An ICT environment to assess and support students' mathematical problem-solving performance in non-routine puzzle-like word problems

Paper to present in TSG19 at ICME11
Theme (d): *Research and development in problem solving with ICT technology*

Angeliki Kolovou*
Marja van den Heuvel-Panhuizen*#
Arthur Bakker*
Iliada Elia[x]

*FIsme, Utrecht University, the Netherlands
#IQB, Humboldt University Berlin, Germany
[x]Department of Education, University of Cyprus

Summary

This paper reports on a small-scale study on primary school students' problem-solving performance. In the study, problem solving is understood as solving non-routine puzzle-like word problems. The problems require dealing simultaneously with multiple, interrelated variables. The study employed an ICT environment both as a tool to support students' learning by offering them opportunities to produce solutions, experiment and reflect on solutions, and as a tool to monitor and assess the students' problem solving processes. In the study, 24 fourth-graders were involved from two schools in the Netherlands. Half of the students who belonged to the experimental group worked in pairs in the ICT environment. The analysis of the students' dialogues and actions provided us with a detailed picture of students' problem solving and revealed some interesting processes, for example, the bouncing effect that means that the students first come with a correct solution and later give again an incorrect solution. The test data collected before and after this "treatment" did not offer us a sufficient basis to draw conclusions about the power of ICT environment to improve the students' problem-solving performance.

1. INTRODUCTION

Problem solving is a major goal of mathematics education and an activity that can be seen as the essence of mathematical thinking (Halmos, 1980; NCTM, 2000). With problems tackled in problem solving typically defined as non-routine (Kantowski, 1977), it is not surprising that students tend to find mathematical problem solving challenging and that teachers have difficulties preparing students for it. Despite the growing body of research literature in the area (Lesh & Zawojewski, 2007, Lester & Kehle, 2003, Schoenfeld, 1985), there is still much that we do not know about how students attempt to tackle mathematical problems and how to support students in solving non-routine problems.

In order to get a better understanding of Dutch primary school students' competences in mathematical problem solving, the POPO study started in 2004. In this study, 152 fourth-grade students who are high achievers in mathematics were administered a paper-and-pencil test on non-routine problem solving. In a few items, students were asked to show their solutions strategies. The results were disappointing. Students did not show a high performance in problem solving, despite their high mathematics ability (Van den Heuvel-Panhuizen, Bakker, Kolovou, & Elia, in preparation). Although the students' scribbling on the scrap paper gave us important information about their solution strategies, we were left with questions about their solution processes. Moreover, after recognizing that even very able students have difficulties with solving the problems, we wondered what kind of learning environment could help students to improve their problem solving performance. The POPO study thus yielded a series of questions. To answer these questions we started the iPOPO study which – in accordance with the two main questions that emerged from the POPO study – implied a dual research goal.

First, the iPOPO study aimed at gaining a deeper understanding of the primary school students' problem solving processes, and, second, it explored how their problem-solving skills can be improved. For this dual goal of assessing and teaching, the study employed ICT both as a tool to support students' learning by offering them opportunities to produce solutions, experiment and reflect on solutions, and as a tool to monitor and assess the students' problem solving processes. In particular, we designed a dynamic applet called *Hit the target,* which is based on one of the paper-and-pencil items used in the POPO study. Like several of these items, it requires students to deal with multiple, interrelated variables simultaneously and thus prepares for algebraic thinking.

This paper focuses on the following two research questions: Which problem-solving strategies do fourth-grade students deploy in this *Hit the target* environment? Does this ICT environment support the students' problem-solving performance?

2. THEORETICAL BACKGROUND

2.1. Mathematical problem solving

The term "problem solving" is used for solving a variety of mathematical problems, ranging from real-life problems to puzzle-like problems. Our focus is on the latter. We consider problem solving as a cognitive activity that entails strategic thinking, and that includes more than just carrying out calculations. An episode of problem solving may be considered as a small model of a learning process (D'Amore, & Zan, 1996). In problem solving, the solution process often requires several steps. First the students have to unravel the problem situation. Subsequently, they have to find a way to solve the problem by seeking patterns, trying out possibilities systematically, trying special cases, and so on. While doing this they have to coordinate relevant mathematical knowledge, organize the different steps to arrive at a solution and record their thinking. In sum, in our view problem solving is a complex activity that requires higher order thinking and goes beyond standard procedural skills (cf., Kantowski, 1977).

An example of a mathematical problem used in the POPO study is shown in Figure 1. Someone who knows elementary algebra might use this knowledge to find the answer to this problem by, for example, solving the equation $2x - 1(10 - x) = 8$. Fourth-grade students, however, have not yet learned such techniques, but can still use other strategies such as systematic listing of possible solutions or trial and error. Grappling with such problems might be a worthwhile experiential base for learning algebra in secondary school (cf., Van Amerom, 2002).

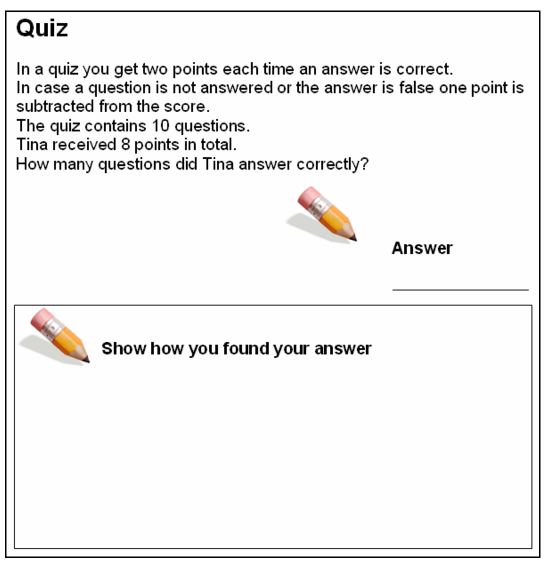

Figure 1: Problem used in the POPO Study

Within the complexity that characterizes problem-solving activity, D'Amore and Zan (1996) identify the involvement of three interrelated discrete variables, as follows: the subject who solves the task; the task; and the environment in which the subject solves the task. This study primarily focuses on the third variable, referring to the conditions, which may help a subject to improve his problem solving abilities.

The research questions stated in Section 1 address two different aspects that are closely related: monitoring learning and supporting that learning. We have chosen to use ICT for both of these aspects, because – as Clements (1998) recognized – ICT (1) can provide students with an environment for doing mathematics and (2) can give the possibility of tracing the students' work.

2.2. ICT as a tool for supporting mathematical problem solving

A considerable body of research literature has shown that computers can support children in developing higher-order mathematical thinking (Suppes, 1966; Papert, 1980; Clements & Meredith, 1993; Sfard & Leron, 1996; Clements, 2000; Clements, 2002). Logo programming, for example, is a rich environment that elicits reflection on mathematics and one's own problem-solving (Clements, 2000). Suitable computer software can offer unique opportunities for learning through exploration and creative problem solving. It can also help students make the transition from arithmetic to algebraic reasoning, and emphasize conceptual thinking and problem solving. According to the Principles and Standards of the National Council of Teachers of Mathematics (NCTM, 2000) technology supports decision-making, reflection, reasoning and problem solving.

Among the unique contributions of computers is that they also provide students with an environment for testing their ideas and giving them feedback (Clements, 2000). In fact, feedback is crucial for learning and technology can supply this feedback (NCTM, 2000). Computer-assisted feedback is one of the most effective forms of feedback because "it helps students in building cues and information regarding erroneous hypotheses"; thus it can "lead to the development of more effective and efficient strategies for processing and understanding" (Hattie & Timperley, 2007, p.102). More generally, computer-based applications can have significant effects on what children learn because of "the computer's capacity for simulation, dynamically linked notations, and interactivity" (Rochelle, Pea, Hoadley, Gordin, & Means, 2000, p. 86).

This learning effect can be enhanced by peer interaction. Pair and group work with computer software can make students more skilful at solving problems, because they are stimulated to articulate and explain their strategies and solutions (Wegerif & Dawes, 2004). Provided there is a classroom culture in which students are willing to provide explanations, justifications, and arguments to each other, we can expect better learning.

2.3. ICT as a window onto students' problem solving

Several researchers have emphasized that technology-rich environments allow us to track the processes students use in problem-solving (Bennet & Persky, 2002). ICT can provide mirrors to mathematical thinking (Clements, 2000) and can offer a *window* onto mathematical meaning under construction (Hoyles & Noss, 2003, p. 325). The potential of computer environments to provide insight into students' cognitive processes makes them a fruitful setting for research on how this learning takes place.

Because software enables us to record every command students make within an ICT environment, such registration software allows us to assess their problem solving strategies in more precise ways than can paper-and-pencil tasks. Therefore, computer-based tasks as opposed to conventional paper-and-pencil means have received growing interest in the research literature for the purposes of better assessment (Clements 1998; Pellegrino, Chudowsky, & Glaser, 2001; Bennet & Persky, 2002; Burkhardt & Pead, 2003; Threlfall, Pool, Homer, & Swinnerton, 2007; Van den Heuvel-Panhuizen, 2007).

Where early-generation software just mimicked the paper-and-pencil tasks, recent research shows that suitable tasks in rich ICT environments can also bring about higher-order problem solving. For example, Bennet and Persky (2002) claimed that technology-rich environments tap important emerging skills. They offer us the opportunity to describe performance with something more than a single summary score. Furthermore, a series of studies indicated that the use of ICT facilitates the assessment of creative and critical thinking by providing rich environments for problem solving (Harlen & Deakin Crick, 2003).

By stimulating peer interaction we also expect that students will articulate more clearly their thinking than when working individually. Thus, student collaboration has a twofold role: it helps them shape and broaden their mathematical understandings and it offers researchers and teachers a nicely bounded setting in order to observe collaboration and peer interaction (Mercer & Littleton, 2007).

3. METHOD

3.1. Research design and subjects

The part of the iPOPO study described in this paper is a small-scale quasi-experiment following a pre-test-post-test control group design. In total, 24 fourth-graders from two schools in Utrecht participated in the study. In each school, 12 students who belonged to the A level according the Mid Grade 4 CITO test – in other words to the 25% best students according to a national norm – were involved. Actually, the range of the scores that correspond to level A of the Mid Grade 4 CITO test is between 102 and 151 points. In both schools, the average mathematics CITO score of the class was A and the average "formation weight" of the class and the school was 1. This means that the students were of Dutch parentage and came from families in which the parents had at least secondary education. First, of each school six students were selected for the experimental group. Later on, the group of students was extended with six students to be in the control group. These students also belonged to the A level, but, unfortunately, their average score was lower than that of the experimental group. The teacher obviously selected the more able students first.

An ICT environment was especially developed for this study to function as a treatment for the experimental group. Before and after the treatment, a test was administered as pre-test and post-test. The control group did the test also two times, but did not get the treatment in between. The quasi-experiment was carried out in March-April 2008. The complete experiment took about four weeks: in the first week the students did the test, in the second week the experimental group worked in the ICT environment and in the fourth week the students did again the test.

3.2. Pre-test and post-test

The test that was used as pre-test and post-test was a paper-and-pencil test consisting of three non-routine puzzle-like word problems, titled Quiz (see Figure 1), Ages, and Coins. The problems are of the same type and require that the students deal with interrelated variables. The test sheets contain a work area on which the students had to show how they found the answers. The students' responses were coded according to a framework that was developed in our earlier POPO study. The framework covers different response characteristics including whether the students gave specific strategy information, how they represented that strategy and what kind of problem-solving strategies they applied.

3.3. Applet used as treatment

The treatment consisted of a Java applet called *Hit the target.*[1] It is a simulation of an arrow shooting game. The screen shows a target board, a score board featuring the number of gained points, and the number of hit and missed arrows, a frame that contains the rules for gaining or loosing points, and an area in which the number of arrows to be shot can be filled in. A hit means that the arrow hits the yellow circle in the middle of the target board; then the arrow becomes green. A miss means that the arrow hits the gray area of the board; in that case, the arrow becomes red.

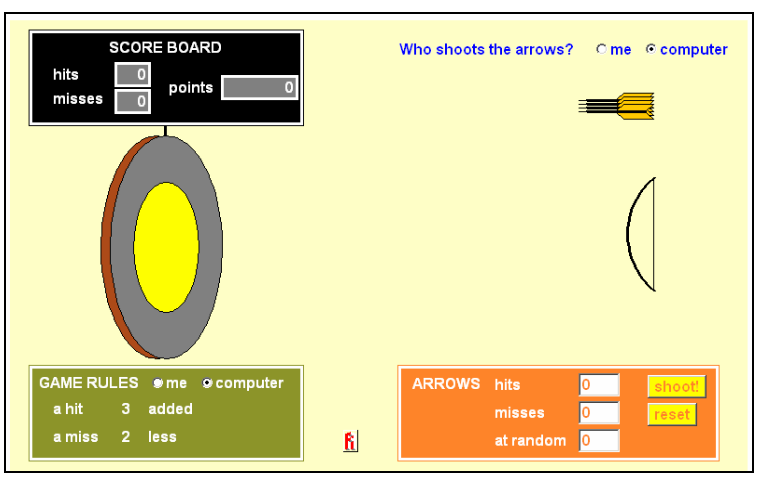

Figure 2: Screen view of applet in the computer-shooting mode

[1] The applet has been programmed by Huub Nilwik.

The applet has two modes of shooting: a player shoots arrows by him or herself or lets the computer do the shooting (see Figure 2). In case the player shoots, he or she has to drag the arrows to the bow and then draw and unbend the bow. The computer can do the shooting if the player selects the computer-shooting mode and fills in the number of arrows to be shot. Regarding the rules for gaining points there are also two modes: the player determines the rules or the computer does this. The maximum number of arrows is 150 and the maximum number of points the player can get by one shot is 1000.

As the player shoots arrows or lets the computer do so, the total score on the scoreboard changes according to the number of arrows shot and the rules of the game. The player can actually see on the scoreboard how the score and the number of hits and misses change during the shooting. The player can also remove arrows from the target board, which is again followed by a change in the total score. When the player wants to start a new shooting round, he or she must click on the reset button. The player can change the shooting mode or the rules of the game at any time during the game.

The aim of the applet is that the students obtain experience in working with variables and realize that the variables are interrelated (see Figure 3); a change in one variable affects the other variables. For example, if the rules of the game are changed, then the number of arrows should be also changed to keep the total points constant.

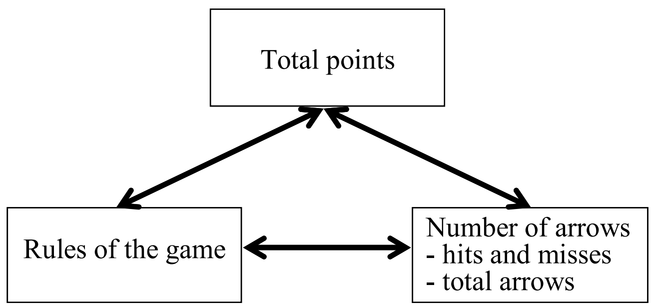

Figure 3: Variables involved

The 12 students of the experimental group worked for about 30 minutes in pairs with the applet. The pairs were chosen by the researcher in such a way that all of them would have about the same average CITO score and consisted of a boy and a girl. The dialogue between the students and their actions on the applet were recorded by Camtasia software, which captures the screen views and the sound in a video file. Scrap paper was also available to the students. Before the students started working, it was explained to them that they should work together, use the mouse in turns, explain their thinking to each other, and justify their ideas.

The work with the applet started with five minutes of free playing in which the students could explore the applet. Then, they had to follow a pre-defined scenario containing a number of directed activities and three questions (see Table 1). The first two questions (A and B) are about the arrows while the rules of the game and the gained points are known. In the third question (C), which consists of two parts, the rules of the game are unknown.

Table 1: Questions in the pre-defined scenario

Arrows	Rules	Gained points
A. How many hits and misses?	Hit +3 points, miss –1 point	15 points
B. How many hits and misses?	Hit +3 points, miss +1 point	15 points
15 hits and 15 misses	**C1. What are the rules?**	15 points
	C2. Are other rules possible to get the result 15 hits-15 misses-15 points?	

The directed activities were meant to assure that all the students had all the necessary experiences with the applet. During these activities, the students carried out a number of assignments in order to become familiar with the various features of the applet: the player-shooting mode, the computer-shooting mode, the rules of the game, and the total score. First, the students had to shoot one arrow, followed by shooting two arrows and then a few more, in order to get five arrows on the target board. Their attention was then drawn to the scoreboard; they had five hits and zero misses and their total score was zero since the rules of the game had been initially set to zero. After that, the rules were changed so that a hit meant that three points were added. Then, the students had to shoot again five arrows in both shooting modes, each resulting in a total score of 15 points. Afterward, the rule was changed again. A miss then meant that one point had to be subtracted. At this point, Question A was asked, followed by Questions B and C.

4. RESULTS

4.1. The students' problem-solving strategies in the ICT environment

All pairs were successful in answering the Questions A, B, and C. The solutions were found based on discussions and sharing ideas for solutions. In all cases, explanations were provided and the talk between the students stimulated the generation of hypotheses and solutions. However, some students provided more elaborate explanations and suggested more successful problem-solving strategies than others.

In order to identify the problem-solving strategies the students applied, we analyzed all dialogues between the students. In this paper, however, we will only discuss our findings with respect to Questions C1 and C2, which triggered the richest dialogues.

Characteristic for Question C is that the number of hits and misses, and the number of points were given, but that the students had to find the rules. All pairs were able to answer Questions C1 and C2, and most of them could generalize to all possible solutions ("It is always possible if you do one less"), albeit on different levels of generalization. The Tables 2 and 3 show which strategies the pairs used when solving Questions C1 and C2. Each pair of students is denoted with a Roman numeral. Pairs I, II, and III belong to school A, while Pairs IV, V, and VI belong to school B.

Table 2: Problem-solving strategies when solving C1

	Strategy	Pairs					
		I	II	III	IV	V	VI
		Average CITO score per pair					
		111	111	114	110	111	107
1a	Directly testing a correct solution (+2 –1 or +1 +0)	1*	1	1		1	
2a	Testing incorrect *canceling-out solution* (+1 –1)				1		
2b	Testing other incorrect solution(s)						1
3	Adapting the rules of the game until a correct solution is reached				2		2
Number of trials		1	1	1	2	1	3

* The numbers in the cell indicate the order in which the strategies were applied

When answering Question C1 (see Table 2), four out of the six pairs directly came up with a correct solution. Pair VI found the correct solution in the third trial. The most interesting strategy came from Pair IV. This pair found the correct solution in the second trial. The pair started with a canceling-out solution (+1 –1) resulting in a total score of zero and then changed the solution to get 15 points in total.

Table 3 shows that having found a correct solution in C1 did not mean that the students had discovered the general principle (or the correct solution rule) of getting "15 hits-15 misses-15 points". Even after finding the correct solution rule and generating a series of correct solutions, some students tested wrong solutions again (we could call this the "bouncing effect"). Perhaps they were not aware that there is only one correct solution rule; the difference between the number of points added for every hit and the number of points subtracted for every miss (or vice versa) should be 1, or the difference between the number of hit-points and miss-points should be 15 points. The highest level of solution was demonstrated by Pair VI, who recognized that the difference between the points added and the points subtracted should be 15 (and that explains why the difference between the number of points added for every hit and the number of points subtracted for every miss – or vice versa – should be 1). A clever mathematical solution came from the Pairs I and II. These students just used the correct solution to C1 in the reverse way to get the required result of 15 points in total.

Table 3: Problem-solving strategies when solving C2

	Strategy	Pairs					
		I	II	III	IV	V	VI
		Average CITO score per pair					
		111	111	114	110	111	107
4a	Repeating the correct solution to C1					2	
4b	Reversing the correct solution to C1 to find another correct solution (–1 +2 or –0 +1/+0 +1)	1*	1/3				
5a	Generating a correct solution rule based on testing of (a) correct solution(s) for which the difference between the number of points added for every hit and the number of points subtracted for every miss (or vice versa) is 1	2	4	6	1	4	
5b	Generating a correct solution rule based on understanding that the difference between hit-points and miss-points is 15						1
5c	Generating a general correct solution rule (“the difference of 1 also applies to 16-16-16”)			8			
6	Testing more correct solutions from a correct solution rule	3		7	2		2
2b	Testing other incorrect solution(s)	4	2	1/3/5		1/3/5	
7	Generating an incorrect solution rule (keeping ratio 2:1 or using rule +even number –odd number) based on correct solution(s)			2/4			

* The numbers in the cell indicate the order in which the strategies were applied

Besides strategies that directly or indirectly lead to a correct solution or rule, some other characteristics were found in the solution processes (see Table 4). Four pairs altered or ignored information given in the problem description. It is noteworthy that during subsequent attempts to answer Question C2, some students insisted on keeping the rules constant and changing the number of hits and misses in order to get a total of 15 points. Pair V, for example, changed the problem information (15 hits and 15 misses) and started C2 with trying out the solution 1 hit is 15 point added and 1 miss is 15 points subtracted. The total score then became zero; subsequently, they set the number of hits to 30 and the number of misses to 15, which resulted into a high score. Even though at that point the researcher repeated the correct problem information, the students ignored it persistently. In their third attempt, they changed the number of hits and misses to 1 and 0 respectively and the total score became 15 instead of the reverse (15 hits and 15 misses resulting in 15 points). Only when the researcher repeated the question they considered the correct information and tried out the solution +4 –2 with 15 hits and 15 misses. However, the total score was 30 points and they suggested doubling the number of misses to 30 so that the number of total points would be halved. This is clearly an example of a wrong adaptation. Another example is from Pair VI. After having +3 and –1 as the rule of the game, resulting in a total of 30 points, the students change the number of hits into 10 in order to get 15 points as the result but forgetting that the number of hits should be 15.

Table 4: Other characteristics of the solution processes

Characteristics	C1						C2					
	Pairs						Pairs					
	I	II	III	IV	V	VI	I	II	III	IV	V	VI
Altering or ignoring information			X			X	X				X	
Exploring large numbers (≥1000)								X	X	X	X	X

Another characteristic of the solution processes was testing rules including large numbers. Four of the six pairs tried out numbers bigger than 1000. These explorations all took place when answering the second part of Question C. The students found this working with large numbers quite amusing, since they then could get a large amount of total points. That the students worked with numbers larger than 1000 was quite remarkable, because it was not possible to fill in numbers of this size in the applet. Consequently, the students had to work out the results mentally. It is also worth noting that some students understood that one could go on until one million or one trillion (Pair IV). This means that several students knew that there are infinite solutions, as it was made explicit by one pair (see Pair II). Furthermore, most of the students used whole numbers and no one used negative numbers. In one occasion, a student (from Pair II) suggested adding 1½ points for a hit, but the applet does not have the possibility to test solutions with fractions or decimals.

Observing the students while working on the applet revealed that the students demonstrated different levels of problem-solving activity. For example, there were students that checked the correctness of their hypotheses by mental calculation, while others just tried out rules with the help of the applet. None of them questioned the infallibility of the applet; when they used the applet after they had found out that a rule was wrong, they did this to make sure that they were *really* wrong. Furthermore, the students also showed differences in the more or less general way in which they expressed their findings. One of the students articulated that the general rule "a hit is one point more (added) than the number of points (subtracted) by a miss" also applies to other triads such as 16 hits-16 misses-16 points and in general to all triads of equal numbers.

To conclude this section about the ICT environment, we must say that observing the students while working with the applet gave us quite a good opportunity to get closer to the students' problem-solving processes.

4.2. Does the ICT environment support the students' problem-solving performance?

In this section, we discuss the results from the pre-test and the post-test in the experimental and control group. Figure 4 shows the average number of correct answers per student in both groups in school A and school B.

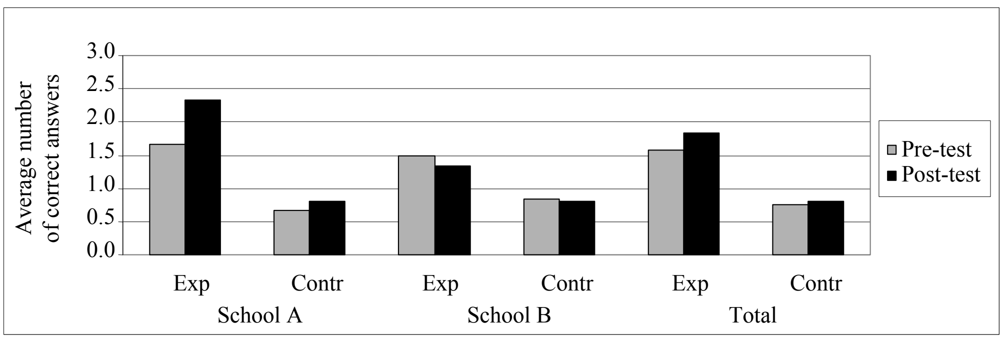

Figure 4: Average number of correct answers per student in the pre and the post-test in both groups

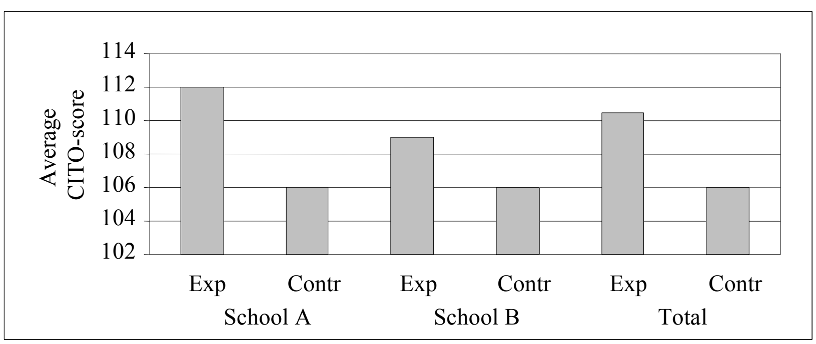

Figure 5: Average CITO score of the experimental and control group

As can been seen in Figure 4, if the group of students is taken as a whole, the experimental group gained slightly from the treatment. However, we have too few data to give a reliable answer to the research question. Only 12 students from school A and 12 students from school B were involved in this study and among these schools, the results were quite different. Only in school A, there is a considerable improvement in the scores of the post-test. Another issue is the mismatch between experimental and control group (see also Section 3.1). In both schools, the control group scored lower than the experimental group. This mismatch was more evident in school A. A plausible explanation for these differences could be that although all students had an A score in mathematics, the average CITO scores of the experimental group and the control group were different in school A and school B (see Figure 5).

In fact, the differences between the average CITO score of the experimental and control group in each school, presented in Figure 5, are similar to the differences between the average scores of these groups in the paper-and-pencil test. In school A, the control group has a lower CITO score than the experimental group. The same holds for school B, but there the difference is smaller than in school A.

5. DISCUSSION

We started this study with two questions that emerged from the earlier POPO study. To investigate these questions, we set up, as a start, a small-scale study in which an ICT environment played a crucial role. The dialogues between the students and their actions when working in the ICT environment gave us a first answer to the first research question. The collected data provided us with a detailed picture of students' problem solving and revealed some interesting processes, for example, the bouncing effect and the making of wrong adaptations.

Our second question is difficult to answer. The sample size, and the number of the test items were not sufficient to get reliable results and the time we had at our disposal was not enough to gather and analyze more data. Moreover, the time that the experimental group worked in the ICT environment was rather limited to expect an effect. Despite these shortcomings, we decided to carry out a small-scale study in order to try out the test items and the ICT environment with a small group of students first.

Clearly, more data (more students, more schools and more problems) are needed to confirm or reject our conjecture that having experience with interrelated variables in a dynamic, interactive ICT environment leads to an improvement in problem solving performance. For this reason, we will extend our study to larger groups of students, involving students of higher grades and different mathematical ability levels as well. Moreover, to see more of an effect we will enlarge the working in the ICT environment substantially. In addition, we will extend the study by analyzing the students' problem-solving strategies when solving paper-and-pencil problems. Our experiences from the present study will serve as a basis for doing this future research.

REFERENCES

Bennett, R. E., & Persky, H. (2002). Problem solving in technology-rich environments. In C. Richardson (Ed.), *Assessing gifted and talented children* (pp. 19-33). London: Qualifications and Curriculum Authority.

Burkhardt, H., & Pead, D. (2003). Computer-based assessment: a platform for better tests? In C. Richardson (Ed.), *Whither Assessment?* (pp. 133-148). London: Qualifications and Curriculum Authority, UK. Retrieved from http://www.nottingham.ac.uk/education/MARS/papers/whither/index.htm.

Clements, D. H., & Meredith, J. S. (1993). Research on Logo: Effects and Efficacy. *Journal of Computing in Childhood Education, 4*(3-4), 263-290.

Clements, D. H. (1998). Computers in Mathematics Education Assessment. In G. W. Bright & J. N. Joyner, *Classroom Assessment in Mathematics: Views from a National Science Foundation Working Conference* (pp. 153-159). Lanham, MD: University Press of America.

Clements, D. H. (2000). From exercises and tasks to problems and projects. *Journal for Research in Mathematics Education, 19*(1), 9-47.

Clements, D. H. (2002). Computers in Early Childhood Mathematics. *Contemporary Issues in Early Childhood, 3*(2), 160-181.

D'Amore, B, & Zan, R. (1996). Italian research on problem solving 1988-1995. In A. Gagatsis & L. Rogers (Eds.), *Didactics and History of Mathematics* (pp. 35-51). Thessaloniki: Erasmus ICP-95-G-2011/11.

Halmos, P. (1980). The heart of mathematics. *American Mathematical Monthly, 87*(7), 519-524.

Harlen, W., & Deakin Crick, R. (2003). A systematic review of the impact on students and teachers of the use of ICT for assessment of creative and critical thinking skills. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London. Retrieved from http://eppi.ioe.ac.uk.

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research.* 77(1), 81-112.

Hoyles, C., & Noss, R. (2003). What can digital technologies take from and bring to research in mathematics education? In A. J. Bishop (Ed.), *Second International Handbook of Mathematics Education* (pp. 323-349). Dordrecht; Boston: Kluwer Academic Publishers.

Kantowski, M. G. (1977). Processes Involved in Mathematical Problem Solving. *Journal for Research in Mathematics Education, 8*(3), 163-180.

Lesh, R., & Zawojewski, J. S. (2007). Problem solving and modeling. In F. K. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (Vol. 2, pp. 763-804). Charlotte, NC: Information Age Publishing.

Lester, F. K., & Kehle, P. E. (2003). From problem solving to modeling: The evolution of thinking about research on complex mathematical activity. In R. Lesh & H. H. Doerr (Eds.), *Beyond Constructivism: Models and Modeling Perspectives on Mathematics Problem Solving, Learning and Teaching* (pp. 501-518). Mahwah, NJ: Erlbaum.

Mercer, N., & Littleton, K. (2007). *Dialogue and the Development of Children's Thinking: a Sociocultural Approach.* London and New York: Routledge.

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

Papert, S. (1980). *Mindstorms: Children, Computers, and Powerful Ideas.* New York: Basic Books.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). *Knowing What Students Know: The Science and Design of Educational Assessment.* Washington, DC: National Academy Press.

Roschelle, J. M., Pea, R. D., Hoadley, C. M., Gordin, D. N., & Means, B. M. (2000). Changing how and what children learn in school with computer-based technologies. *The Future of children 10*(2), 76-101.

Schoenfeld, A. H. (1985). *Mathematical Problem Solving.* New York: Academic Press.

Sfard, A., & Leron, U. (1996). Just Give me a Computer and I will Move the Earth: Programming as a Catalyst of a Cultural Revolution in the Mathematics Classroom. *International Journal of Computers for Mathematical Learning, 1*(2), 189-195.

Suppes, P., & Morningstar, M. (1969). Computer-assisted instruction. *Science 166*(903), 343-350.

Threlfall, J., Pool, P., Homer, M., & Swinnerton, B. (2007). Implicit aspects of paper and pencil mathematics assessment that come to light through the use of the computer. *Educational Studies of Mathematics 66*(3), 335-348

Van Amerom, B. A. (2002). *Reinvention of early algebra: developmental research on the transition from arithmetic to algebra*. Utrecht: CD-ß Press, Center for Science and Mathematics Education, Utrecht University.

Van den Heuvel-Panhuizen, M. (2007). Changes in purpose, content and format: Trends in assessment in mathematics education. In B. Choksi & C. Natarajan (Eds.), *The epiSTEME Reviews, Vol. 2, Research Trends in Science, Technology and Mathematics Education* (pp. 241-266). New Delhi, India: Macmillan.

Van den Heuvel-Panhuizen, M., Bakker, A., Kolovou, A., & Elia I. (in preparation) *Mathematical problem solving by high achieving primary school students.*

Wegerif, R., & Dawes, L. (2004). *Thinking and Learning with ICT. Raising Achievement in Primary Classrooms.* London and New York: Routledge Falmer.