Kristina Reiss, Aiso Heinze, Oldenburg; Eckhard Klieme, Frankfurt

ARGUMENTATION, PROOF AND THE UNDERSTANDING OF PROOF

Abstract:
Argumentation and proof play an important role in mathematics. In recent years several empirical studies have revealed deficits in students' abilities in logical argumentation and in their understanding of mathematical proofs. In an empirical survey with upper secondary students we investigated different components of both mathematical, and general, competence like declarative knowledge, methodological knowledge, metacognition and spatial abilities in relation to students' performance for geometrical proof items. The results indicate that these components explain a comparatively high proportion of inter-individual differences.

1 Theoretical framework

1.1 Argumentation and proof in the mathematics classroom

Logical argumentation, reasoning and proof may be regarded as highly important topics in mathematics. Despite the fact that mathematics may even be regarded as a *proving science* the role of proof in the school curriculum did not always reflect that importance. In the last few years there was a significant change in the teaching of reasoning and proof in the mathematics classroom. During the lively discussions of the 70s and 80s as to whether proofs should be part of the mathematics curriculum in secondary schools, mathematics educators argued that proving in the classroom had developed into a topic that particularly emphasised formal aspects but disregarded mathematical understanding (Hanna, 1983). In the 90s the situation changed: proof is again regarded as an important topic in the mathematics curriculum (NCTM, 2000) and is an essential aspect of mathematical competence. However *proof* was not necessarily used as a synonym for *formal proof*. Several authors like Hanna and Jahnke (1993), Hersh (1993), Moore (1994), Hoyles, (1997), Harel and Sowder (1998) pointed out that proving spans a broad range of formal and informal arguments.

Understanding or generating proofs is a significant component of mathematical competence. Moreover, mathematical argumentation has been identified as the essential element of higher order mathematical competence in the TIMSS study. Nonetheless not only the TIMSS study (cf. Klieme 2000) but also other empirical findings have revealed wide gaps in students' understanding of argumentation and proofs (Senk, 1985; Usiskin, 1987; Healy and Hoyles, 1998). In a systematic investigation with high-attaining grade 10 students in Great Britain, Healy and Hoyles (1998) showed that there are deficits in their understanding of proofs, their ability to construct proofs, and their views on the role of proof. Even high-attaining students were far from proficient in constructing mathematical proofs, and were more likely to rely on empirical verification. However, most of them were well aware that once a statement has been proved it holds for all cases within its domain of validity. Moreover, they were frequently able to recognise a correct proof, though their choices were influenced by factors other than correctness, such as perceived teacher preference.

1.2 Geometrical thinking and geometrical competence

Geometry is a good starting point to teach and learn mathematical argumentation, to explore mathematical concepts, to fill the gap between everyday life and mathematics, and to value mathematics as a part of human culture. Accordingly, geometrical competence can be regarded as an important prerequisite of understanding mathematics. In recent years research results in the field of cognitive psychology gave new ideas for describing the cognitive processes and specific knowledge structures needed to solve geometry problems. In particular, geometrical reasoning has been investigated in detail by setting various types of test items and observing students working on these items. In the model of Greeno (1980), with respect to geometrical knowledge, one may distinguish between (a) theorems and rules, (b) visual patterns — such as the image of corresponding angles — and (c) "strategic principles" which, for example, govern the construction of proofs.

Ideas about the proving process of experts are described by Koedinger and Anderson (1990). When experts construct geometrical proofs, they do not merely retrieve declarative knowledge like definitions and theorems from the memory and combine these with logical deductions. On the contrary, they outline their argumentation in broad terms, taking a constructivist approach. They use visual models, in which they are able to "see" properties and connections, and "pragmatic reasoning schemas" such as set patterns for individual steps in the proving process. This indicates that geometrical competence is not merely a question of talent, but of specific skills and knowledge (Koedinger, 1998).

Nevertheless, geometrical competence requires specific knowledge; it is based on general psychological mechanisms that are central to other domains of mathematics as well as to thinking and problem solving in general. Where geometrical knowledge is concerned, a distinction must be drawn between declarative knowledge and methodological knowledge (e.g., knowledge of construction procedures and the principles of mathematical proof). As a general mechanism, on the one hand, metacognition can be identified, i.e., the active steering and control of one's own cognitive processes. On the other hand, various components of general intelligence are relevant; according to Clements and Battista (1992), spatial reasoning is of particular importance for geometrical competence. The last point is also described by Klieme (1986), who showed by a meta-analysis that there is a relationship between spatial abilities and performance in mathematics.

1.3 Research questions and hypotheses

The present study integrates the lines of research described above. The students' proof performance, geometrical competence and its cognitive prerequisites are investigated by reference to TIMSS items, with a particular focus on respondents' understanding of proof, metacognition and spatial abilities. The hypotheses to be addressed in this article are the following:

(i) The geometrical knowledge of German upper secondary students is comparatively low. This does not only apply to declarative knowledge, but also to methodological knowledge. Insofar, the results of the TIMSS study can be confirmed.

(ii) The understanding of geometrical concepts (we investigated the concept "congruence" as an example) is comparatively poor. Basically the students know definitions or symbols. The number of correct answers in the case of examples, drawings, applications and connections is rather small. Students with a correct understanding of geometrical concepts undertake the geometrical items significantly better.

(iii) For the students it is easier to evaluate the correctness of proofs than to perform proof items. However there is a relationship between the quality of students' proof and the evaluation of the correctness of proofs.

(iv) There is a significant correlation between the performance of geometrical items and spatial abilities.

2. Design of the study

2.1 Instruments

The questionnaire the students were presented with consisted of five parts. The first part was a test to measure their performance in geometrical problem solving. There were nine TIMSS items allocated to two proficiency levels. The six easier items were assigned to the lower level which corresponded to proficiency levels I and II in Klieme (2000) while the three more difficult ones were assigned to the higher level which corresponded to levels III and IV in Klieme (2000). The lower level items were answered correctly by more than half of the students in the international TIMSS population and the three higher level items by one third or less of the students.

In the second part, declarative geometrical knowledge was measured as a prerequisite of geometrical competence. For the evaluation of students' declarative knowledge we chose the concept "congruence", a central concept of school geometry. Students were asked to give a definition, an example, and a visual or graphic portrayal of the word "congruent", and to name a mathematical theorem in which the concept features. The students' open-ended answers were coded according to a specially developed category system; one point could be earned for each of the four aspects.

Methodological knowledge, in the form of knowledge about the validity of mathematical arguments, was assessed using an item from Healy and Hoyles' (1998) proof questionnaire. The item dealt with the proof problem as to whether or not a given triangle was isosceles. The students were presented with one correct formal proof, one correct narrative proof, one solution with an empirical argument and one circular solution. They were then asked to assess the correctness and generality of each of the four arguments. These eight assessments were marked either right or wrong. The total mark is then taken as an index of methodical knowledge (understanding of proof).

To assess an important aspect of metacognition, the students were asked to rate each of the items in terms of whether they think the answer they gave was correct or incorrect. These ratings were then compared with the students' actual performance. The number of items correctly rated by the students as having been answered correctly or incorrectly is taken as an index of metacognitive competence.

The last part of the test was the so-called *Schlauchfiguren-Test* (Stumpf & Fay, 1983), an instrument which is well known in Germany. *Schlauchfiguren* presents different views of complex tubular figures, which have to be judged with respect to a specific point of view. Validation studies have shown that the test calls for both spatial ability and deductive reasoning. It is therefore a suitable instrument to capture those aspects of general intellectual ability which are cognitive prerequisites of geometrical competence.

2.2 Study Sample and Administration

The present study was administered in two German *Gymnasium* schools — the most academic type of secondary school. The total sample comprised 81 students (48 female), all of them taking mathematics courses in the 13th grade, i.e. the final year of secondary school. Of the 81 students who participated in the study, 59 attended a regular mathematics course and 22 an advanced course.

The students were first presented with the geometry problems. The items became gradually more difficult as the test went on and the closed and open-ended questions were presented in two separate sections, each with a time limit. There was no indication that students had too little time to complete the test. After the geometry items, the students were presented with the supplementary questions (metacognitive assessment, declarative knowledge, understanding of proof and spatial reasoning). Some time later, after a preliminary analysis of results, about a quarter of the students were asked to attend a face-to-face interview conducted by a member of the research team. In these individual interviews, students were presented with a number of items taken from the written test, and a further TIMSS item in which a proof was to be constructed. The students were asked to work on these items using the think-aloud method, and were videotaped as they did so.

3 Results

3.1 Descriptive Findings on Geometrical Competence

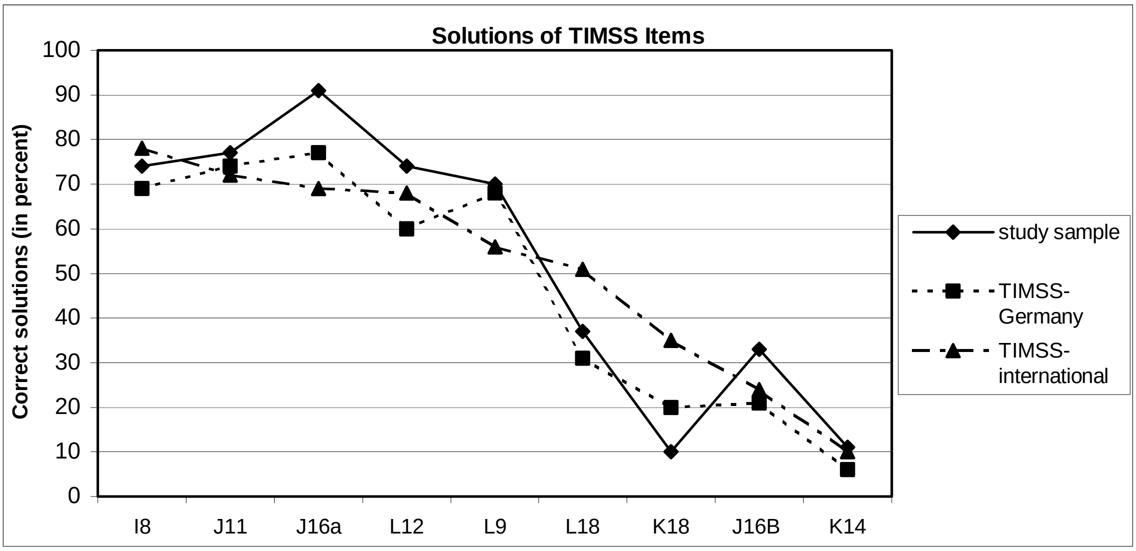
Table 1 shows the percentage of students providing correct solutions for each of the nine geometry items administered in our study, along with the corresponding results for the international TIMSS sample and the German national TIMSS sample.

Items	I8	J11	L12	L9	J16a	J16b	L18	K18	K12
Study sample	74.1%	76.5%	74.1%	70.4%	93.8%	34.6%	42.0%	16.1%	11.1%
TIMSS international	78%	72%	69%	68%	56%	51%	35%	24%	10%
TIMSS Germany	69.4%	73.5%	59.6%	67.6%	77.4%	21.4%	31.0%	18.5%	5.9%

Table 1 Correct answers for the geometry items

The results show a remarkably high level of correspondence across the three samples, both in the average achievement level and in the performance in each of the nine questions (see the diagram on the next page).

Averaged out across the nine items, 53% of the students in our sample provided correct solutions, compared to 51% in the international sample and 47% in the German sample. Across the nine items, the correlations between the performance in our sample and in the representative German and international TIMSS samples amount to .97 and .89 respectively; both of these correlations are highly significant. The relative strengths and weaknesses of the German students — in comparison to the international TIMSS sample — are thus also reflected in our small sample.



In terms of international comparison, German students performed relatively well in the multiple-choice question L9, in which the symmetrical properties of a geometrical figure were to be evaluated. In contrast, they performed relatively poorly in the open-ended questions K18 and L18. In these they were asked to construct a geometrical proof in the tradition of classical Euclidean geometry (K18) and to determine the length of a line segment and to verbally describe the process of construction and argumentation (L18). Notable differences in the performance of the students in our sample and the German TIMSS sample only emerged for items J16a and J16b. Here the participants were asked to draw the triangle resulting from the reflection or rotation of a given triangle in a co-ordinate system.

Overall, however, the results show a remarkably high level of correspondence across the three samples, in both the general level of geometrical competence and the profile of performance in each of the nine items. This suggests that the cognitive mechanisms activated by the students tackling the geometry items were also of a similar nature, meaning that the results of the detailed investigations reported in the present paper can probably be generalised to the entire TIMSS population. The structural similarity observed in the patterns of performance in both our sample and the TIMSS national sample justifies the approach taken in this study i.e. carrying out detailed cognitive analyses of students' ways of tackling TIMSS items in a small supplementary study, but then claiming validity for a much larger population of students.

The following results are of particular interest from the didactical viewpoint:

- Item L12, in which the length of a diagonal is to be determined, is much easier than item L18, in which students are also asked to determine the length of a given line in a geometrical drawing. Both items can be solved by constructing specific triangles and applying the Pythagorean theorem. However the crucial advantage of item L12 is that students are offered a choice of several alternatives, and that they can use their common sense to weigh up these alternatives. Evidently students find it much easier to make such a choice — here using their knowledge about the properties of a regular hexagon and the sides of a right-angled triangle — than to construct and calculate an answer of their own, showing their working. It should be emphasised that this by no means calls the validity of multiple-choice items into question, but it does demonstrate that these questions call for different cognitive subprocesses than open questions, in which students have to give an explicit description of their approach to the problem.

- The construction of an explicit proof, as required in K18, is in turn much more difficult than determining a precise number value and giving arguments for the specific calculations performed, as called for in L18. Probably it is the precise result that makes L18 a simpler task to perform. Proving a statement seems to be regarded by students as a more open-ended activity.
- The most difficult item by far in all three populations is K14. This item can evidently only be solved if the students are able to use visualisation or drawing skills to recognise the appropriate strategy for calculating the length of the string. One can, for example, visualise the surface of the rod as a rectangle, with the string representing the diagonal. The length of the diagonal can then be calculated simply by applying the Pythagorean theorem. Items such as L12 and L18 demonstrate that students have no major difficulties in estimating and/or calculating such a length and it is evident that spatial visualisation skills pose the decisive obstacle here. Moreover the poor performance indicates that students have difficulties in solving non-routine problems which ask for restructuring the elements involved.

3.2 Findings on the Understanding of Proof and Views of the Role of Proof

In view of the observation that very few students (20% and 35% of the representative German and international TIMSS samples respectively) were able to construct correct Euclidean geometry proofs (see previous section), we also expected the levels of performance to be rather unsatisfactory in the questions investigating students' understanding of proof and their views of the role of proof (taken from Healy and Hoyles, 1998). Interestingly, our students also found it much easier to judge given proofs than to construct their own proofs (see Table 2).

Proof / feature	Relative frequency (in percent)	Corrected item-total correlation
Correct formal proof		
/ correct	57	.49
/ general	57	.45
Correct narrative proof		
/ correct	42	.45
/ general	30	.38
Empirical argument		
/ incorrect	46	.13
/ not generalisable	60	.26
Formal, circular argument		
/ incorrect	33	.39
/ not generalisable	27	.40

Table 2 Components of methodological knowledge (understanding proof)

As shown in Table 2, 57% of our respondents recognised the correct formal proof (using congruence) to be correct, and the same proportion of participants correctly appreciated its generality. A similar proportion of respondents recognised a purely exemplary, empirical argument to be incorrect: 46% said that the argument was incorrect, and 60% recognised that it was not generalisable. However the low item-total correlations of these two answers (see right-hand column of Table 2) showed that even students with a low general understanding of proof were aware that the purely empirical argument was incorrect and not generalisable. It could be that they rejected the argument simply because they intuitively recognised that it did not satisfy the formal criteria for proof.

Further findings indicate that a significant minority of students implicitly assume that only formally presented proofs are acceptable. Only 42% recognised the correct narrative proof (using a geometrical diagram) to be correct; moreover, only 30% correctly appreciated its generality. Conversely, only one third of the respondents recognised that a formally presented, but circular argument was incorrect, and 27% that it was not generalisable. In other words, the students find it just as difficult to accept correct proofs with a non-formal presentation as to reject incorrect, but formally-presented, proofs. The high item-total correlations show that, generally, only high-attaining students were able to evaluate these proofs correctly.

Thus far, we have considered various aspects of students' understanding of proof: the ability to construct proofs and to recognise the correctness and/or generality of given proofs. In the following, we turn to student preferences and views of the role of proof (cf. Table 3).

Proof / feature	Relative frequency (in percent)	
	Present study	Healy & Hoyles (1998)
Correct formal proof		
/ useful as explanation	35	
/ assumed teacher choice	58	48
/ student's choice	30	24
Correct narrative proof		
/ useful as explanation	48	
/ assumed teacher choice	12	15
/ student's choice	25	26
Empirical argument		
/ useful as explanation	33	
/ assumed teacher choice	4	19
/ student's choice	10	40
Formal, circular argument		
/ useful as explanation	15	
/ assumed teacher choice	11	18
/ student's choice	28	13

Table 3 Preferences for different proofs and arguments

Apart from evaluating the correctness and/or generality of the proofs, the students were asked to judge whether each of the four arguments was appropriate in order to explain the particular geometrical content to one of their classmates, to state which answer would be given the best mark, and to identify the argument which would be closest to what they would do if asked to answer the question. The most interesting finding, in which our results are entirely in line with those of Healy and Hoyles (1998), is that a clear majority of students assume that their teacher would give the best mark to a correct formal proof, but that the respondents are almost as likely to choose the correct narrative proof as the correct formal proof for their own approach. In our study, the students even thought that the narrative proof was more appropriate to explain the geometrical content to a classmate than the formal proof. Individual preferences for one of the two proofs were related to whether or not it was felt the argument was required to convince or explain. Those who chose the narrative proof as the best explanation were more likely to select it as their own approach, and the same goes for the formal proof.

We interpret the findings, particularly those on assumed teacher preference, as an indication that the majority of students consider a correct (noncircular), formally presented proof to be the mathematically accepted norm. However, they have not entirely adopted this norm in their own attempts at proof or their understanding of convincing mathematical arguments. The majority feel that the correct narrative proof is the best way of explaining geometrical content to their classmates, while one third of respondents selected the purely empirical argument as the best explanation. These figures, particularly the proportion of students selecting the empirical argument, are markedly lower than the corresponding British results (probably due to the age difference — 13th graders were tested in Germany, 10th graders in England and Wales). Nevertheless, in the German group too, there is a clear discrepancy between the perceived “official” norms and the students’ personal preferences for mathematical arguments. In fact, correlation analyses show that it is the students with high levels of geometrical competence and a better understanding of proof who tend to select the empirical argument as the best explanation for their classmates. Thus the ability to differentiate between proofs perceived to be the mathematical norm and other forms of mathematical argument is possibly an important facet of mathematical competence. The fact that 28% of our respondents selected the circular argument for their own approach is a critical finding. The majority of these students stated that their teacher would give the best mark to the correct formal proof, but evidently did not consider circularity to be an obstacle to choosing this argument as their own approach.

Overall, these results show that the proof-related competencies and views of the German *Gymnasium* students participating in this study were insufficiently developed. Even when students recognise the type of mathematical argumentation which is “officially” required — by their teacher, for example, as a representative of the discipline — a large proportion of them still choose (incorrect) empirical or circular arguments for their own approach.

3.3 Descriptive Findings on Declarative Knowledge (Comprehension of Geometrical Concepts)

Our findings on the respondents’ declarative knowledge, summarised in Table 4, are also less than satisfactory from the standpoint of mathematical didactics.

Knowledge component	Relative frequency of correct answers (in percent)	Corrected item-total correlation
Definition	8.6	.16
Example	48.1	.37
Diagram	81.5	.35
Theorem	11.1	.22

<u>Table 4</u> Components of declarative geometrical knowledge, assessed with reference to the concept of “congruence”

When asked to describe the concept of "congruence", 82% of respondents were able to illustrate the concept in a sketch, most of them drawing congruent triangles. Less than half of the respondents were able to give an example of congruence, however. Only about one in ten of the students mastered the central mathematical components of the concept, i.e. were able to provide a definition of the concept and name a mathematical theorem in which it features (e.g. a theorem of triangular congruence). Moreover, the low item-total correlations for these two knowledge components indicate that some of these correct answers may well have been lucky guesses. Instead of providing a definition of congruence, many students simply gave a "translation" of the term, stating, for example, that *kongruent* (a word of Latin origin) means *deckungsgleich* (the German-language equivalent). Such answers were coded as being "incorrect". It is apparent that even students with a vague intuitive understanding of triangular congruence lack exact mathematical knowledge.

3.4 Explaining Geometrical Competence: Declarative Knowledge, Methodological Knowledge, Metacognition and Spatial Reasoning

Having described each of the components of our study design, we will now explore the relations between the scales for geometrical competence, methodological knowledge (understanding of proof) and declarative knowledge (understanding of geometrical concepts). Table 5 shows the intercorrelations, calculated as rank-correlation coefficients (Kendall's tau), on which this discussion is based.

Scale	(2)	(3)	(4)	(5)	(6)	(7)
(1) Geometrical Competence	.76**	.62***	.20*	.24**	.24**	.33***
(2) Geometry, level I/II		.27**	.10	.18*	.21*	.23**
(3) Geometry, level III/IV			.22**	.23**	.21*	.37***
(4) Methodological knowledge				-.01	.05	.12
(5) Declarative knowledge					-.02	.09
(6) Metacognition						.09
(7) Spatial reasoning						

Table 5 Intercorrelations of scales (Kendall's tau – b)
*) p< .05 **) p< .01 ***) p< .001

In addition to the mathematical dimensions of competence and knowledge, the two general psychological predictors — metacognition and spatial reasoning — are also included in the table.

The most important finding is that all four predictors exhibit significant correlations with geometrical competence. This lends support to our basic hypothesis that geometrical competence is dependent on methodological knowledge, declarative knowledge, metacognition and spatial reasoning. The correlation matrix does not actually allow such causal interpretations to be made but, interpreting the results in the light of other research on geometrical knowledge, (Reiss & Abel, 1999), makes it plausible to assume that scales (4) to (7) tap the *prerequisites*, and scales (1) to (3) the *results*, of development of geometrical competence.

As expected, stronger correlations with the predictors emerge at the higher levels of geometrical competence (items on TIMSS proficiency levels III and IV) than at the lower levels of geometrical competence (levels I and II). Understanding of proof is a vital ingredient at the higher levels of competence, but is irrelevant to performance in the easier TIMSS geometry items. This confirms our assumption that the TIMSS proficiency levels really do reflect different standards of (geometrical) competence.

There is no intercorrelation between the four predictors. These evidently represent different prerequisites of geometrical competence, which not only refer to different constructs on the theoretical level, but also are separate on the empirical level. A combination of the four variables can thus be expected to have a much higher predictive efficiency than each of the individual predictors. Accordingly, multiple regression analyses were run for both the entire geometrical competence scale and each of its subscales. The three predictors, “spatial reasoning”, “declarative knowledge” and “metacognition” yield a multiple correlation coefficient of .61 and explain 37% of the variance. Thus a comparatively high proportion of inter-individual differences in geometrical ability can by explained by these three predictor variables. Spatial reasoning, which — as explained above — involves general components of intellectual ability, including deductive reasoning, has the greatest explanatory power. This is true for solving low-level geometrical problems as well as for solving complex geometrical tasks. With respect to complex tasks, spatial reasoning explains a higher proportion of the inter-individual differences in problem solving performance. Declarative knowledge, conceptual knowledge and metacognition however, also explain highly significant, non-redundant amounts of inter-individual variance. The regression analyses for the two subscales again show that performance at the higher proficiency levels calls for different prerequisites than performance at the lower levels of proficiency. While success at the lower levels is largely dependent on a basic cognitive skill (metacognition), performance in the more complex geometrical tasks of levels III and IV has more to do with subject-related conceptual knowledge and methodological knowledge (in this case, the understanding of proof). This confirms Klieme’s (2000) analysis of geometrical competence, in which it was not only postulated that a greater amount of mathematical knowledge must be accessible at higher proficiency levels, but also that more complex processes must be invoked.

4 Discussion

Our analyses have revealed considerable deficits in declarative knowledge (understanding of concepts) and methodological knowledge (understanding of proof). Where declarative knowledge is concerned, it emerged that even students at the end of secondary level often have only a vague intuitive understanding of concepts such as “congruence”, that this understanding is restricted to examples, and that they have no exact mathematical knowledge of the respective definitions and theorems. Similar to the broad representative survey conducted in England and Wales by Healy and Hoyles (on whose methods we have drawn), various misconceptions and misinterpretations in the students’ methodological knowledge were revealed. Many German *Gymnasium* students who have taken advanced mathematics courses seem to assume implicitly that “good” proofs should have a strictly formal presentation. Yet they choose narrative proofs, purely empirical arguments or – in significant numbers – incorrect circular arguments as their own approaches to proofs or in order to explain proofs to their classmates.

These findings, revealing the students' inadequate understanding of proof, can be regarded as an important indication of where the problem areas in mathematics instruction lie. Indeed, this was the basic approach taken by Healy and Hoyles (1998). In the context of a theory of situated cognition, however, the discrepancy between abstract knowledge about the correct construction of proofs on the one hand and (at least partly) erroneous personal preferences on the other hand is easy to understand and can be positively evaluated. Students — especially the more competent ones — bear the context in mind when evaluating differing formulations of mathematical arguments. This is precisely the sort of approach encouraged in modern, reform-oriented conceptions of mathematics instruction. After all, students should not only experience mathematics as a set of fixed rules; rather they should be able to construct appropriate mathematical arguments both in school and in applied contexts. Our findings indicate that the topic of "proof in mathematics instruction" is particularly well suited as an introduction to mathematical argumentation – precisely because of this juxtaposition of views and preferences.

References

BAUMERT, J., BOS, W. & LEHMANN, R. (2000). *TIMSS/III – Dritte Internationale Mathematik- und Naturwissenschaftsstudie. Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Mathematische und physikalische Kompetenzen am Ende der gymnasialen Oberstufe (Bd. 2).* Opladen: Leske + Budrich.

BEATON, A. E.; MULLIS, I.V.S.; MARTIN, M. O.; GONZALEZ, E. J.; KELLY, D. L.; SMITH, T. A. (1996). *Mathematics achievement in the middle school years. IEA's third International Mathematics and Science Study (TIMSS).* Chestnut Hill, MA: Center for the Study of Testing, Evaluation and Educational Policy.

CLEMENS, D. H. & BATTISTA, M. T. (1992). *Geometry and spatial reasoning*. In D. A. Grouws (ed.) *Handbook of research on mathematics teaching and learning*. New York: Macmillian, pp. 420 – 464.

GREENO, J. G. (1980*). Some examples of cognitive task analysis with instructional implications*. In R. E. Snow, P. Frederico, W. E. Montague (Eds.) *Aptitude, learning and instruction*, Vol. 2. Hillsdale, NJ: Erlbaum, pp. 1 – 21.

HANNA, G. (1983). *Rigorous proof in mathematics education***.** Toronto, Ontario: Ontario Institute for Studies in Education.

HANNA, G. (1997). *The ongoing value of proof*. Journal für Mathematikdidaktik, 18, pp. 171-185.

HANNA, G. & JAHNKE, H. N. (1996). *Chapter 23: Proof and Proving*. In Bishop, A. J.; Clements, K.; Keitel, C.; Kilpatrick, J.; Laborde, C. (eds.) *International Handbook on Mathematics Education.* Volume 4, Part 2 Dordrecht: Kluwer, pp. 877 – 908.

HAREL, G. & SOWDER, L. (1998). *Students' Proof Schemes: Results from Exploratory Studies*. In A.H. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Research in Collegiate Mathematics Education*. Providence, RI: American Mathematical Society, pp. 234-283.

HEALY, L & HOYLES, C. (1998). *Justifying and proving in school mathematics. Technical Report on the Nationwide Survey*. Mathematical Science. London: Institute of Education, University of London.

HERSH, R (1993). *Proving is Convincing and Explaining.* Educational Studies in Mathematics 24, pp. 389 – 399.

HOYLES, C. (1997). *The curricular shaping of students' approaches to proof*. For the Learning of Mathematics 17, pp. 7 – 16.

KLIEME, E. (1986). *Bildliches Denken als Mediator für Geschlechterunterschiede beim Lösen mathematischer Probleme*. In H.G. Steiner (ed.) *Grundfragen der Entwicklung mathematischer Fähigkeiten*. Köln: Aulis, pp. 133 – 151.

KLIEME, E. (2000). *Fachleistungen im voruniversitären Mathematik- und Physikunterricht*. In J. Baumert, W. Bos & R. Lehmann, R. *TIMSS – Mathematisch-naturwissenschaftliche Bildung am Ende der Sekundarstufe II*. Opladen: Leske & Budrich

KOEDINGER, K.R. (1998). *Conjecturing and argumentation in high-school geometry students*. In R. Lehrer & D. Chazan (Eds.). *Designing learning environments for developing understanding of geometry and space*. Mahwah, NJ: Erlbaum.

KOEDINGER, K. R. & ANDERSON, J. R. (1990). *Abstract planning and perceptual chunks: Elements of expertise in geometry.* Cognitive Science, 14, pp. 511 – 550.

KUHN, D. (1989). *Children and adults as intuitive scientists*. Psychological Review, 96, pp. 674 – 689.

MOORE, R. C. (1994). *Making transition to formal proof*. Educational Studies in Mathematics, 27, pp. 249 – 266.

National Council of Teachers of Mathematics (Ed.) (2000). *Principles and Standards for School Mathematics*. Reston, VA: NCTM.

REISS, K., & ABEL, J. (1999). *Die Diagnose deklarativen Wissens mit Hilfe von Concept Maps.* In H. Henning (Hrsg.), *Mathematiklernen durch Handeln und Erfahrung.* Oldenburg: Bültmann & Gerriets, pp. 175 - 184.

SENK, S. (1985). *How Well Do Students Write Geometry Proofs?* Mathematics Teacher 78 (6), pp. 448 – 456.

STUMPF, H. & FAY, E. (1983). *Schlauchfiguren.* Göttingen: Hogrefe.

USISKIN, Z. (1987). *Resolving the continuing dilemmas in school geometry*. In *Learning and Teaching geometry, K-12.* Reston VA: NCTM, pp. 17 – 31.

Adresse:
Prof. Dr. Kristina Reiss, Dr. Aiso Heinze
Fachbereich Mathematik
Carl von Ossietzky-Universität Oldenburg
Postfach 2503
26111 Oldenburg, Germany
E-mail:kristina.reiss@uni-oldenburg.de
E-mail:aiso.heinze@uni-oldenburg.de

Prof. Dr. Eckhard Klieme
Deutsches Institut für Internationale
Pädagogische Forschung (DIPF)
Schloßstraße 29
60486 Frankfurt am Main, Germany