoms of understanding mathematical proof as a logical sequence of inferences (it is about whether a participant acknowledges a need to justify validity of individual inferences by specifying the adopted deduction base).
• Determining nature of the argumentation brought up by the participants to justify individual inferences in proof reasoning.
• Revealing how students understand the link between an theorem and its proof.

1.2 Participants

The study was conducted in the years 2014-2015. Students of mathematics of the senior year of university studies (prepared for teaching in junior and senior high school) of one of the universities in one of the largest Polish cities, as well
as students of one of the major universities in Slovakia, were subjected to the research. A total of 74 participants were examined.

1.3 Procedure

Students received a text containing two authentic examples of reasoning of other students (S₁ and S₂), recognized by their authors as proofs of a certain theorem. The participants in the study were asked to evaluate these proofs and to give reasons for their assessment. Two methods were used in the research: the paper and pencil instruments (58 participants) and the method of “think aloud” – interviews (16 participants).

1.4 Instruments

The task material used in the study included two authentic examples of reasoning (of mathematics students: S₁ and S₂) obtained by conducting previous studies of the process of solving open mathematical problems (Ciosek, 2005). In the first round of the research, thirteen participants were given only one of the mentioned examples of reasoning (reasoning by S₁ – see below) to evaluate. In the next round of the research, the participants were given both examples at the same time to evaluate.

Task for students:

Read the following two proofs of the theorem [T], and then evaluate them providing your justification.

[T] If \(a, b, c, d\) belong to the interval \((0, 1)\), then holds the inequality

\[
(1 - a)(1 - b)(1 - c)(1 - d) > 1 - a - b - c - d
\]

Argument provided by S₁:

\[
(1 - a)(1 - b)(1 - c)(1 - d) = [1 + ab - (a + b)][1 + cd - (c + d)] > [1 - (a + b)][1 - (c + d)] = 1 + (a + b)(c + d) - (a + b + c + d) > 1 - (a + b + c + d).
\]

The thesis of the theorem follows from the performed calculation.

Argument provided by S₂:

The following inequalities hold true:

\[\text{The task was used previously in A. Schoenfeld’s studies (1979); students were asked to prove the theorem [T].}\]
How major mathematics students evaluate

\[
1 - a > 1 - a - b - c - d, \\
1 - b > 1 - a - b - c - d, \\
1 - c > 1 - a - b - c - d, \\
1 - d > 1 - a - b - c - d.
\]

It follows there from that:
\[
(1 - a)(1 - b)(1 - c)(1 - d) > 1 - a - b - c - d.
\]

The above reasoning were presented in the original version of their authors, who did not explicitly state the reasons for the correctness of individual inferences.

Both examples of reasoning are incorrect, each for a different reason. Reasoning by S\(_1\) contains a substantial error that results from an incorrect justification for one of the inferences. Reasoning by S\(_2\) does not contain any substantive error (all the records in this reasoning are objectively correct), but the conclusion does not follow logically from the previously written conditions. Incorrectness of reasoning of the student S\(_1\) can thus be demonstrated by providing an appropriate counter-example for inequality C\(_2\) (see below) whereas it is not possible in evaluation of reasoning by S\(_2\).

1.4.1 Remarks on argument by S\(_1\)

This reasoning can be briefly summarized in the following four steps, denoted as C\(_1\), C\(_2\), C\(_3\), C\(_4\), respectively

C\(_1\): \((1 - a)(1 - b)(1 - c)(1 - d) = [1 + ab - (a + b)][1 + cd - (c + d)]\)

C\(_2\): \([1 + ab - (a + b)][1 + cd - (c + d)] > [1 - (a + b)][1 - (c + d)]\)

C\(_3\): \([1 - (a + b)][1 - (c + d)] = 1 + (a + b)(c + d) - (a + b + c + d)\)

C\(_4\): \(1 + (a + b)(c + d) - (a + b + c + d) > 1 - (a + b + c + d)\).

The equality C\(_1\) is correct. The expression \([1 + ab - (a + b)]\) is a result of multiplication of binomials \((1 - a)(1 - b)\), and the expression \([1 + cd - (c + d)]\) is the result of multiplication of \((1 - c)(1 - d)\).

The inequality C\(_2\) is false. Of course, it is true that the value of the expression \(1 + ab - (a + b)\) is greater than the value of \(1 - (a + b)\), as \(a, b\) are positive. Similarly, the value of the expression \(1 + cd - (c + d)\) is greater than the value of the expression \(1 - (c + d)\) for each \(c, d\) satisfying the assumptions of the theorem. However, the product on the left side of the inequality C\(_2\) does not need to take values greater than the product on the right side of the inequality. And indeed, for example for \(a = \frac{1}{2}, b = \frac{2}{3}, c = \frac{3}{4}, d = \frac{4}{5}\), the value of the expression on the left side of the inequality is less than the value of the expression on its...
right side. Student $S_1$ probably thought that for all $a, b, c, d \in (0, 1)$, the value of the expression on the left side of the inequality is greater than the value of the expression on the right side of this inequality.

The equality $C_3$ is correct. Probably, the right side of this equality derived from the rule “we multiply each word of one expression by each word of the second expression” and then partial results were ordered in memory.

The inequality $C_4$ is also correct. Student $S_2$ noticed that the right side of the inequality from the thesis of the theorem can be obtained by leaving out the positive component $(a + b)(c + d)$ of the algebraic sum $1 + (a + b)(c + d) - (a + b + c + d)$, which results in the expression with a smaller value. Ultimately, it should be stated that argument by $S_1$ is false. The essence of error in reasoning by $S_1$ can be uncovered as a result of analysis of inequality $C_2$:

\[
(1 + ab - (a + b))[1 + cd - (c + d)] > [1 - (a + b)][1 - (c + d)].
\]

Its shape indicates that $S_1$ probably holds the following false belief: Reducing factors decreases the product. The above rule applies when all four factors in $C_2$ are of positive values but if factors on the right side are of negative values, the mentioned rule can not be fulfilled. This false belief probably derives from many years of the student’s experience in the arithmetic of natural numbers (and then positive real numbers). It remained in the student’s mind so strongly that it is transferred (incorrectly) to the wider “world of real numbers”.

### 1.4.2 Remarks on argument by $S_2$

Student $S_2$ wrote four real inequalities, denoted as follows:

- **Claim 1**: $1 - a > 1 - a - b - c - d$,
- **Claim 2**: $1 - b > 1 - a - b - c - d$,
- **Claim 3**: $1 - c > 1 - a - b - c - d$,
- **Claim 4**: $1 - d > 1 - a - b - c - d$.

Each of these inequalities was probably created by $S_2$ as a result of paying attention to the form of inequality from the thesis of the theorem $[T]$. This student considered it obvious that each factor of the product from the conjecture of the theorem takes a value greater than the value of the expression on the right side of the analyzed inequality. Indeed, this condition is satisfied. The student then wrote down the final inequality (denoted as Claim 5):

**Claim 5**: $(1 - a)(1 - b)(1 - c)(1 - d) > 1 - a - b - c - d$,

probably judging by the fact that since each of the four numbers is greater than a certain number $m$, thus the product of those numbers is also greater than the number $m$. It can not be ruled out that the student $S_2$ thought: “I know
that $1 - a > 1 - a - b - c - d$. If the left side of this inequality is multiplied by $(1 - b)$, then I obtain more than $(1 - a)$. Then, $(1 - a)(1 - b) > 1 - a - b - c - d$. I know that the expression $(1 - a)(1 - b)(1 - c)$ takes even higher values than $(1 - a)(1 - b)$, which means that it will be even greater than $1 - a - b - c - d$. But $(1 - a)(1 - b)(1 - c)(1 - d)$ is greater than $(1 - a)(1 - b)(1 - c)$; it means that the final inequality $(1 - a)(1 - b)(1 - c)(1 - d) > 1 - a - b - c - d$ is correct”.

This kind of inference would be correct if each factor was a number greater than 1. In this analyzed example, this is not the case though; factors of the product are the numbers from the interval $(0, 1)$, so the result of multiplying each of them by the second one is less than each factor.

Analysis by student S2 can be interpreted so that he holds a false belief: Multiplication makes bigger. The existence of such a false belief has already been revealed in previous empirical studies (e.g. Fischbein et al., 1985; Bell, 1988; Graeber, Tirosch, 1989; Ciosek, Turnau, 2015). In Fischbein’s opinion, this false belief is intuitive in its nature: “[The] primitive model associated with multiplication is repeated addition, in which a number of a collections of things of the same size are put together. [...] The initial didactic models seemed to become so deeply rooted in the learner’s mind that they continue to exert an unconscious control over mental behavior even after the learner has acquired formal mathematical notions that are solid and correct” (Fischbein, 1985, p. 6, 16).

2 Results and Analysis

2.1 Analysis of students’ procedure in validation of reasoning by student S1

A group of 74 participants assessed reasoning by student S1 (sixteen of whom were subjected to individual observation). Seven participants did not submit any assessment of the correctness of reasoning by S1; four of them did not take any action, and the other three tried to construct their own proof of the theorem [T].

Below, there is a description of other 67 students’ performance:

1. Students started by verifying the calculations carried out by S1 with respect to the equality C1. They multiplied $(1 - a)(1 - b)$ and $(1 - c)(1 - d)$, and compared their result with the result recorded by S1. Some wrote down their computation on a piece of paper, others only evaluated visually whether the results of their mental calculus were the same as those of reasoning by S1.
2. Then, after positive verification of the equality $C_1$, the students checked whether the inequality $C_2$ held. First, they compared the product factors on the left and right sides of the inequality. They noticed that the value of the expression $[1 - (a + b)]$ was less than the value of the expression $[(1 + ab - (a + b))]$, and similarly, the value of the expression $[1 - (c + d)]$ was less than the value of the expression $[(1 + cd - (c + d))]$. After this observation, several respondents considered that the product on the right side of the inequality $C_2$ was less than the product on the left of this inequality. It seemed so obvious that they failed to record why the inequality $C_2$ held. Only some of them noted, for example: “Student $S_1$ reduces the factors, so the product is also reduced”.

Several participants expressed their doubts as to the correctness of the inequality $C_2$. These doubts were settled in various ways, and with different results. Two participants, after empirical examination (on one numerical example) if the examined inequality held, found that: “The inequality $C_2$ is OK”. Other students were trying to find a counterexample.

Others who expressed their doubts as to the inequality $C_2$ carried out a certain algebraic calculus and eventually found that the inequality was satisfied.

3. Having recognized the inequality $C_2$ as true or false, students went on to the validation of $C_3$ and $C_4$ in the reasoning by student $S_1$. Both of them were considered correct.

Numerical distribution of the students’ final answers to the question of correctness of reasoning by student $S_1$ is presented in the following table (Tab. 1).

<table>
<thead>
<tr>
<th>Assessment of reasoning by student $S_1$</th>
<th>60</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reasoning by $S_1$ is correct</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reasoning by $S_1$ is incorrect</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No validation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Table 1. Numerical distribution of the answers regarding reasoning by student $S_1$. |

2.1.1 Methods for justifying the opinion “Reasoning by student $S_1$ is correct”

Recognizing the inequality $C_2$ as true or false in fact determined the evaluation of the whole reasoning to be “correct” or “incorrect”. Below there is a qualitative analysis of the participants’ opinions, from the viewpoint their justification of the assessment of validity of the inequality $C_2$. 
Three different types of attempts to justify the statement: The inequality $C_2$ is true can be distinguished:

1. Applying (false) multiplication property
   The participants considered the following statement to be true: *If we reduce the factors, we also reduce the product.* They used it as justification for the validity of the inequality $C_2$.

2. Examining signs of numerical values of the expressions on both sides of the inequality $C_2$
   The participants dealt with algebraic expressions on both sides of the inequality and tried to answer what numbers (positive or negative) may be the values of these expressions.

3. Verification on examples
   The participants selected the example of the numbers $a, b, c, d$ that satisfied the assumption of the theorem and, using the calculation, verified whether the inequality was satisfied.

Numerical distribution of the individual types of justification is presented in the following table (Tab. 2).

<table>
<thead>
<tr>
<th>Justification of the opinion “the inequality $C_2$ is true” by:</th>
<th>55</th>
<th>3</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applying (false) multiplication property</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Examining signs of numerical values of the expressions on both sides of the inequality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verification on examples</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Numerical distribution of types of justification for the truth of the inequality $C_2$.

What students wrote to justify that inequality $C_2$ is true, is shown in the following examples.

**Example 1** (applying false multiplication property):

$$[1 + ab - (a + b)][1 + cd - (c + d)] > [1 - (a + b)][1 - (c + d)].$$

*This inequality is true: $S_1$ reduced the factors so he reduced the product because of the assumption $a, b, c, d \in (0, 1)$.*

**Example 2** (examining signs of numerical values of the expressions on both sides of the inequality):

$$[1 + ab - (a + b)][1 + cd - (c + d)] > [1 - (a + b)][1 - (c + d)].$$

*The left side of this inequality always takes positive values. But right side as equal to $1 + (a + b)(c + d) - (a + b + c + d)$, can take negative values.*
Comment: The student expressed the following view: The expression that can take negative or zero values is always less than the expression that takes only positive values.

Example 3 (verification on examples):

Let us check if the inequality is true for: \(a = \frac{1}{2}, b = \frac{1}{4}, c = \frac{1}{2}, d = \frac{1}{4}\).

\[
\begin{align*}
[1 + \frac{1}{8} - (\frac{1}{2} + \frac{1}{4})][1 + \frac{1}{8} - (\frac{1}{2} + \frac{1}{4})] &= \frac{3}{8} \cdot \frac{3}{8} = \frac{9}{64}, \\
[1 - (\frac{1}{2} + \frac{1}{4})][1 - (\frac{1}{2} + \frac{1}{4})] &= \frac{1}{16}.
\end{align*}
\]

It is evident that \(\frac{9}{64} > \frac{1}{16}\), so the inequality is true.

2.1.2 Ways of justifying the opinion “Reasoning by student S_1 is incorrect”

As regards recognizing the inequality \(C_2\) as incorrect, one person failed to provide any justification for his/her opinion. Others tried to explain why the analyzed inequality was incorrect. Two different methods were distinguished:

I. Using a well-known way of comparing algebraic expressions (to evaluate the inequality such as \(A > B\), where \(A\) and \(B\) are algebraic expressions, it would beenough to present \(A\) in the form \(B + X\), where the expression \(X\) always takes positive values).

II. Analyzing signs of numerical values of the expressions in the inequality \(C_2\) and providing a counter-example.

Frequency of using those two method by participants is shown in the Table 3.

<table>
<thead>
<tr>
<th>Justification of the opinion “the inequality (C_2) is false” by:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Using a well-known method of comparing algebraic expressions</td>
<td>Analyzing signs of numerical values of factors on the right side of the inequality and providing a counter-example</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 3. Numerical distribution of methods of justification that the inequality \(C_2\) is false.

Below two examples of the students’ justification of the opinion “the inequality \(C_2\) is false” with using method I. or method II. are presented.
Example 4 (using a well-known way of comparing algebraic expressions):

Student $S_1$ can omit $ab$ from the first bracket and also $cd$ from the second one, provided that each of obtained numbers are positive. But we have:

$$[1 + ab - (a + b)][1 + cd - (c + d)] = 1 + cd - c - d + ab + abcd - abc +$$

$$-abd - a - acd + ac + ad - b - bcd + bc + bd,$$

and

$$[1 - (a + b)][1 - (c + d)] = 1 - a - b - c - d + ac + ad + bc + bd.$$

In order the inequality was true the value of the expression: $cd + ab + abcd - abc - abd - acd - bcd$ should be positive. However, we do not know if it is the case, so $S_1$ can not get the solution in this way.

Comment: The student transformed the left side of the inequality to the form $B + X$, where $B$ is equal to the right side of this inequality. Then he noticed that the expression $X$ does not need to take only positive values.

Example 5 (providing a counter-example):

Looking at the inequality $C_2$ student said: The inequality is given by the following symbolic form: $A \cdot B > (A - X)(B - Y)$, where $A, B, X$ and $Y$ are algebraic expressions that always take positive values. Then, the students stated: Opinion on the inequality is true or false depends on whether numerical values of the expressions $A - X$ and $B - Y$ are positive or negative. If these expressions took only positive values the inequality is true. But what would happen if both of these expressions were negative? Next the solver fixed: $a = b = c = d = \frac{5}{6}$, and having performed the suitable calculation, this person found that the left side of the inequality $C_2$ takes the form $(\frac{1}{6})^4$, whereas the value of the right side is $\frac{4}{9}$.

2.2 Analysis of students’ procedure in validation of reasoning by student $S_2$

Out of 74 participants who was to assess of reasoning by student $S_1$, 62 assessed reasoning by $S_2$ too. There were no signs of any attempts to solve the task in seven works. The other students first verified the validity of the inequalities
denoted as Claim 1 – Claim 4, and stated that they all held. The observed
participants justified their opinions e.g. as follows: “On both sides of the first
inequality there is an expression \((1 - a)\), and on the right side are subtracted
in sequence the numbers \(b, c\) and \(d\), which are positive. Therefore, on the right
side we obtain an expression whose numerical value is less than the value on the
left side. It is a similar case in the other three examples”. Then, the students
noted down: “The inequalities Claim 1 – Claim 4 are correct”. Further, having
read the phrase “It follows therefrom that...” (see description of reasoning by
\(S_2\)), the participants began to wonder how it could be explained. They asked
themselves a question: “What is the relationship between Claim 1 – Claim 4
and the inequality Claim 5?” While trying to find the answer to this question,
the students acted in different ways. Five of them stated that they could not
assess the correctness of reasoning by \(S_2\) and finished their work on the task.

Analysis of how the remaining participants acted formed the basis for
identifying the following three approaches to answer this question.

Approach 1.

Stating lack of justification for the inequality Claim 5 following
from Claim 1 – Claim 4 in reasoning by student \(S_2\)

Approach 2.

Trying to complete the missing link in argumentation

Approach 3.

Validating the inequality Claim 5 regardless of its relationship with
the inequalities Claim 1 – Claim 4

All the students who chose Approach 1 finished their work formulating
the final opinion: “Reasoning by \(S_2\) is incorrect”. Using Approaches 2 some
students came to the statement “Reasoning by \(S_2\) is correct”, and others to
the opposed statement. It was similarly with reference to Approach 3.

Distribution of subjects’ using Approaches 1–3 to get the final opinion
about reasoning by \(S_2\) is shown in Table 4.

<table>
<thead>
<tr>
<th>Approaches used to get answer</th>
<th>Reasoning by (S_2) is correct</th>
<th>Reasoning by (S_2) is incorrect</th>
<th>I am not able to evaluate</th>
</tr>
</thead>
<tbody>
<tr>
<td>App.1</td>
<td>14</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>App.2</td>
<td>12</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>App.3</td>
<td>19</td>
<td>12</td>
<td>5</td>
</tr>
</tbody>
</table>

**Table 4.** Distribution of evaluations of reasoning by \(S_2\) and approaches to reach them.
2.2.1 Approach 1

Nineteen students who did not recognize any explicit basis for the inequality Claim 5 to follow from the preceding Claims, decided that there was no such following and so the reasoning is incorrect. The following example illustrate this opinion.

Example 6

*Reasoning* $S_2$ is incorrect, because there is no explanation why does the thesis of theorem $[T]$ come from.

2.2.2 Approach 2

Thirty-one students stated that $S_2$ did not justify why the inequality Claim 5 followed from the inequalities Claim 1 – Claim 4, and that it was not obvious whether such following occurred at all. We can say, using terms of Rav (1999), that those students arrived at an impasse, not seeing why Claim 5 is to follow from Claim 1 – Claim 4 as $S_2$ affirms. These students attempted to complete this argumentation. Two directions of such attempts could be distinguished.

1. Thirteen people focused on analyzing the structure of the expression on the left side of the inequality Claim 5. They stated that it resulted from multiplication of $(1-a)$ by the product $(1-b)(1-c)(1-d)$. Most students in this group (10 of them) concluded that the value of the expression $(1-a)(1-b)(1-c)(1-d)$ was greater than the value of $(1-a)$. As a consequence, they recognized the inequality Claim 5 as true, because the inequality Claim 1 was also true. The other three participants of the group considered now found that the product on the left side of the inequality Claim 5 took smaller values than $(1-a)$. They justified their opinion that the expression $(1-a)$ was multiplied by an expression whose value was less than 1. They finally concluded that reasoning by $S_2$ was incorrect.

2. Some of the students assumed that $S_2$ could think of “multiplying both sides” of the inequalities Claim 1 – Claim 4, because the left side of the inequality Claim 5 was the product of the expressions that occurred on the left sides of Claim 1 – Claim 4. Nine participants stated that $S_2$ did not perform the multiplication correctly because the right side of the inequality Claim 5 did not have the form as it should have had, which
was \((1 - a - b - c - d)^4\). Their answer was that the whole reasoning was incorrect.

The other two students stated that reasoning by \(S_2\) was correct, because the inequality \((1 - a - b - c - d)^4 > 1 - a - b - c - d\) was satisfied.

Below, there are examples illustrating the students’ standpoints mentioned above.

**Example 7**

*Reasoning by \(S_2\) is correct. Each of the factors \(1 - a, 1 - b, 1 - c\) and \(1 - d\) is greater than \(1 - a - b - c - d\), their product is also greater than \(1 - a - b - c - d\).*

**Example 8**

*Student \(S_2\) made an error because having four (correct) inequalities he multiplied only the left their sides but the right side was written only once. Thus his conclusion is not logical and nothing justifies it.*

### 2.2.3 Approach 3

Two directions of validating the inequality Claim 5 can be distinguished.

1. One of them involved an attempt to construct a counter-example for the inequality Claim 5. Five students provided examples of such values of \(a, b, c, d\) which they considered to be counter-examples (but, in fact, they did not be counter-example). One of the selected systems of the values of \(a, b, c, d\) was as follows:

   \[
   a = 0,1; \quad b = 0,01; \quad c = 0,001; \quad d = 0,0001.
   \]

   The person who provided this “counter-example” did not verify whether selection of these values was appropriate. The other four students did not perform any verification, either.

2. The second direction was to analyze the signs of the numerical values that algebraic expressions can take on both sides of the inequality Claim 5 (2 students). One student wrote his reasoning as follows: “Because \(a, b, c, d \in (0,1)\), we know that the expression on the left side of the inequality is always positive. Therefore, the sign of inequality is correctly written because the expression on the right side is negative. Thus, if each expression \((1 - a), (1 - b), (1 - c), (1 - d)\) whose result is positive is multiplied by itself, a positive number will be obtained, which is greater than the negative number \(1 - a - b - c - d\).”
2.3 Evaluating the correctness of reasoning (acknowledged by its author as proof of the theorem) as solving a mathematical problem

A mathematical task, although perhaps not a typical one, is obviously evaluating the correctness of another person’s reasoning. Let us recall that most of the students actually solved the following tasks:

(I) Evaluate the correctness of reasoning by $S_1$ and justify your answer;
(II) Evaluate the correctness of reasoning by $S_2$ and justify your answer;
(III) Evaluate the correctness of two different examples of reasoning concerning the justification of the same theorem.

The following are synthetic remarks on the students’ solutions of those tasks.

2.3.1 Task (I)

From the above-described tables and descriptions of students’ approaches it can be noted that 74 participants were assigned to solve Task (I), with 67 of them solving this Task. Only seven students managed to provide the answers which could be recognized as correct. They included both the answer: “Reasoning by $S_1$ is incorrect” and the correct justification for this answer. Each of these seven students stated that the inequality $C_2$ did not follow from the assumption of the theorem $[T]$, because there were such values of $a, b, c, d$ (satisfying the assumption of the theorem) for which this inequality was not true.

Surprisingly, as many as 60 people erroneously solved the task (I) and provided the answer: “Reasoning by $S_1$ is correct”. Many students in this group (55 of them) most probably reckoned that the inequality $C_2$ was obtained by applying the rule: If we reduce factors, we also reduce the product. In this way, these students shared false belief of $S_1$: Reducing factors decreases the product. Analysis of students’ solutions also revealed other substantive errors, one of which draws special attention. Namely, in the justification that reasoning by $S_1$ was correct, the student applied the following false belief: The algebraic expression taking only positive values is greater than the expression that can take both positive and negative values.

2.3.2 Task (II)

Out of the 62 participants who were asked to solve Task (II), 55 submitted their answers in writing. The correct solution of this task is the one whose author
wrote: “Reasoning by $S_2$ is incorrect”, and justified the answer by providing a statement: “The conjecture of the theorem [T] cannot be concluded from the inequalities Claim 1 – Claim 4 (although the theorem itself is correct)”. This correct solution was presented by 22 students.

However, many students provided a wrong solution to Task (II), stating: “Reasoning by $S_2$ is correct”. Let us recall at this point that some participants thought that the student $S_2$ could use the inequality

$$1 - a < (1 - a)(1 - b)(1 - c)(1 - d),$$

which they thought was true, as one of the reasoning links. Such a way of thinking may result from the previously mentioned false belief of the subjects: Multiplication makes bigger.

The students who found that the inequality

$$(1 - a)(1 - b)(1 - c)(1 - d) > (1 - a - b - c - d)^4$$

followed from the inequalities Claim 1 – Claim 4 were probably convinced that inequalities can be multiplied by sides, without any additional assumptions about the signs of the expressions in them. In fact, this statement of the expressions of the students signals their false belief, which is an algebraic equivalent of the false belief: Reducing factors decreases the product.

It is worth noticing that two students, who wrote double inequality:

$$(1 - a)(1 - b)(1 - c)(1 - d) > (1 - a - b - c - d)^4 > 1 - a - b - c - d$$

revealed in this way that they held both of the above-mentioned false beliefs simultaneously.

2.3.3 Task (III)

The results of the evaluation of both examples of reasoning, presented finally by 45 students, are shown in the following table.

<table>
<thead>
<tr>
<th>Students’ opinions on reasoning by $S_1$ and $S_2$</th>
<th>$S_1$ – correct</th>
<th>$S_1$ – incorrect</th>
<th>$S_2$ – correct</th>
<th>$S_2$ – incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of students</td>
<td>13</td>
<td>28</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>

The data contained in the table demonstrate that only four students assessed rightly that both of reasoning (by $S_1$ and $S_2$) were incorrect and they justified their point.
The answers of five out of twenty-eight students in the table above require a special comment. Each of these five students reckoned that reasoning by S\textsubscript{1} was proof of the theorem [T]. The consequence of such an opinion should be their recognition of the theorem [T] as true. However, these students, while analyzing reasoning by student S\textsubscript{2}, searched for a counter-example for this theorem, and finally concluded that the theorem [T] was false. This suggests that these students treated the consequences, which each of the justifications of this theorem led to, as completely independent of each other.

3 Conclusions

In this section, the authors of this paper will present the analysis of students’ approaches with respect to the revealed symptoms of understanding or not understanding the essence of mathematical proof.

- It should be noted that in both examples of reasoning presented in the research task, there is no explicit indication of the arguments (in the form of theorems, definitions and formulas) which would justify particular inferences. Indication of the basis of inference is treated as a characteristic and necessary feature of mathematical proof (for example Avigad, 2006; Krygowska, 1977). If an author fails to indicate proof of explanation for each inference, it is difficult to judge the validity of this proof. The participants recognized this feature of proof to be essential, and expressed it in two ways: they argued that “given reasoning can not be evaluated” or themselves tried to complete the explanation of the next steps of proof. The inner need to justify subsequent steps in proof-type reasoning should be treated as a symptom of understanding that it is necessary to indicate the basis of inference for each inference in mathematical proof.

- The explanation of inference in proof raises a question of whether the basis of inference referred to by the author of this proof is “the law” in mathematics, or it is subjective knowledge of its author. Some elements of this subjective knowledge of an individual can be inconsistent with objective mathematics. They are variously called, for example: misconceptions, false beliefs, misbeliefs, faulty beliefs (Graeber, Tirosh, 1990; Bell, 1988) or false convictions (Pawlik, 2007). According to Ciosek, Turnau (2015) “false belief is a relationship or regularity wrongly taken as generally true. In may be revealed when the solver of a problem refers to it (explicitly or implicitly) while justifying (wrongly) his/her solution”
Elements of such knowledge are considered by an individual to be certain, obvious, and therefore requiring no verification. As it was already mentioned earlier, such knowledge was used by both \( S_1 \) and \( S_2 \) to justify inferences in their proofs. Let us recall that it is about two false statements: (1) \textit{Reducing factors decreases the product}, and (2) \textit{Multiplication makes bigger}. The study revealed that approximately 75\% of the students evaluating reasoning by student \( S_1 \), and about 18\% of the students evaluating reasoning by student \( S_2 \), held false beliefs (1). This is evidenced by authentic exemplary statements made by the participants:

- “\( S_1 \) reduces the factors, so the product is also reduced”;
- “It can be noted that the factors on the right side of the inequality [2] are less than the corresponding factors on the right side of this inequality, so the product on the right is less than the product on the left”;
- “Reasoning by \( S_2 \) is incorrect because he/she has written four inequalities, then multiplied the left sides, and the right side he/she wrote “once”, and he/she should also multiply the right sides;
- “\( S_2 \) has done wrong, because after multiplying by sides, he/she should get: \((1 - a)(1 - b)(1 - c)(1 - d) > (1 - a - b - c - d)^4\), but \((1 - a - b - c - d)^4\) is not greater than \(1 - a - b - c - d\)”.

The ten students (accounting for about 16\%) who found reasoning by \( S_2 \) a correct proof, revealed a false belief (2). This is evidenced by their statements such as: \textit{If we multiply} \((1 - a)\) by \((1 - b)(1 - c)(1 - d)\), \textit{then we get more than} \((1 - a)\). The authors of these claims probably were not aware that the statement they considered obvious “if \( A, B > 0 \) then \( A \cdot B > A^2 \)” does not hold true for any positive numbers \( A \) and \( B \); it is false when \( A > 0 \) and \( 0 < B < 1 \).

The approach described above turns attention to yet another important aspect of mathematical proof: for given inference in proof to be considered valid, it is necessary to verify whether the self-imposing argument for its justification is an element of objective knowledge. As it was already demonstrated earlier, this can be extremely difficult to verify because the use of subjective, often intuitive, knowledge (e.g. false beliefs), does not raise any doubts as to its correctness. Accepting self-imposing arguments that we ad hoc consider to be justifications for individual inferences is a symptom of misunderstanding that it is necessary to verify whether the explanation provided is mathematically correct.
Let us recall that the research task contained two examples of reasoning from elementary mathematics, regarding the same theorem previously unknown to the participating students. These examples of reasoning were obtained from other students (S₁, S₂) solving the task:

Prove the following theorem [T]:

If \(a, b, c, d\) belong to the interval \((0,1)\), then holds the inequality

\[(1 - a)(1 - b)(1 - c)(1 - d) > 1 - a - b - c - d\]

The examples of reasoning by S₁ and S₂ were deliberately prepared to be analyzed one by one. It was about gaining insight into how a student understands the relationship between proof of the theorem and the truth of this theorem. It turned out that not all the students understood this relationship well. Some of them found that proof written by S₁ was correct; it can therefore be presumed that the consequence of such an evaluation is acknowledging the truth of the theorem [T]. The same students, while examining the correctness of proof written by S₂, attempted to find a counter-example for the theorem [T]. Such an approach is treated as a symptom of erroneous understanding of the relationship between the theorem and its proof. This erroneous understanding of the analyzed relationship by an individual is manifested in recognizing a given theorem as true (accepting its proof), and at the same time casts doubt on the truth of the same theorem in the context of its another proof.

The research clearly revealed one more difficulty for students in assessing the correctness of reasoning recognized by their authors as proofs of a given theorem. A symptom of this difficulty is as follows: in the case of uncertainty as to the correctness of proof, the students assessed the validity of the original theorem [T] (for example using numerical examples). If the theorem was presumed to be true, the analyzed proof was considered valid. This kind of approach is treated as a signal that the individual may accept the following inference as valid: It may be concluded from acknowledging the truth of the theorem that the reasoning that has been presented as proof of this theorem is correct.

4 Didactical implications

Identifying difficulties in understanding mathematical proof is equally important to pointing out to the causes of their occurrence. In this context, it is
worth quoting W. P. Thurston, who was awarded the Fields Medal, stating that in fact even mathematical studies did not shape sufficiently the notion of mathematical proof: “When I started as graduate student at Berkeley, I had trouble imagining how I could “prove” a new and interesting mathematical theorem. I didn’t really understand what a “proof” was” (Thurston, 2006, p. 45).

It seems that one of the reasons for such state of affairs is that university students in the course of mathematics very often learn mathematics as a “ready made knowledge”. Semadeni (2015) stated: Mathematics is usually presented to students (from primary school to graduate level) as a ready-made edifice of knowledge. This condensed, elegant form often fails to reflect the dramatic intellectual struggle which lasted for centuries to clarify some very basis concepts (Semadeni, 2015, p. 39). A similar opinion on the issue of proof is expressed by Weber (2010, p. 307): “Advanced mathematics courses are typically thought in a “definition-theorem-prove” format. Much of the lectures in advanced mathematics courses consist of professors presenting proof of theorem to students”.

The above statements suggest that students generally are provided with ready-made proofs of theorems, so they do not participate in the process of their development. This means that during mathematics lectures and classes students have no opportunity to present their subjective (sometimes erroneous) views on what constitutes a mathematical proof.

Perhaps because of the large scope of the learned content, university students, potentially future teachers of mathematics, have little opportunity to reflect on the fact that mathematics is not just a collection of definitions, properties and patterns, but it is also a specific, creative human activity. Other authors of the works also point this out (e.g. Weber, 2010; Boyle at al., 2015). Boyle at al. stated: However, mathematics teachers have been provided limited opportunities as learners to construct arguments and critique the reasoning of others, and hence have developed perceptions of proof (p. 32).

What was said before suggests the need to search for methods to introduce changes in current mathematical education with reference to the problem of proof in mathematics. The literature of mathematical education contains descriptions of certain concepts of courses and didactic approaches addressed to university students who aim to improve their understanding of the concept of proof (e.g. Jones, 2000, Knut, 2002; Weber, 2010; Brandt, Rimmensch, 2012; Boyle at al., 2015).

In the authors’ opinion there is a need to develop a conception addressed specially to the major students of mathematics – future teachers of these subject. Such a conception should emphasize on understanding of mathematical methodology, among others through engaging students in creative mathema-
tical activity. In the works (Sekerák, Šveda, 2008; Ciosek, Żeromska, 2013) one can find examples of mathematical problems, which could be helpful in creating the mentioned conception.

References


C i o s e k, M., T u r n a u, S.: 2004, Which aims of mathematics teaching should be kept as fundamental and universally important?, in: Giménez J., Fitz Simons G. E., Hahn C (Eds.), *A challenge for mathematics education: To reconcile commonalities and differences*; [CIEAEM 54], Biblioteca de Uno 195. Graó, Barcelona 2004, 161–165.


C i o s e k, M., Ż e r o m s k a, A. K.: 2013, *Rozumowania w matematyce elementarnej. Hipotezy, twierdzenia, dowody*, Wydawnictwo Naukowe Uniwersytetu Pedagogicznego, Kraków.


Gardiner, T.: 1999, Conflicts between mathematics graduates’ proof behaviours and their stated beliefs and proof, 31–36 academia.edu/6300598/Measuring_Mathematical_Beliefs_and_Their_Impact_on_the_Learning_of_Mathematics_A_New_Approach


How major mathematics students evaluate


Rav, Y.: 1999, Why do we prove theorems, Philosophia Mathematica, 7(11), 5-41.


Solomon, Y.: 2006, Deficit or difference? The role of students epistemologies of mathematics in their interactions with proof, Educational Studies in


Jak studenci matematyki oceniamy poprawność dowodu matematycznego

**Streszczenie**

Artykuł prezentuje wyniki badania umiejętności oceniania poprawności dowodów matematycznych przez 74 studentów matematyki (polskich i słowackich) – przyszłych nauczycieli matematyki. Badani mieli za zadanie zapoznać się z dwoma autentycznymi rozumowaniami pochodzącymi od innych studentów matematyki (S₁ i S₂), uznawanymi przez ich autorów za dowody pewnego twierdzenia. Obydwa rozumowania były błędne, choć każde z innego powodu. Analiza zebranego materiału badawczego była nastawiona na zidentyfikowanie i scharakteryzowanie trudności w rozumieniu dowodu matematycznego oraz symptomów rozumienia istotnych cech dowodu w matematyce. Narzędzie badawcze zostało tak dobrane, by umożliwić pewien wgląd w pojmowanie przez studentów związku między twierdzeniem a jego dowodem w matematyce.

**Marianna Ciosek**
Pedagogical University of Krakow (retired)
Poland
e-mail: ciosek.maria@gmail.com

**Anna K. Żeromska**
AGH University of Science and Technology
Krakow, Poland
e-mail: zeromska@agh.edu.pl

**Dušan Šveda**
Universita Pavla Jozefa Šafárika v Košiciach
Slovakia
e-mail: dusan.sveda@upjs.sk