

Birgit Burchartz; Martin Stein, Münster

REASONING AND PROBLEM SOLVING PROCESSES OF PRIMARY AND SECONDARY STUDENTS WORKING ON TASKS WITH A GOAL THAT CANNOT BE REACHED

Abstract

The project *pupils work on problems with a goal which cannot be reached* („invisible wall project“) analyses problem solving processes of children in grades 3-4 and 8-9. The central idea of the research is to use sets of tasks which are all *unsolvable* which means they have a goal which can not be reached. The unsolvability, however, is of a kind which can be understood even by younger children. In our case, the tasks are geometrical puzzles which have to be covered with given tiles, and arithmetical tasks (for instance: Find exactly for different numbers out of {1, 2, 3, 4, 5} which have the sum 9). To the unprepared problem solver, the unsolvability is like an „invisible wall“, against which he bounces in his attempt to tackle the task.

On the basis of altogether 280 fully transcribed interviews, we analyse the pupils' *reasoning behavior* and analyse the *structures of problem solving processes* of primary and lower secondary pupils.

1. Introduction

In the history of mathematics as well as mathematics teaching, *problem solving* always has played an important role, since all creative mathematical work demands actions of problem solving. So, it is not astonishing that problem solving has been analysed from many different points of view, and in many different fields. Surveys of the relevant literature can be found in Schoenfeld 1983, Orehovec 1991, whereas Lester 1994 gives an overview of the development of the field from 1970 to 1994.

Despite the large number of studies on „problem solving“, and despite a large quantity of studies about „heuristics of problem solving“ or „heuristic techniques“, there are not so many studies which try to describe or analyse individual problem solving processes and their mechanisms *in full*. To give some examples for what is meant here, some studies from Rutgers University shall be mentioned (see, e.g., Goldin 1984, Goldin/Landis 1985, Maher/Davis/Alston 1991, and Maher/Speiser 1997). Another recent example which shows how complex such analyses can become is DeBellis 1996.

The *invisible wall project* - which has been running since 1992 - is placed in those less crowded areas of research on problem solving. The focus of the research is on such components of problem solving ability which are *not* subject of mathematics lessons. As a consequence, the project does not deal with word problems and other „classical“ problems.

The central idea of the research is to use sets of tasks which are all *unsolvable* which means they have a goal which can not be reached. The unsolvability, however, is of a kind which can be understood even by younger children (e.g.: try to find exactly 4 different numbers out of the set {1, 2, 3, 4, 5} which add up to 9). To the unprepared problem solver, the unsolvability is like an „invisible wall“, against which he bounces in his attempt to tackle the task.

Since there *is* a solution to the task - the insight that it can not be done, and why - in the following text we prefer the wording *impossible task* to *unsolvable task*.

In the *first phase* of the project, elementary components of problem solving behaviour which

are actually used by younger children were identified. The results were described in three papers: Stein 1995 describes *gestalt aspects* of problem solving, Stein 1996 describes *problem solving techniques* which are used even by primary students, and Stein 1999 describes aspects of *reasoning and logical thinking* with respect to problem solving. The empirical design of all of those studies was *qualitative*, using transcripts of videotapes to describe the observed phenomena.

In the present paper we report about some of the results of the *second phase* of the project which was carried out in the years 1996 to 1998. This study used mainly quantitative empirical methods. It was supported by *Deutsche Forschungsgemeinschaft* - German Research Foundation). We ask:

- (1) What are the *problem solving abilities* of primary and lower secondary students with respect to our tasks?
- (2) Are there differences in the *structure of the problem solving processes* of primary and lower secondary students?

2. Design of the study

2.1 Tasks

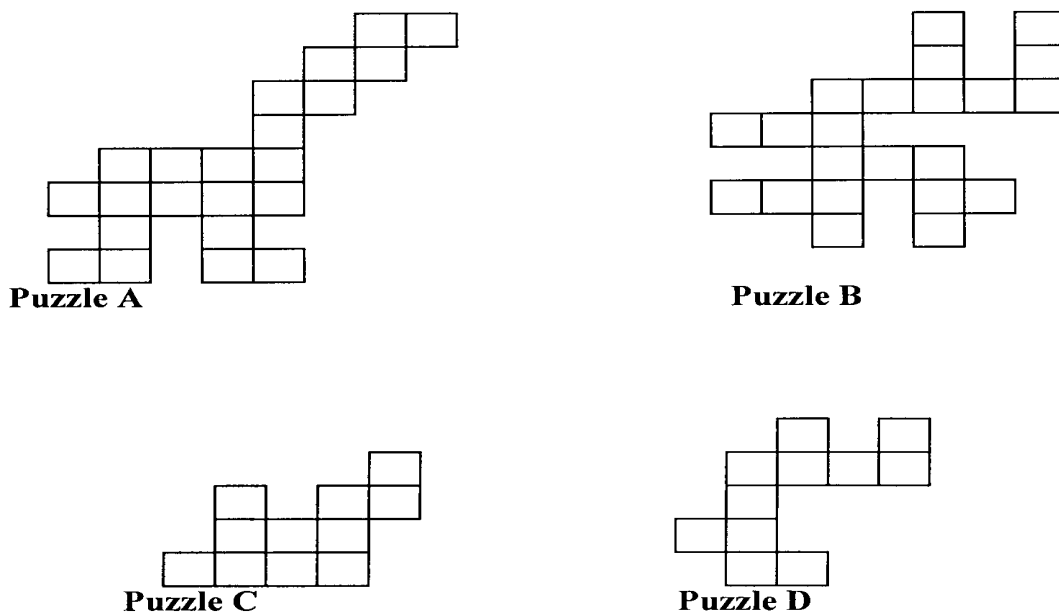


Figure 1: Small puzzles

We work with *puzzles* and *arithmetical tasks*. First we describe the *puzzles*. We use the word *puzzle* since the tasks remind us of *jigsaw-puzzles*. The word is *not* used as in *Tower-of-Hanoi-puzzle* or *missionary-and-cannibal-puzzle*. The idea for using impossible puzzles (in our sense) goes back to Jeffery (1978).

There are two „big puzzles” and two „small puzzles”.

The puzzles are drawn on cardboard. They shall be covered with tiles which are cut out from cardboard.

The small puzzles shall be covered exactly with the following tiles (each is given as often as shown).

The big puzzles shall be covered exactly with the following tiles (each is given once, the tiles have different colours).

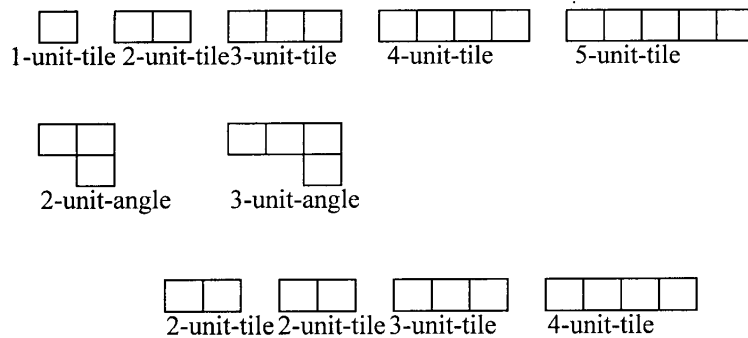


Figure 4: Puzzle tiles big puzzles

There are three *arithmetical tasks*.

Given are round tokens (each only once), bearing the numbers 1, 2, 3, 4 and 5 (task 1) resp. 1, 2, 3, 4, 5, 6 (task 2). The problem solver gets a pattern which informs him/her how many tokens he/she may use. In the first impossible task the pattern shows 4 circles. Consequently, the problem solver shall lay down 4 tokens (for each circle one token), *which sum up to 9* (This task will be called *sum-9-task*).

Figure 5: Sum-14-task



In the *sum-5-task*, there are three circles in the pattern, and the numbers to choose from are in {1, 2, 3, 4}. In the *sum-14-task*, there are five circles in the pattern, and the numbers to choose from are in {1, 2, 3, 4, 5, 6}.

2.2 Some notes on task construction

We have discussed the principles of task construction as well as the reasons for choosing unsolvable tasks in many other publications (see, e.g., Burchartz & Stein, 1998), so only the most important reason for choosing unsolvable tasks shall be mentioned here: *they give many occasions for reasoning*. Since the goal of the task can not be reached, pupils see a *need to give reasons* for their „failure“. Thus, the tasks contain an *implicit request to give (good) reasons*.

2.3 Subjects and task presentation

Due to an underlying „teacher variable“, pupils who come out of one class sometimes behave quite similar. To make sure that this effect does not spoil the value of the study, we did not make more than 5 interviews in one class. The tasks had to be solved by pairs of pupils. Each pair of pupils got 4 tasks, the first two of them always were solvable. The pupils either had to

work on the arithmetical tasks, or on the puzzles.

With the puzzles, each pair of pupils got one big and one small puzzle. The order (big puzzle first, or big puzzle as second task) was randomly assigned to the pairs of pupils.

With the arithmetic tasks, each pair of pupils got one big and one small task. „Big” and „Small” had different meanings for primary and lower secondary students: in primary classes, the number-5-task was the small one, the number-9-task was the big one. In lower secondary classes the number-9-task was the „small” one, the number-15-task was the „big” task. The order (big task first, or big task second) was randomly assigned to the pairs of pupils.

The interviews were taken from summer 1996 to winter 1997. 17 primary schools and 12 lower secondary schools took part (some of them with more than one class). Altogether, interviews taken from 140 groups of students were fully transcribed and analysed in the study (this makes 280 transcripts).

	Puzzles	Arithmetical tasks
Primary schools (grades 3 / 4)	82	50
Lower secondary schools (grades 7 / 8)	90	58

Table 1: Numbers of interviews

2.4 Interview schema

As has been said, all problems are solved by groups of two pupils. The pupils are informed that some of the tasks can be solved, other tasks not, and that they have to find out what is the case and why.

Since every question of the interviewer influences the problem solving process, the interviewers are restricted to a very limited set of interactions with the pupils. Most of the time, the interviewers *only watch the pupils work*. When the pupils *say* that the task can not be solved, or *show some „suspicion”* the interviewers ask „*and if you try it a different way?*” (This is interviewer question I1).

If the pupils refuse to proceed, the interviewer asks them to explain why they think that the task can not be done. In most cases, however, pupils resume work. Again, when they *say* that the task can not be solved, or *show some „suspicion”* the interviewers ask „*perhaps there are other possibilities?*” (This is interviewer question I2). If the pupils refuse to proceed, the interviewer asks them to explain why they think that the task can not be done.

In most cases, however, the pupils resume work again. The next time they demonstrate that they are not willing to proceed, the interviewer asks „*So what do you think: can the task be solved, or not?*”. (This is interviewer question I3).

I3 is omitted if the pupils by themselves say that they refuse to work on. Anyway, the pupils’ answer to I3 always is: „*The task cannot be done*”.

So, finally the interviewer says: „*This is right. Please explain why you think the task can not be done.*” (This is interviewer remark I4).

Only at this end of the interview the interviewers are permitted to talk with the children about their understanding of the situation and about the process of solution. This phase of the interview is called *explanation phase*.

The interviews are filmed with a video camera. The videos of all interviews were checked by

one of the authors. Interviews which did not meet the above described standards of non-interaction, were eliminated from the study and not analysed. The final number stated above is the number of valid interviews.

2.5 Transcription

The final *transcription* of the video has the character of a script for a movie which allows to replay the interview. Every spoken word and every action with the puzzle pieces respective number tokens is protocolled.

The coding was partly done by student teachers as part of their thesis for the so called *states exam*, partly by paid staff. Since the transcripts of videos are the basis of all following analyses, utmost care was observed to get precise transcripts, using a system of checks and double checks.

3. Category systems

3.1 Geometrical tasks, first category system: grading pupils' answers

The interviews starts with the pupils' answer to interviewer question I3: „*The task cannot be done*”. Interviewer question I4 then is: „*This is right. Please explain why you think the task can not be done.*”

The following coding system assesses the pupils' answers to this question. The *assessment only considers the outspoken words and the final covering of the puzzle*. Actions of the pupils are not considered, and the process which led to the answers is as well *not* looked at.

Students' answers were classified by three categories.

- I: *Complete proof* (for instance: „the 5-unit-tile and the 4-unit-tile have this unique position, and then we have no space left for the 3-unit-tile”)
- II: *Demonstration of the way the problem was solved*
„If you cover it this way, you cannot place the 5-unit-tile”)
- III: *Naming isolated facts* (for instance: „This can't be done since the parts are too long”)

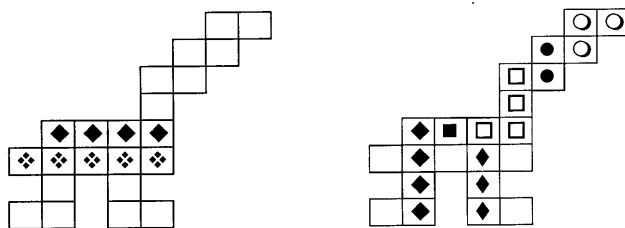


Figure 3: Illustrations for category I (left) and category II (right)

The coding system was developed by B. Burchartz, who also did all the coding. The reliability of coding was – after an introduction to the coding method – tested independently by M. Stein. A set of 20 randomly chosen answers was coded. There were no differences in both coders' results.

Though one might presume that a proof by cases (Puzzle B, puzzle D) is more complicated than a linear proof, there is actually no difference at all in the answers in both school levels. For reasons of printing space, we only give one example: primary students' solutions of the *big puzzles* were graded as follows:

Category I	:	puzzle A: 2	puzzle B: 3
Category II	:	puzzle A: 9	puzzle B: 6
Category III	:	puzzle A: 10	puzzle B: 11

From this point of the study, we therefore chose to put all results of puzzle A and puzzle B (resp. Puzzle C and D) together. In future, we shall only use the notion *small puzzles* (C and D) and *big puzzles* (A and B).

3.2 Geometrical tasks, second category system: grading answers *and* problem solving process

The category system of section 3.1 only regards the spoken words of the students. Pupils' *actions* during the final explanation phase are *not considered*. Information contained in the *problem solving gets lost*. To overcome this, B. Burchartz developed a second category system. Coding with respect to the new category system demands an analysis of the fully transcribed protocols. We have four categories:

- I: *Complete proof* (for instance: „the 5-unit-tile has this unique position, and then we do not have enough 2-unit-tiles to fill the gap”).
For this category, only the final explanation phase is used.
- I*: *The spoken words form an incomplete proof. The non-verbal parts of the communication, however, fill in the gaps.* Considering those non-verbal parts, the proof is complete.
For this category, only the final explanation phase is used.

If the pupils' answer is neither in category I or I*, the full transcript up to the explanation phase is analysed. There are two remaining categories.

- L: *Reasoning using tacit assumptions.* Pupils in this category refer in their explanation to situations they have constructed recently, or perhaps some time earlier, in the problem solving process. During the explanation phase, the argumentation itself is not completed by actions.
- ?: *Not understandable*

Again, coding was very reliable. The reliability of coding was – after an introduction to the coding method – tested independently by M. Stein. A set of 20 randomly chosen answers was coded. There was only 1 difference in both coders' results.

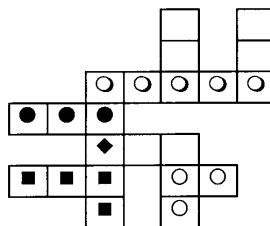


Figure 4: Example for category L

Examples

Example for category L The situation is as shown in the Figure 7. The pupil says: „Yes, if you put it down that way...“... points at the puzzle... „...then this would have to be cut in the middle“, and points at the 4-unit-tile.

*Example for category I** (see Figure 8) The pupils start giving their reasons. The situation is as shown in picture 1 (By using different symbols on the puzzle pieces (●, ○, ■) we intend to make the positions of the puzzle pieces easier to detect for the reader).

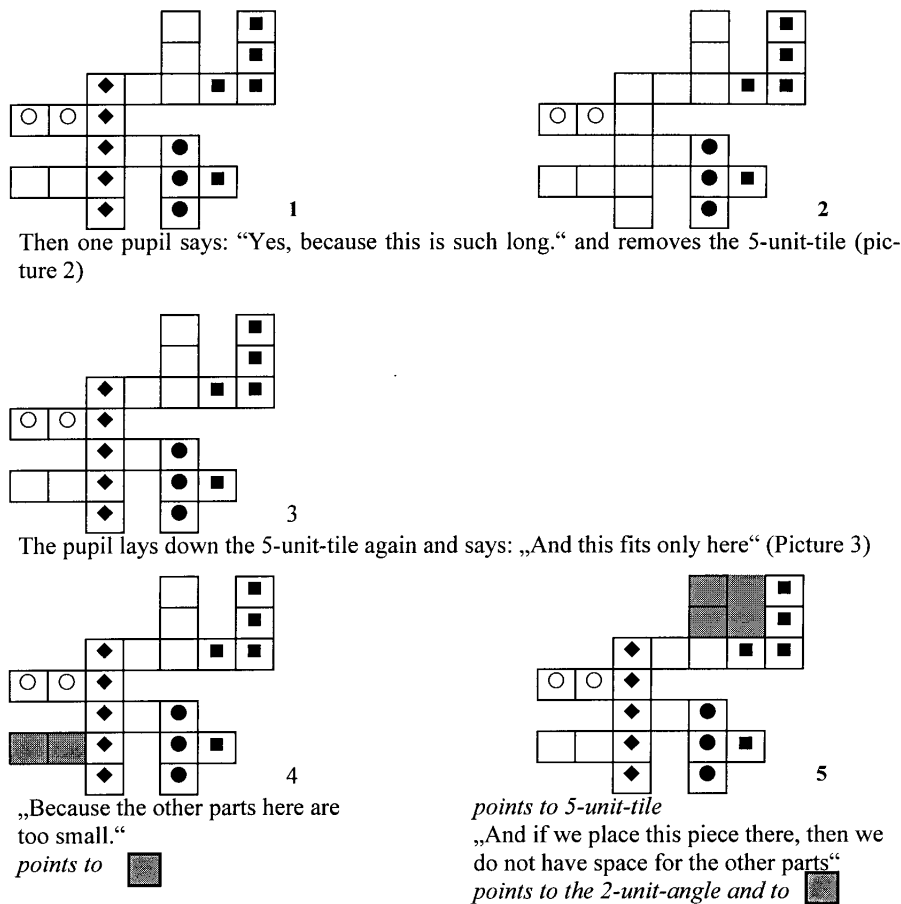


Figure 8: Example for category I*

3.3 Coding system for arithmetical tasks

The category system for the arithmetical tasks is quite similar to the first system of the geometrical tasks. We have four categories:

- I: *Complete proof*: the sum of the lowest numbers, or the full set of possible combinations of numbers is given as proof.
- II: *Demonstration of the way the problem was solved*. („No combination made 9”; „If I put down the numbers this way, then it is too much”)
- III: *Naming isolated facts*. („The numbers are too big”; „There is a number missing”)

IV: *Talking round the matter.* („It can't be done because 14 cannot be divided by 5")

4. Results

4.1 Geometrical tasks – first coding system

A comparison of students answers showed for both types of puzzles *that students in lower secondary give „better“ answers than students in primary schools.* The difference is significant only for the big puzzles. Statistical analysis used a test by Raatz for grouped ordinal data, as shown in Lienert (1973). The level of Significance was 5%.

Small Puzzles			Big Puzzles		
(not sig.)	P	S	(sig.)	P	S
I	13	23	I	5	20
II	18	5	II	15	6
III	10	17	III	21	19

Table 2 Comparison primary vs. lower secondary students

The performances of both groups with respect to big vs. small puzzles differ significantly (S: Small puzzles; B: Big puzzles): *primary students solve the small puzzles significantly better than the big ones.* Secondary students show no differences in solving small vs. big puzzles.

4.2 Geometrical tasks – second coding system

The main result of the re-coding using the second category system is that - *using the process information contained in the transcripts - we get a totally new impression about the pupils' abilities.* The following table shows the migration from old categories to new categories. Category I is not listed because category I remains untouched.

		Primary	Secondary	
	Category I*:	30	11	
Category II	↗			
	Category L :	2	--	
	↘			
	Category ?:	1	--	
Category III				
	↗	Category I*:	2	7
	Category L :	18	25	
	↘			
	Category ?:	11	4	

Table 3 Migration from old to new categories

For a *statistical analysis*, we cannot assume an ordinal ordering from I to I* to L to ?:

- We can not say that I is „better“ than I*.
- There is a big step from I* to L since I* uses the final explanation phase of the transcript only, whereas L uses the full transcript.

- We cannot say that L is „better” than ?. „L” means that the coder was able to understand the „hidden information” in the transcript, „?” means, the coder was not able to do so. As a consequence, the distinction between L and ? may be more a problem of the coder than of the problem solver.

There *is*, however, a clear distinction between I and I* vs. L and ? since I and I* only use the verbal and non-verbal information given during the explanation phase. In the following, we test I/I* against L/?. Having the above table in mind, it is not astonishing that there is now far less distinction between the groups of primary and secondary students.

Small Puzzles			Big Puzzles		
	P	SI		P	SI
I / I*	30 (13/17)	34 (23/11)	I / I*	20 (5/15)	27 (20/7)
L / ?	11 (7/4)	11 (11/0)	L / ?	21 (13/8)	18 (14/4)

Table 4 Results for new categories

For the small puzzles, there is no difference at all between groups, for the big puzzles, the group of secondary students is a bit „above” the group of primary students, but this difference is not significant.

Of special interest is, *how big the contrast* is for primary students, if we compare „good explanations against the rest” in the old category system (i.e.: category I vs. II/III) with „good explanations against the rest” in the new category system (i.e.: categories I/I* vs. L/?), bearing in mind that good explanation in the new category system just means that we look at what the pupil *does* while he is giving his final explanation.

Small Puzzles			Big Puzzles		
	Old system	New system		Old system	New system
„good”	13	30	„good”	5	20
„bad”	28	11	„bad”	36	21

Table 5 Comparison of old and new categorisation system

4.3 Arithmetical tasks

The differences between primary and secondary students are not very marked. We think that the following tables make a statistical analysis obsolete (i.e.: no explanation need be given):

Primary			Secondary		
	Sum-5	Sum-9		Sum-9	Sum-14
I	6	4	I	8	6
II	10	10	II	7	7
III	7	10	III	13	15
IV	1	--			
n.e.	1	1			

Table 6: Results for arithmetical tasks

4.4 Didactical consequences

The results of sections 4.1 and 4.2 are of special interest for our understanding of problem solving and reasoning in primary education. Taking the final verbal argumentation „as it is”, the results of the primary students are not too bad but not very impressive. The new coding now shows that the *potential* of primary students is *far higher* than can be seen in the verbally expressed answers.

In understanding the unsolvability of small resp. big puzzles, primary students' abilities are near to secondary students' abilities. Primary students

- have, however, more difficulties to express their insight verbally (this is group I*),
- or they „loose” their insight during the process of argumentation, or in their argumentation they tacitly assume that the interviewer knows all they know (this is category L).

The result of section 4.3 shows that reasoning about unsolvability is not impossible for younger children. Teachers have to take care, however, that the structure of the needed argumentation is not too complex.

5. Process Analysis

Since the abilities of primary and secondary students are such close together with regard to the verbal explanation of the unsolvability of our arithmetical tasks, it is of interest to have a look at the problem solving *processes*. (For the *puzzles*, the respective analysis will be contained in Burchartz/Stein 2001.).

To be able to compare the processes, we developed so called *profiles*. Figure 9 shows one such profile.

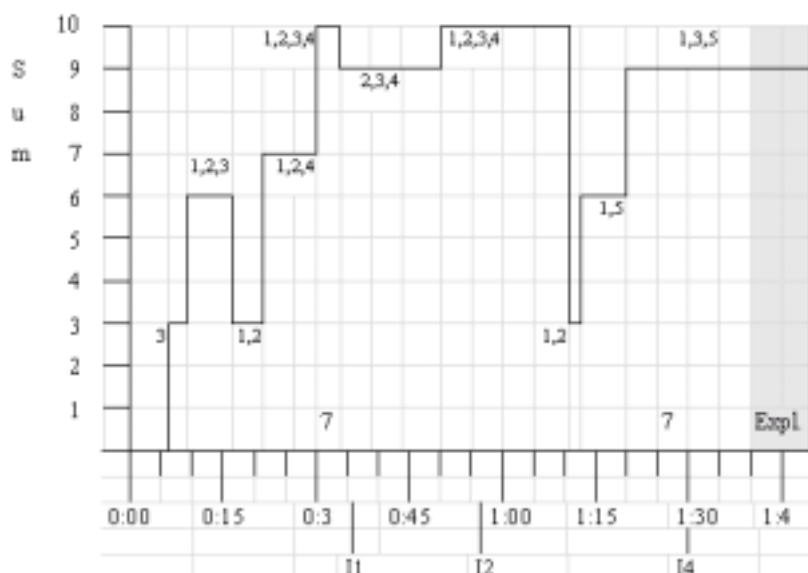


Figure 9: Sums laid by the pupils: how they were built and how long they remained laid down. Expl.: Explanation phase

In those profiles, we can now count the number of building-up-sequences. A building-up-sequence *ends* when at least two numbers are removed. The process then ends, or a new building-up-sequence starts (The profile in Figure 9 has two building-up-sequences).

We can now compare the number of building-up-sequences for primary and secondary students. *Secondary students show significantly less such sequences*. To simplify the analysis, the data are grouped.

	Primary	Secondary
„short“: 0-3 building-up-sequences	21	38
„medium“: 4-6 building-up-sequences	21	15
„long“: 7 or more building-up-sequences	8	3

Significant

Table 7 Numbers of building-up-sequences

6. Summary

6.1 What are the *problem solving abilities* of primary and lower secondary students with respect to our tasks?

For both types of tasks – the geometrical as well as the arithmetical – our results show that *even primary students are able to understand the concept of unsolvability*. The *number-9-task* shows that *both groups are not far away of each other if – once the insight was gained – the amount of reasoning is not too high and the structure of the expected answer is not too complex*.

The same holds for the geometrical tasks: for the *small puzzles* the needed chain of argumentation is not very long, and though there is some difference between the quality of answers in the primary group vs. the group of secondary students, it is not significant. 30% of the primary students give a full explanation for the unsolvability of the task.

In case of the *big puzzles*, however, a somewhat longer chain of argumentation is needed. The results of the coding using the first category system show that complete correct explanations of the unsolvability are nearly out of reach for primary students.

The second category system now shows that the *potential* of primary students is *far higher* than can be seen in the verbally expressed answers. In *understanding* the unsolvability of small resp. big puzzles, primary students' abilities are near to secondary students' abilities (For related results see Anderson et al. 1997 and McCoy 1994). We see

- that many younger students tend to use *actions as part of their reasoning* (this is the case for those students who migrate to category I*). If we accept those actions as part of the final argumentation, we get a complete and logically valid argument.
- that there is another big group which uses „hidden arguments“ for which they possibly assume that they are known to the interviewer (this is the case for those students who migrate to category L).

6.2 Are there differences in the structure of the problem solving processes of primary and lower secondary students?

We have seen that for both types of tasks students' reasoning abilities in primary and secondary education are quite near to each other if we accept *actions* as part of the argumentation. The question now arises whether there are any differences between the way primary vs. secondary students work on the tasks.

For the arithmetic tasks we found that younger pupils run through sequences of problem solving activities which are more complex than those of the older students, in the following meaning: primary students have to run through the same sequences of actions as the secondary students do, *but more often*. There are significantly more building-up-sequences in the group of primary students. By this process of building up and then reducing the sums they seem to

build up a good „feeling” for the impossibility of the task.

6.3 Consequences for mathematics teaching

The results which were reported here have some important consequences for primary education:

- (1) We *underestimate* the reasoning abilities if we only regard their verbally formulated responses to a task. (Though this was not analysed in this research, we are very sure that this holds for *written* answers as well.) *Actions* can form an important part of a child’s attempt to communicate an argumentation.
- (2) When reasoning about their findings, *children often use tacit assumptions*. To understand those assumptions in full, the teacher has to observe the full length of the child’s problem solving process.
- (3) Children have to build up a feeling for the structure of the problem by approaching it several times. It is important to *give them enough time* to work on a problem.

References

- ANDERSON, R. C., CHINN, CL., CHANG, J., WAGGONER, M., & YI, H. (1997). *On the logical integrity of children's arguments*, in: *Cognition and Instruction*, 15, p. 135-167.
- BELL, A. (1976). *The learning of general mathematical strategies*. Dissertation, Nottingham 1976
- BURCHARTZ, B.; STEIN, M.(1998). *The invisible wall project: Problem solving processes of pupils who work on problems with a goal which can not be reached*, in: *Proceedings of MERGA 21*, Brisbane
- BURCHARTZ, B.; STEIN, M. (2001). *The invisible wall project: Reasoning and problem solving processes of primary and lower secondary students*. Under review
- BURKHARDT, H. e.a. (eds.) (1984). *Problem Solving - a World View*. Proceedings of Problem Solving Theme Group, ICME 5
- DAVIS, G.; PEPPER, K. (1992). *Mathematical Problem solving by pre-school children*, in: *Educational Studies in Mathematics*, 23, p. 397 – 415
- DEBELLIS, V.; GOLDIN, G. A. (1993). *Analysis of interactions between affect and cognition in Elementary school children during problem solving*, in: *Proceedings of the 15th annual meeting of PME-NA*, vol. 2, p. 56-62
- DEBELLIS, V. (1996). *Interaction between affect and cognition during mathematical problem solving: a two year case study of four elementary children*. Dissertation, Rutgers University
- GOLDIN, G. A. (1984). *The structure of heuristic processes*, in: *Proceedings of the 6th annual meeting of PME-NA*, p. 184-188
- GOLDIN, G.A., LANDIS, J.H. (1985). *A problem solving interview with Stan (age 11)*, in: *Proceedings of the 7. annual meeting of the North American Chapter of PME*, 1985, p. 100 – 105
- JEFFERY, R. (1978). *A study of the generalisation and explanation strategies of 10 and 11 year old children in Mathematics*, Thesis, Nottingham
- LESTER, F. K. (1994). *Musings about Mathematical Problem Solving Research: 1970 - 1994*, in: *Journal for Research in Mathematics Education*, 25, p. 660-675

- LIENERT, G.A. (1973). *Verteilungsfreie Methoden in der Biostatistik*, vol. 1. Verlag Anton Hain, Meisenheim am Glan
- MAHER, C. A.; DAVIS, R. B.; ALSTON, A. (1991). *Brian's representation and development of mathematical knowledge: a 4-year study*, in: *Journal of Mathematical Behavior*, 10, p. 163 – 210
- MAHER, C. A.; SPEISER, R. (1997). *How far can you go with block towers*, in: *Journal of mathematical behavior*, 16, p. 125 – 132
- MCCOY, L. P. (1994). *Mathematical Problem-Solving Processes of Elementary Male and Female Students*, in: *School Science and Mathematics*, 94(5), p. 266-270
- OREHOVEC, J.P. (1983). *Implications of the development of mathematical problem solving. 1894-1983*. Dissertation, Michigan State University, 1991
- SCHOENFELD, A. H. (1983). *Problem Solving in the Mathematics Curriculum. A Report, Recommendations, and an Annotated Bibliography*. The Mathematical Association of America, Committee on the Teaching of Undergraduate Mathematics, MAA Notes Number 1
- SCHOENFELD, A. H. (1985 a). *Mathematical Problem Solving*. Academic Press, Orlando
- SCHOENFELD, A. H. (1985 b). *Making sense of „out loud“ problem-solving protocols*, in: *Journal of Mathematical Behavior*, 4, p. 171-191
- STEIN, M. (1995). *Elementare Bausteine von Problemlöseprozessen: Gestaltorientierte Verhaltensweisen*, in: *math. didact.*, 18; vol. 2, p. 59 – 84
- STEIN, M. (1996). *Elementare Bausteine der Problemlösefähigkeit: Problemlösetechniken*, in: *JMD*, 17, vol. 2, p. 123 – 146
- STEIN, M. (1999). *Elementare Bausteine der Problemlösefähigkeit: logisches Denken und Argumentieren*, in: *Journal für Mathematikdidaktik*, 20, No. 1, S. 3 – 27

Prof. Dr. Martin Stein
Didaktik d. Mathematik
Einsteinstr. 62
48149 Münster