Mathematical problem solving in primary school

ico

This research was carried out in the context of the Dutch Interuniversity Centre for Educational Research.

Kolovou, Angeliki

Mathematical problem solving in primary school / A. Kolovou – Utrecht: Freudenthal Institute for Science and Mathematics Education, Faculty of Science, Utrecht University / FIsme Scientific Library (formerly published as CD-β Scientific Library), no 66, 2011.

Dissertation Utrecht University. With references. Met een samenvatting in het Nederlands.

ISBN: 978-90-73346-72-7
Key words: mathematical problem solving / early algebra / textbooks / ICT / online game / feedback / strategies / gender

Cover design: Vormgeving Faculteit Bètawetenschappen
Printed by: Drukkerij Wilco BV

MATHEMATICAL PROBLEM SOLVING IN PRIMARY SCHOOL

WISKUNDIG PROBLEEM OPLOSSEN IN HET PRIMAIR ONDERWIJS

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof. dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op maandag 4 juli 2011 des middags te 12.45 uur

door
Angeliki Kolovou

geboren op 30 mei 1974
te Athene, Griekenland

Promotoren: Prof. dr. M.H.A.M. van den Heuvel-Panhuizen
Prof. dr. J.A. van Maanen

This thesis was (partly) accomplished with financial support from the Greek Ministry of Education: Paid educational leave Φ.12/6976/1-10-2007.

Table of contents

For Georgia and Ilias

Ik weet helemaal niet meer hoe ik aan het volgend vraagstuk moet beginnen – kun je me helpen ?
Piet vertrekt om 10 uur en rijdt met gemiddelde snelheid van 8km per uur, Jan vertrekt om 10.15 uur en rijdt met gemiddelde snelheid van 12km per uur tot hij Piet tegenkomt.
Wanneer komen ze mekaar tegen, hoever en hoe lang hebben zowel Jan als Piet dan gereden.
Ik loop gewoon vast in mijn eigen gedachten

I do not know how to start with the following problem – can you help me?
Piet departures at 10 o'clock and rides with an average speed of 8 km per hour, Jan departures at 10.15 and rides with an average speed of 12 km per hour until he meets Piet.
When do they meet each other, how far and for how long did they ride.
I am just stuck.

(Question sent by student to the Freudenthal Institute)

Chapter 1
Introduction

Introduction

1 Background of the PhD study

Much of the early research on mathematical problem solving has drawn on the seminal work of Pólya (1945) on heuristics or general problem solving strategies. Pólya pointed out that besides knowledge of the subject-matter, problem solving requires the application of heuristics, which he defined as “mental operations useful in solving problems” (p. 130). Early attempts to teach problem solving focused on heuristic training (Kantowski, 1977; Schoenfeld, 1979), however, without much success (Lester, 1994; Schoenfeld, 1992). Other studies focused on problem solving behavior and investigated problem solving expertise in order to figure out what instruction is necessary for successful problem solving (Schoenfeld & Hermann, 1982; Silver, 1979; Sweller, Mawer, & Ward, 1983). Another line of inquiry placed emphasis on the role of metacognitive skills and on beliefs related to problem solving as well (Lester, Garofallo, & Kroll, 1989; Schoenfeld, 1982; 1983). All these perspectives in researching problem solving were gradually integrated in a framework including domain-specific knowledge, heuristic strategies, metacognitive skills and affective components (De Corte, Greer, & Verschaffel, 1996; Schoenfeld, 1985; 1992).

After the first period of intense interest starting in the mid-1970s, research on problem solving fell into decline in the mid-1990s (Schoenfeld, 2007). More recently, there is a growing body of documents for integrating problem solving into the primary school mathematics curriculum and beyond (Arcavi & Friedlander, 2007; Kilpatrick, Swafford, & Findell, 2001; NCTM, 2000). However, more research is needed to specify what is to be learned and at what age, in order to support teachers fulfilling this goal (Stacey, 2005). At the same time, research has emphasized the role of technology tools for enhancing students’ problem solving performance (Clements, 2000; Hoyles & Noss, 2003) but these claims need to be well evidenced by more empirical studies (Kilpatrick et al., 2001).

The PhD study described in this thesis focuses on problem solving in primary school with the use of new technological tools. The study emerged from the first phase of the POPO project (Probleem Oplossing in het Primair Onderwijs) that was set up in 2004 to investigate problem solving by high-achieving Dutch primary school students. The initial findings of the POPO project concerning students from grade 4 yielded some unexpected results (Van den Heuvel-Panhuizen & Bodin, 2004). Despite their high scores on a general mathematics test, the students encountered difficulties in solving challenging non-routine contextual number problems. These complex problems required more than just carrying out calculations. The students had to develop a good understanding of how the numbers in the problems were related. Another observed shortcoming in the performance was related to the students’ counterproductive attitudes. For instance, students were not inclined to use their scrap paper to write down in-between answers or make a sketch of the problem situation, which could have helped them to solve the problems. Moreover, very often students did not even attempt to solve the problems and when they did, they did not persist in pursuing a solution.

Low results in problem solving were also reported by the Programme for International Student Assessment (PISA). Among 40 OECD-countries, Dutch 15-year-old students were ranked fourth in mathematics, but placed twelfth in problem solving (OECD, 2004). It is also striking that, among all OECD-countries, the difference between scores for mathematics and problem solving was the largest in the Netherlands (PISA-NL-team, 2006).

Therefore, the purpose of this PhD study was to gain more knowledge about the problem solving ability of primary school students in the Netherlands. In particular, the study was designed to enrich our understanding of students' problem solving performance and strategies, their opportunities to learn problem solving, and characterize possibilities to enhance their performance on problem solving.

2 What is meant by problem solving?

One could not agree more that the primary goal of mathematics education should be to make students independent problem solvers. However, the definition of problem solving and its role in school mathematics are much less unequivocal. Although this perspective was noted by various researchers as early as the beginning of the 1990s (Lester, 1994; Schoenfeld, 1992), to date – almost two decades later – problem solving is still an ambiguous concept and researchers are still concerned with similar issues. Probably, its complex and multifaceted character has made it difficult to provide adequate definitions of problem solving, set clear goals for educational practice, and achieve a clear understanding of how to support the development of problem solving.

Very often the terms 'problem' and 'non-routine problem' are used interchangeably in opposition to what is commonly called a 'routine problem'. Routine problems – usually one- or two-step problems – require the reproduction and application of a fixed solution procedure, whereas non-routine problems require productive thinking and can be approached in more or less sophisticated ways.

In several definitions of problem solving, the non-straightforward character of the solution procedure is underlined. For example, Kantowski (1977) stated that a task is a problem for an individual when his immediate knowledge is not adequate to solve it, or when he does not have an algorithm at his disposal that leads straightforwardly to the solution. Similarly, Schoenfeld (1985) argued that a problem is a task the individual wants to complete, and for which he or she does not have access to a straightforward means of solution.

Problem solving is also seen as a high-level skill; it is considered to be the heart of mathematics (Halmos, 1980). As a matter of fact, the distinction between higher and lower types of thinking is a recurrent theme in the psychology of learning. Skemp (1976) made a distinction between relational (i.e., *knowing both what to do and why*) and instrumental understanding (i.e., *rules without reasons*). Similarly, Hiebert and Lefevre (1987) distinguished between conceptual knowledge (which is rich in relationships) and procedural knowledge (which consists of formal language and algorithms). Stein, Schwan, Henningsen, and Silver (2000) differentiated between mathematical tasks with high-level and low-level demands. The former category of tasks requires engagement with conceptual ideas and complex thinking, whereas the latter includes memorization and algorithmic tasks unrelated to the underlying meaning. However, the two types of knowledge do not develop independently; students' conceptual understanding influences the procedures they apply and vice versa (Rittle-Johnson & Alibali, 1999).

Problem solving is considered a complex activity that requires much more than a simple recall of facts and procedures. Boaler (2002) argued that in problem solving various mathematical areas have to be linked, which is an action that extends beyond knowledge. Lester and Kehle (2003) claimed that problem solving is a function of several interdependent factors such as knowledge acquisition and utilization, control, beliefs, affects, and various representational modes.

A broader definition of problem solving is given by Lesh and Zawojewski (2007) who emphasized that problem solving is a goal-directed activity that entails interpreting a situation mathematically, which means modeling. For them problem solving is a complex activity that exceeds the realm of school mathematics and "involves iterative cycles of expressing, testing, and revising mathematical interpretations" (p. 728).

3 Problem solving in the present study

A characteristic of problem solving as investigated in this study is that it refers to non-routine puzzle-like problems. Solving these problems is a cognitive activity that requires an insightful approach to the problem situation and strategic thinking. It entails more than a direct application of an algorithm, formula or procedure. The solution process often requires many steps back and forth until the student is able to unravel the complexity of the problem situation. Furthermore, the students have to be aware of how the numbers or quantities relate to one another in order to find a way to the solution (O'Brien & Moss, 2007). However, we need to be aware of the subjectivity that lies in this definition of problem solving, as Schoenfeld (1985, p. 74) pointed out: "It is a particular relationship between the individual and the task that makes a task a problem for that person".

More specifically, the non-routine problems in this study contain interrelated variables and can be solved by means of an algebraic procedure. However, when primary school students encounter such problems they can only apply informal, context-connected solutions and employ various problem solving strategies, such as systematically trying out possible solutions or trial-and-error. Yet, students' informal strategies and notation when solving such problems can provide important entry points for developing algebraic reasoning. For the sake of conciseness we call these problems *early algebra problems*. In our study the terms "non-routine puzzle-like problems" and "early algebra problems" are used interchangeably.

The purpose of this study is, thus, to investigate Dutch primary school students' competence in solving problems with interrelated variables. In addition, we aim to examine students' opportunities to learn problem solving within the current curriculum and to provide them with support in solving problems.

4 Structure of the thesis

The PhD thesis comprises a series of articles each addressing a different perspective of this study on problem solving by Dutch primary school students. Table 1 illustrates the structure of the thesis and shows which research question is answered in each chapter.

Table 1
Structure of thesis

<table>
<tr><th>Chapter</th><th></th></tr>
<tr><td>1</td><td>Introduction</td></tr>
<tr><td>2</td><td>Research Question 1
How able are Dutch primary school students in solving non-routine puzzle-like problems?</td></tr>
<tr><td>3</td><td>Research Question 2
To what degree do textbooks offer students opportunities to learn solving non-routine puzzle-like problems?</td></tr>
<tr><td>4</td><td rowspan="4">Research Question 3
How can we support students in solving non-routine puzzle-like problems?</td></tr>
<tr><td>5</td></tr>
<tr><td>6</td></tr>
<tr><td>7</td></tr>
<tr><td>8</td><td>Conclusion</td></tr>
</table>

Chapter 2 focuses on the analysis of data stemming from the first phase of the POPO project. This chapter mainly addresses the first research question:

How able are the Dutch primary school students in non-routine problem solving?

a. How do students perform in non-routine problem solving?

b. Which strategies do students apply in non-routine problem solving?

The data collected in the first phase of the POPO project gave us an initial glimpse in the difficulties that Dutch students encounter in problem solving. In order to investigate strategy use in non-routine problem solving, the responses to three non-routine problems from the data set collected in the first phase of the POPO project were further analyzed. The responses were examined with respect to the strategies that students applied and their flexibility in strategy use.

Chapter 3 reports on the textbook analysis and addresses the second research question:

To what degree do textbooks offer students opportunities to learn non-routine problem solving?

The input from mathematics textbooks is an important factor in realizing students' opportunities to learn problem solving. The analysis of the content of textbooks can, therefore, provide insights into students' low performance in problem solving. To this end, a mathematical problem solving framework was developed to determine the amount and the cognitive demand of problem solving tasks in the textbook series for grade 4.

After finding that non-routine problems raised serious obstacles even for high-achieving students, we wondered what kind of learning environment could support students in solving early algebra problems. Chapters 4, 5, 6 and 7 deal with the third research question:

How can we support students in non-routine problem solving?

Chapter 4 reports on a dynamic computer game, called *Hit the target* that was employed to support learning by offering students opportunities to experiment, produce and reflect on their solutions. The chapter addresses the following research questions:

Which problem solving strategies do fourth-grade students deploy in the 'Hit the target' environment?

Does this ICT environment support students' problem solving performance?

Chapter 5 describes a study in which we further examined issues that came up in the aforementioned study by including more students and offering them an online environment. The resulting large-scale study focused on the role of feedback and more particularly, the computer-generated feedback and how it is related to students' solving processes. The chapter addresses the following research questions:

How can feedback resulting from playing an online game at home support students' solving processes in early algebra problems?

How do solving processes differ from situations where no feedback is available?

Chapter 6 examines the effects of working in the online environment on students' performance in problem solving. In particular, it compares the performance of the students that participated in different experimental conditions. Moreover, it investigates the role of prior mathematical performance, gender, and effort put into working online on students' posttest problem solving performance. In particular, the following research questions are addressed:

Does an intervention including an online game enhance students' performance in early algebra?

What is the effect of the effort indicators of the online work on the achieved performance in early algebra problem solving?

Do boys and girls differ in their achieved performance in early algebra as a result of the intervention and is gender related to the effort indicators?

Chapter 7 zooms in on the strategies that students in grades 4, 5, and 6 applied while working online and the relationships between strategy use and performance in early algebra problems. In particular, we seek to answer the following research questions:

How do fourth to sixth graders utilize an interactive online environment including a dynamic game to solve early algebra problems?

Is students' work in this environment related to their performance in a written test on early algebra problem solving?

Chapter 8 synthesizes the gains from the series of studies on problem solving carried out in this PhD study and discusses what lessons can be drawn from the findings of these studies for making Dutch arithmetic education more mathematical, in order to provide students with more opportunities for developing mathematical reasoning and to prepare them for learning algebra in secondary school.

References

Arcavi, A., & Friedlander, A. (2007). Curriculum developers and problem solving: the case of Israeli elementary school projects. *ZDM The International Journal on Mathematics Education, 39*, 355–364.

Boaler, J. (2002). Exploring the nature of mathematical activity: Using theory, research and 'working hypotheses' to broaden conceptions of mathematics knowing. *Educational Studies in Mathematics, 51*, 3−21.

Clements, D.H. (2000). From exercises and tasks to problems and projects: Unique contributions of computers to innovative mathematics education. *Journal of Mathematical Behavior, 19*, 9−47.

De Corte, E., Greer B., & Verschaffel, L. (1996). Mathematics teaching and learning. In D. Berliner & R. Calfee (Eds.), *Handbook of educational psychology* (pp. 491−549). New York: Macmillan.

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

Hiebert, J., & Lefevre, P. (1987). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and Procedural Knowledge: The Case of Mathematics* (pp. 1−27). Hillsdale, NJ: Lawrence Erlbaum Associates.

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*, 163−180.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.) (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.

Lesh, R., & Zawojewski, J.S. (2007). Problem solving and modeling. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 2, pp. 763–804). Charlotte, NC: Information Age Publishing.

Lester, F.K. (1994). Musings about mathematical problem-solving research: The first 25 years in JRME. *Journal for Research in Mathematics Education, 25*, 660–675.

Lester, F.K. Garofallo, J., & Kroll, D.L. (1989). Self-confidence, interest, beliefs, and metacognition: Key influences on problem-solving behavior. In D. B. McLeod & V. M. Adams (Eds.), *Affect and mathematical problem solving: A new perspective* (pp. 75−88). New York: Springer-Verlag.

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. Doerr, H. (Eds.), *Beyond Constructivism: Models and Modeling Perspectives on Mathematics Problem Solving, Learning and Teaching* (pp. 501−518). Mahwah, NJ: Erlbaum.

National Council of Teachers of Mathematics (NCTM) (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.

O'Brien, T.C., & Moss, A. (2007). What's basic in mathematics? *Mathematics Teaching*, *202*, May, 26−27.

OECD, (2004). *Problem Solving for Tomorrow's World: First Measures of Cross-Curricular Competences from PISA 2003.* Paris: Author. Retrieved February 5, 2011, from http://www.oecd.org/dataoecd/25/12/34009000.pdf.

PISA-NL-team (2006). *Wiskundige geletterdheid volgens PISA. 1. Analyse - hoe staat de vlag erbij?* [Mathematical literacy according to PISA. 1. Analysis - what is the situation?] Utrecht, the Netherlands: Freudenthal Institute.

Pólya, G. (1945). *How to solve it*. Princeton: Princeton University Press.

Rittle-Johnson, B., & Alibali, M.W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology, 91,* 175−189.

Schoenfeld, A.H. (1979). Explicit heuristic training as a variable in problem-solving performance. *Journal for Research in Mathematics Education, 10*, 173−187.

Schoenfeld, A.H. (1982). Some thoughts on problem-solving research and mathematics education. In F. K. Lester. & J. Garofalo (Eds.), *Mathematical problem solving: Issues in research* (pp. 27−37). Philadelphia, PA: The Franklin Institute Press.

Schoenfeld, A.H. (1983). Beyond the purely cognitive: Belief systems, social cognitions, and metacognitions as driving forces in intellectual performance. *Cognitive science, 7*, 329−363.

Schoenfeld, A.H. (1985). *Mathematical problem solving*. New York: Academic Press.

Schoenfeld, A.H. (1992). Learning to think mathematically: Problem solving, metacognition, sense-making in mathematics. In D. Grouws (Ed.), *Handbook for Research on Mathematics Teaching and Learning* (pp. 334–370). New York: MacMillan.

Schoenfeld, A.H. (2007). Problem solving in the United States, 1970–2008: Research and theory, practice and politics. *ZDM The International Journal on Mathematics Education*, *39*, 537–551.

Schoenfeld, A.H. &. Hermann, D.J (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*, 484–494.

Silver, E.A. (1979). Student perceptions of relatedness among mathematical verbal problems. *Journal for Research in Mathematics Education, 10*, 195–210.

Skemp R.R. (1976). Relational understanding and instrumental understanding, *Mathematics Teaching, 77*, 20−26.

Stacey, K. (2005). The place of problem solving in contemporary mathematics curriculum documents. *Journal of Mathematical Behavior, 24*, 341−350.

Stein, M.K., Schwan, S., Henningsen, A., & Silver, E.A. (2000). *Implementing Standards-Based Mathematics Instruction*. New York: Teachers College Press.

Sweller, J., Mawer, R.F., & Ward, M.R. (1983). Development of expertise in mathematical problem solving. *Journal of Experimental Psychology: General*, *112*, 639–661.

Van den Heuvel-Panhuizen, M., & Bodin-Baarends, C. (2004). All or nothing: Problem solving by high achievers in mathematics. *Journal of the Korea Society of Mathematical Education, 8*(3), 115−121.

Chapter 2

Exploring strategy use and strategy flexibility in non-routine problem solving by primary school high achievers in mathematics

Iliada Elia [a, b], Marja van den Heuvel-Panhuizen [a, c], and Angeliki Kolovou [b]

[a] *Centre of Educational Research and Evaluation, Cyprus Pedagogical Institute, Cyprus*
[b] *Freudenthal Institute for Science and Mathematics Education, Utrecht University, the Netherlands*
[c] *IQB, Humboldt University Berlin, Germany*

Elia, I., Van den Heuvel-Panhuizen, M., & Kolovou, A. (2009). Exploring strategy use and strategy flexibility in non-routine problem solving by primary school high achievers in mathematics. *ZDM The International Journal on Mathematics Education, 41*, 605–618.

Exploring strategy use and strategy flexibility in non-routine problem solving by primary school high achievers in mathematics

1 Introduction

Problem solving is considered the most significant cognitive activity in everyday and professional environments (Jonassen, 2000). An attribute which is considered integral to the problem solving process is strategic behavior (Polya, 1957; Schoenfeld, 1992). Another important characteristic of problem solving is that people are able to work in a flexible way and can modify their behavior according to changing situations and conditions. In fact, a person's flexibility determines to a large degree how well he or she can cope with a new situation. As Demetriou (2004) emphasized, more flexible thinkers can develop more refined concepts that are better adjusted to the special features of the environment and produce more creative and appropriate solutions to problems.

What is true for solving problems in general also applies to mathematical problem solving. Numerous studies in mathematics education (e.g., Pape & Wang 2003; Verschaffel et al., 1999) hold strategy use central to processing mathematical problems. A well-documented finding is that success in solving a mathematical problem is positively related to the students' use of problem solving strategies (Cai, 2003; Kantowski, 1977). In mathematics education, though, students continuously face new situations and new problems (Stanic & Kilpatrick, 1988), which require them not only to know and apply various strategies, but also to be flexible (e.g., Baroody, 2003; Silver, 1997). What they have learned in one situation and what applies to one problem, will not necessarily fit another situation or be appropriate for another problem. As a result, in the mathematics education community considerable research has been devoted to studies on strategy flexibility. However, most of these studies focused on children's strategies related to arithmetic concepts and skills (e.g., Baroody 2003; Beishuizen, Van Putten, & Van Mulken 1997). Less attention has been devoted to the study of flexibility in using heuristic strategies in mathematical non-routine problem solving (e.g., Kaizer & Shore, 1995), especially among primary school children. More information is needed to understand how flexibility in using heuristic strategies occurs in non-routine problem solving and how it is associated with performance.

This paper aims to give insight into the strategy use and strategy flexibility of high achievers in primary school mathematics in non-routine problem solving. The theoretical value of our study lies in that it may contribute to the formation of an operational definition of strategy flexibility in non-routine problems, as it proposes and explores two distinct aspects of strategy flexibility in students' problem solving behavior: strategy flexibility *between* different problems and *within* a problem. Furthermore, the study may clarify the interrelations of strategy use and strategy flexibility within problems and across problems with problem solving success. From a practical point of view, knowledge about the above may contribute to the interpretations of individual differences in problem solving and provide suggestions on how to support student development in solving non-routine problems.

In our study, we interpret and use the terms ‘non-routine problem’ and ‘problem’ interchangeably, on the basis of Schoenfeld’s (1983) definition of a problem, that is, as an unfamiliar situation for which an individual does not know how to carry out its solution. In other words, he or she is unable to solve the situation comfortably using routine or familiar procedures (Carlson & Bloom, 2005). Furthermore, the term ‘strategies’ refers to problem solving strategies or heuristics, in the sense given by Schoenfeld (1992) and Verschaffel et al. (1999), such as drawing a picture, making a list or a table, or guessing and checking.

2 Theoretical framework

2.1 Flexibility

The term ‘flexibility’ has been extensively used by researchers in the field of cognitive and developmental psychology (Demetriou, 2004; Krems, 1995) on the one hand, and of mathematics education (Krutetskii, 1976; Verschaffel, Luwel, Torbeyns, & Van Dooren, in press), on the other hand.

According to Demetriou (2004) flexibility refers to the quantity of variations that can be introduced by a person in the concepts and mental operations he or she already possesses. Krems (1995) defines cognitive flexibility “as a person’s ability to adjust his or her problem solving as task demands are modified” (p. 202). Mathematics educators highlight the educational value of recognizing and promoting flexibility in children’s self-constructed strategies and have developed and implemented instructional materials and interventions planned for the improvement of such flexibility (e.g., Freudenthal, 1991; Torbeyns, Desmedt, Ghesquière, & Verschaffel, in press).

A term that is closely related to flexibility is ‘adaptivity’. An in-depth analysis by Verschaffel et al. (in press) of how these terms are currently used in the literature, suggests that the term ‘flexibility’ is primarily used to refer to using multiple strategies and switching between them, whereas the term ‘adaptivity’ emphasizes the ability to consciously or unconsciously select and use the most appropriate approach for solving a certain mathematical item or problem, by a particular person, in a given socio-cultural context. In the present study, which we see as the start of a program to investigate flexibility and adaptivity in non-routine problem solving, we focus on flexibility in strategy use.

In the light of the above, we consider strategy flexibility as the behavior of switching strategies during the solution of a problem, i.e. intra-task strategy flexibility, or between problems, i.e. inter-task strategy flexibility. This is a broad operational definition of flexibility, which includes different patterns of changing strategies. It is noteworthy that, in this definition, we do not connect strategy flexibility with the appropriateness of the problem solving strategies. Connecting strategy appropriateness with strategy adaptivity in problem solving would be an interesting topic to investigate. However, this is not our focus in the present study.

2.2 Mathematical problem solving

2.2.1 What counts as a problem?

A (non-routine) problem appears when an individual encounters a given situation, intends to reach a required situation, but does not know a direct way of accessing or fulfilling his or her goal. A central issue is the problem solver's ignorance with respect to a solution method (Mayer, 1985). In contrast with routine problems which involve the application of routine calculations, non-routine problems do not have a straightforward solution, but require creative thinking and the application of a certain heuristic strategy to understand the problem situation and find a way to solve the problem (Pantziara, Gagatsis & Elia, in press). Therefore, non-routine problems are considered more complicated and difficult than routine problems. However, Polya observed that, although routine problems can be used to fulfill particular didactical functions of teaching students to apply a certain procedure or a definition correctly, only through the careful use of non-routine problems can students develop their problem solving ability (Stanic & Kilpatrick, 1988).

In our study, we concentrate on non-routine problems which involve interrelated variables and require the understanding that a change in one variable affects the other variables. Since the participants in this study – fourth graders – do not have any algebraic tools at their disposal, they cannot apply a routine algebraic method, but have to confront these tasks by a heuristic or problem solving strategy, such as trial-and-error or systematic listing of possible solutions. A more detailed description of the problems is given in the *Methods* section.

2.2.2 Strategy use in problem solving

Problem solving strategies constitute a fundamental aspect of mathematical thinking (Schoenfeld, 1992). Previous research has shown that students' use of heuristic strategies was positively related to performance on problem solving tests, but the effects were only marginal (Kantowski, 1977). The results of more recent studies have provided stronger evidence for the use of heuristic strategies as a means to enhance problem solving. Specifically, the problem solvers' ability to try possible solution approaches and to assess the likely outcome of each one has been found to play an important role to their efficient decision making and problem solving success (Carlson & Bloom, 2005). Altun and Sezgin-Memnun (2008) have found that among mathematics teacher trainees, the strategies used had a dominant and decisive role in determining success on a problem.

Several mathematical problem solving strategies can be introduced in primary or middle school mathematics teaching, such as: guess-check-revise, draw a picture, act out the problem, use objects, choose an operation, solve a simpler problem, make a table, look for a pattern, make an organized list, write an equation, use logical reasoning, work backward (Charles, Lester, & O'Daffer, 1992).

Although many students do not spontaneously use valuable heuristic strategies when dealing with unfamiliar complex problems (De Bock, Verschaffel, & Janssens, 1998; Schoenfeld, 1992), a number of researchers have found evidence that primary school students are capable of using strategies (like the ones noted above) to solve problems (e.g., Cai 2003; Follmer, 2000). Moreover, Verschaffel et al. (1999) found that problem solving instruction addressed to fifth- grade students contributed to the improvement of their ability to solve mathematical application problems and to use problem solving strategies. Follmer's (2000) findings showed that in the fourth grade, the instruction on non-routine problems had a positive impact on students' use of cognitive strategies and their awareness of how they solved the problems.

Besides heuristic (cognitive) strategies, solving a problem also requires metacognitive strategies. Metacognitive strategies involve self-regulatory actions, such as decomposing the problem, monitoring the solution process, evaluating and verifying results (Schoenfeld, 1992; Verschaffel et al., 1999). These strategies play a crucial role in achieving problem solving success (Schoenfeld, 1992; Carlson & Bloom, 2005). However, a number of studies have shown that students display deficiencies in applying these strategies in their solution efforts (Schoenfeld, 1985; 1992).

2.3 Strategy flexibility and problem solving

Strategy flexibility appears to be strongly interconnected with problem solving activities and performance. On the one hand, from a developmental perspective, developing and excelling in problem solving is to a considerable extent a function of increase of flexibility. On the other hand, from a differential point of view, individual differences at the same age level result from variations in flexibility, which enables the individual to carry out strategy alternations on the basis of the requirements of particular problems (Demetriou, 2004). Given that problem solving performance is improved when task requirements and problem solving methods are coordinated (Krems, 1995), flexibility in strategy use may significantly contribute to success.

Martinsen and Kaufmann (1991) distinguished between solvers who extended or persevered on the use of a particular problem solving strategy ('assimilators') and solvers who varied their problem solving strategy more frequently ('explorers'), even when a shift was not essential. Flexibility, however, especially characterizes competent students (Shore & Kanevsky, 1993). Kaizer and Shore (1995) compared the flexibility of solution strategies between mathematically competent and less competent 11th-grade students on mathematical non-routine word problems which were drawn from Kruteskii's (1976) work. A distinction was made between verbal-logical methods, visual strategies and trial-and-error procedures. Across the problems, students employed different strategies. Competent students switched primarily between verbal-logical and visual methods, whereas less competent students alternated equally between verbal-logical and visual strategies, or between verbal-logical methods and trial-and-error.

Flexibility can be activated not only across problems, but also within a problem. Krems (1995) suggests that a type of mechanism that is important for flexible strategy use in solving a problem is modifying strategies. A flexible problem solver can modify strategies to correspond with alterations in resources and task requirements and can use several different techniques to find an answer. With development, thinkers become more competent in observing each of the components of a problem separately and reconstruct their structure (or relations) according to the plans or objectives of the particular situation (Demetriou, 2004). As the person's knowledge of the problem's components (mental representation of the problem) becomes more complete and interconnected, he or she can more easily invent and use strategies, determine the most efficient solution path for the intended goal or model and flexibly determine alternative solutions (Baroody, 2003; Demetriou, 2004). Kaizer and Shore (1995) suggest that alternative solution strategies in problem solving may occur when the students experience difficulties with a problem or at any stage of the problem solving process. Nevertheless, Muir and Beswick's (2005) study has shown that most sixth-grade students were not capable of reflecting on the appropriateness of the strategy they had chosen, or display any inclination to use an alternative strategy, even when the initial strategy was not working. Furthermore, students were unwilling to attempt to confirm the appropriateness of the answer using an alternative method.

3 Research questions and predictions

The present study aims to contribute to our understanding of strategy use and strategy flexibility in non-routine problem solving. Our research questions, which refer to primary school high achievers in mathematics, are distinguished into two thematic groups. The first group of questions is concerned with strategy use, while the second group refers to strategy flexibility. For each question, a prediction has been formulated on the basis of the theoretical background presented above and the setting of the study.

3.1 Strategy use

Question 1: To what extent do the students apply strategies to solve non-routine problems?
Prediction 1: We expect that only a small number of students will use strategies to solve non-routine problems. Previous research has shown that many students do not spontaneously use heuristic strategies to tackle unfamiliar complex problems (De Bock et al., 1998; Schoenfeld,

1992). This hypothesis is also based on the setting of this study, which cannot be ignored, since the current Dutch textbooks in mathematics include only a small set of non-routine problems (Kolovou, Van den Heuvel-Panhuizen, & Bakker, in press).

Question 2: What is the relation between strategy use and answer success?
Prediction 2: De Corte (2007) suggests that, while heuristic strategies do not guarantee a correct solution, they significantly strengthen the potential of providing one. Some possible strategies that could be used in solving the problems of this study are: Trial-and-error, systematic listing of possible solutions, calculating extreme values. These strategies vary in their cognitive demands, thus the use of each one of them may have a different effect on problem solving success. The trial-and-error strategy, for example, is a strategy without high cognitive demands, which is commonly used in mathematics classrooms and in everyday life. Stacey (1991) characterizes trial-and-error as an intuitive strategy that everyone can use. In applying the trial-and-error strategy, one has to try possible solutions and compare the results with the intended results. If a match does not occur, then one has to try other solutions by adopting the processes used according to the requirements of the task. Students may be more skilful and experienced in this strategy than in the more sophisticated and less familiar ones. Thus, we expect that the trial-and-error strategy is more likely to lead to success than the other strategies.

3.2 Strategy flexibility

Question 3: To what degree does inter-task strategy flexibility occur among the students?
Prediction 3: We anticipate that only a few of the students will demonstrate strong inter-task strategy flexibility, that is, modify strategies in every problem situation. To be able to show inter-task strategy flexibility, one should have a rich repertoire of strategies. Thus, we ascribe prediction 3 to the Dutch mathematics curriculum, which does not substantially contribute to the development of students' repertoire of heuristics (Kolovou et al., in press) that would enable them to flexibly use a variety of strategies.

Question 4: To what degree does intra-task strategy flexibility occur among the students?
Prediction 4: As a person's mental representation of the problem becomes more complete and interconnected, he or she can more easily invent and use strategies, determine the most efficient solution path for the anticipated goal and flexibly establish alternative solutions (Baroody 2003; Demetriou 2004). Understanding or building a mental representation of a (non-routine) problem depends on the existing cognitive schemes of the solvers (Mayer 1985). Thus, taking into account Dutch students' marginal learning experiences with non-routine problems at school (Kolovou et al. in press), we expect that the majority of the participants who will make a solution attempt for a problem, will use mainly one strategy and fail to consider alternatives even if they encounter difficulties in the problem solving process. In other words, we anticipate that students will rarely demonstrate intra-task strategy flexibility when solving the problems.

Question 5: Are there any differences between students who show inter-task flexibility and students who use a single strategy across the problems in their problem solving performance? According to Demetriou (2004) individual differences in performance at the same age level are closely related to variations in flexibility, which enables the individual to alternate between strategies on the basis of the requirements of particular problems. On the basis of the above, we formulated the following prediction:
Prediction 5: We expect that students who flexibly switch strategies across problems will outperform students who persevere with the same strategy.

Question 6: Are there any differences between students who show intra-task flexibility and students who use a single strategy within problems in their problem solving performance?
Prediction 6: We anticipate that students who use different strategies when solving a problem will exhibit greater problem solving performance than students who use a single strategy within problems. The above prediction is based on a previous research finding that the solvers' ability to try different solution strategies for a problem contributes to their efficient decision making and problem solving success (Carlson & Bloom 2005).

4 Methods

4.1 Participants

A total of 152 high-achieving students (97 boys and 55 girls) in grade 4 (9-10 years of age) from 22 different schools in the Netherlands were examined. The students belonged to the top 25% ability range in mathematics, and were selected by their teachers on the basis of their mathematics score. In most cases this was the students' mathematics score on the CITO Student Monitoring Test. In some cases the so-called DLE score was used. In the data preparation, the CITO scores were converted into DLE scores[1].

4.2 Tasks

Three non-routine problems were given to the students[2]. Instructions involved an explicit request for showing the solution strategy. The three tasks were:

> *1. Angela is 15 years now and Johan is 3 years. In how many years will Angela be twice as old as Johan? (Age problem)*
> *2. Liam has tokens of value 5 and 10 only. In total he has 18 tokens. The total value of these tokens is 150. How many tokens of value 5 does Liam have? (Coins problem)*
> *3. In a quiz you get two points for each correct answer. If a question is not answered or the answer is wrong, one point is subtracted from your score. The quiz contains 10 questions. Tina received 8 points in total. How many questions did Tina answer correctly? (Quiz problem)*

A major characteristic of these problems is that they do not have a straightforward solution, but require a good understanding and modeling of the situation, that is, recognizing how different variables covariate. It is evident that a person who knows elementary algebra might use this knowledge to find the answer to the problems. The third problem, for example, could be tackled by solving the equation $2x - 1(10 - x) = 8$. But, as fourth graders have not yet learned such techniques, they have to coordinate several pieces of information and use other strategies, such as systematic listing of possible solutions or trial-and-error, to solve the problems. These features make the tasks non-routine problems for the students. On the other hand, despite their complexity, the problems are accessible, as they involve small numbers and do not entail difficult calculations.

4.3 Procedure

As only about a quarter of each fourth-grade class participated in the study, the tasks were administered to these students either by their teacher or by another member of the educational staff of the school. These tasks were administered in the middle of the school year, during regular school hours in a quiet place at school. Students were given enough time to finish all three tasks. The teachers were instructed that students must work on their own and no assistance could be given to them. Students were not allowed to use a calculator and were instructed that if they needed to do a calculation, they could use the test sheet.

4.4 Strategy analysis and scoring

As already noted, students were asked to write down their solution process for the problems. For the analysis of the students' responses, a coding scheme was formulated[3] for each problem. After the student work on the three problems was coded, we asked a judge to do a second coding on the responses of twenty randomly chosen students for one problem, which involved 400 dichotomous codes[4]. The interrater reliability was measured with Cohen's Kappa (.83) from which we concluded that the coding scheme was reliable.

In the present study we concentrate on the codification and analysis of the strategies which were visible on the students' test sheets. The strategies that we identified in students' test sheets for the three problems, their explanation and variables' names are presented in Table 1. Examples of these strategies are illustrated in Section 5.2.4.

As for the correctness of the answers, each correct answer on a problem was scored as 1, and each wrong or no answer as 0. An answer was assessed as correct when the accurate numerical result was written on the test sheet.

Table 1
Coding scheme for the strategies across the three problems

Category	Repeating information	Trial-and-error	Systematic listing	Calculating-an-extreme	Proof-or-check	Halving the number of coins or value[a]
Variable name	Rep	TE	Sys	Extr	PoC	Half
Explanation: The response involves…	Repeated information from the problem formulation.	Two or more trials and the last one is the given answer. The steps are not of the same size each time and the "movements" of the trials do not need to go in one direction.	A systematic listing strategy which entails at least three elements (including the final answer). The step size is stable (mostly 1) and the "movement" of the list goes in one direction.	Use of extreme values, e.g. in the coins problem the students start with calculating the value of 18 5-cent coins, or 18 10-cent coins, or calculate how many 5-cent coins or 10-cent coins are in €1,50.	A kind of proof of the result or a check of the calculation. Students explain that the given answer is correct (according to them).	The assumption that there were 9 10-cent coins and 9 5-cent coins, i.e. halving the total number of coins or the total amount of money.

[a] This strategy is applied only in the coins problem.

5 Results

The results are organized into two subsections, which correspond to the two thematic groups within the research questions. That is, we will first present our findings about strategy use and then we will turn to strategy flexibility within and across problems.

5.1 Strategy use

5.1.1 Students' strategies

The non-routine problems turned out to be difficult for the majority of students despite their high general mathematical ability as measured by the CITO Student Monitoring Test and the DLE Test. Only 35 students out of 152 (23%) provided a correct solution to the age problem, 61 students (40%) solved the coins problem correctly and 30 students (20%) succeeded in the quiz problem. The relatively lower success rates at the age and the quiz problem are probably due to their higher complexity in comparison to the coins problem. The two problems required the understanding and coordination of some additional data components and relations. The age problem entailed understanding the same change in time for both children's ages, while the quiz problem involved the coordination of two variables which changed in two different directions, one increasing and another decreasing.

Table 2 shows the percentages of the students who used each of the five strategies per problem and the strategy of halving the total number of coins or the total amount of money in the coins problem. Although repeating information helps students confirm that they use the data given, only a small number of students repeated the problems' information components. The quiz problem appears to elicit this strategy more often (10%), probably because of its high complexity and longer statement. The calculating-an-extreme strategy was hardly used (1-7%). More commonly used strategies were – what we have called – proof-or-check (15-32%), trial-and-error (P1: 8%, P2: 16%, P3: 7%) and systematic listing (P1: 15%, P2: 3%, P3: 11%). As for the 'halving strategy' in the coins problem, it was used only by four students (3%). In general, traces of strategies were only found on about half of the test sheets. These findings provide evidence to prediction 1 suggesting that only a small number of the students could spontaneously apply strategies to tackle the problems.

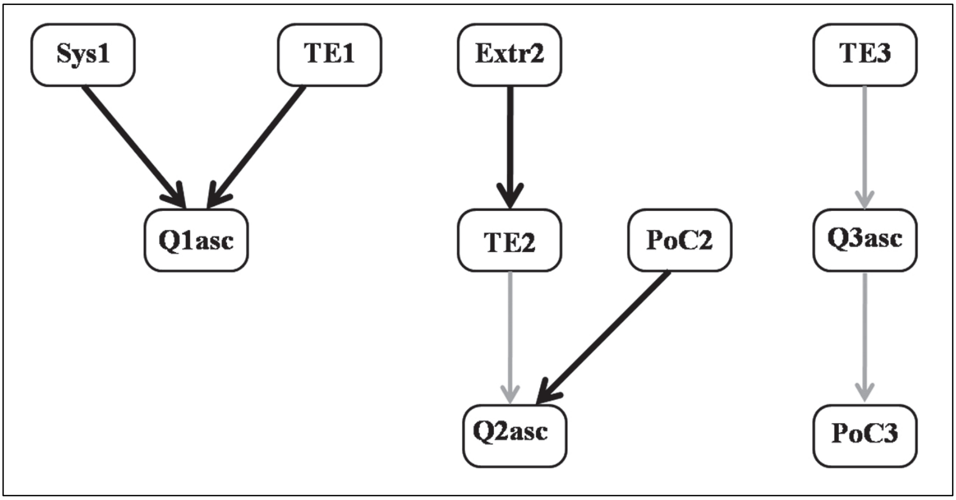
Table 2
Students' strategy distribution for each problem

Strategy	Age problem		Coins problem		Quiz problem	
	f[a] (n = 152)	%	f (n = 152)	%	f (n = 152)	%
Repeat	7	5	4	3	15	10
Trial-and-error	12	8	24	16	10	7
Extreme	2	1	7	5	10	7
Systematic listing	22	15	5	3	16	11
Proof-or-check	22	15	36	24	49	32
Halving coins /value	–	–	4	3	–	–

[a] Students could use more than one strategy (see sections 5.2.2 and 5.2.4)

5.1.2 Successfulness of students' strategies

To determine the strategies that can be considered successful in the solution of the three problems, we performed the implicative statistical method using a computer software called C.H.I.C. (Classification Hiérarchique, Implicative et Cohésitive) (Bodin, Coutourier and Gras, 2000). This method of analysis determines the implicative relations between variables (Elia, Panaoura, Gagatsis, Gravvani, & Spyrou, 2008; Gras, Suzuki, Guillet, & Spagnolo, 2008), which give a statistical meaning to expressions such as: "If we observe the variable *a* in a subject, then in general we observe the variable *b* in the same subject." The underlying principle of the implicative analysis is based on the quasi-implication: "If *a* is true then *b* is more or less true." The implicative diagram represents graphically the network of the quasi-implicative relations among the variables of the study. Figure 1 shows the implicative relations among the variables of students' strategies (see Table 1) and their success on the three problems.



Notes

a) *Q1asc*, *Q2asc* and *Q3asc* refer to students' success on the three problems respectively.

b) The numbers 1, 2 and 3 next to the variable names correspond to the three problems respectively.

c) The thickness of the arrows is a function of the strength of the implications, whose estimated probabilities are 99%, 95% and 90% respectively.

Fig. 1 Implicative diagram of strategies and correct answers to the problems

Three groups of implications were found, each of which links the variables of a different problem. First, systematic listing and trial-and-error strategy appear to be crucial for solving the age problem. Second, success on the coins problem is found to be a function of using primarily the strategy of giving a proof of the result or checking the calculation, and secondly the trial-and-error process. Students who used the strategy of calculating an extreme in the coins problem employed also the trial-and-error process in their solution. Third, using trial-and-error is the major strategy that leads to success in the quiz problem. An important process feature of successful students in the quiz problem is their effort to prove or check their answer. A noteworthy commonality among the three groups of implications is the link between the trial-and-error process and the problem's solution, indicating that using the trial-and-error strategy implied success in all three problems.

The implicative analysis provided evidence for prediction 2 suggesting that the trial-and-error strategy would have a strong potential to lead to success. Specifically, it is the only strategy that led the students to a correct solution for all three problems. Thus, students used the trial-and-error strategy more widely and competently in the problems, independent of the situation or the numbers involved. In the age problem, systematic listing was equally successful. In the coins problem explaining and checking the correctness of a given solution was found to be more successful than trial-and-error. A hypothetical explanation might be that students mentally used a strategy, such as trial-and-error or systematic listing, and reached an answer, which they just checked in the end, without writing down their strategy.

5.2 Strategy flexibility

5.2.1 Inter-task strategy flexibility

Having three problems in total means that there are three different pairs of problems (1-2, 1-3, and 2-3) in which the students can show a difference in strategies. A difference or a change in strategies between each set of two problems was considered to occur if the strategy or the strategies used were not the same between the problems. It should be noted here that this method of analyzing inter-task flexibility does not take into account the effects the order of the problems may have on students' strategies. This is a methodological limitation of our study that should not be disregarded when discussing inter-task strategy flexibility.

A total of 51 students (34%) demonstrated *inter-task strategy flexibility*. Most of these students (n = 27, 18%) changed strategies in only one pair of problems, while they did not exhibit any traces of strategy use in the other problem (score: 1). An equal proportion of students alternated between strategies in two problem pairs (score: 2) (n = 12, 8%) and three problem pairs (score: 3) (n = 12, 8%) respectively. The frequencies and percentages in strategy alternations are shown in Table 3. These findings lend support to prediction 3 which suggests that only a small number of students would display high levels of inter-task strategy flexibility, that is, change strategies in every problem situation. A larger number of students, though, had a low score in inter-task strategy flexibility. It is also noteworthy that only 11 students (7%) were found to persevere with the use of a single strategy across the problems.

Table 3
Distribution of students' inter-task strategy flexibility scores

Score of changes in strategy use over the three problem pairs	f (n=152)	%
Non-applicable	90[a]	59
0	11	7
1	27	18
2	12	8
3	12	8

[a] These students did not exhibit strategy use in more than one problem.

5.2.2 Intra-task strategy flexibility

Regarding the *intra-task strategy flexibility*, only 36 students (24%) alternated strategies in one or more problems. Specifically, 28 students (18%) displayed intra-task strategy flexibility in one problem, seven students (5%) in two problems and one student (1%) in all three problems. Table 4 shows the frequencies and percentages of students who used a particular number (zero to three) of strategies per problem. A similar pattern appears in all three problems. Only a small number of students, i.e. 6, 18 and 21, switched between two or three strategies, in each of the three problems respectively. A single strategy was employed by a considerably larger number of students in each problem, while the vast majority of the students did not apply or did not report on strategies, as shown in their test sheets. These findings verify prediction 4 stating that intra-task flexibility would be rarely detected in students' solutions.

Table 4
Distribution of Students' Intra-task Strategy Flexibility Scores per Problem

	Age problem		Coins problem		Quiz problem	
Number of strategies	*f* (*n*=152)	%	*f* (*n*=152)	%	*f* (*n*=152)	%
0	93	61	92	61	75	49
1	53	35	42	28	56	37
2	6	4	16	11	19	13
3	–	–	2	1	2	1

5.2.3 Inter-task strategy flexibility and success

To explore the differences between students who showed inter-task strategy flexibility and students who used a single strategy across the problems on their problem solving performance we used the *t*-criterion for independent samples. The results provided evidence for prediction 5, since students who switched strategies across the problems ($\bar{x}$ = 1.35, *SD* = .82, *n* = 51) performed significantly better in problem solving (t = 2.70, p < .01) than students who persevered with the same strategy across the problems ($\bar{x}$ = 0.64, *SD* = 0.67, *n* = 11).

Moving a step forward, we examined whether students' performance varied as a function of the inter-task strategy flexibility score. As illustrated in Table 5, using analysis of variance (ANOVA) with performance as the dependent variable and score of inter-task strategy flexibility as the independent variable showed that problem solving performance was *generally higher* when the inter-task strategy flexibility was higher.

Table 5
Students' Mean Performance and Standard Deviation in Problem Solving per Inter-task Strategy Flexibility Score

Score of inter-task strategy flexibility	$\bar{x}$	*SD*	*n*
0	0.64	0.67	11
1	1.15	0.79	27
2	1.58	0.67	12
3	1.58	1.00	12

Post hoc analysis showed that there were statistically significant differences among the mean performances of the students who used the same strategy across the problems (Score: 0) and the students who changed strategies in two (Score: 2) or three (Score: 3) problem pairs. Students who exhibited high scores of flexibility, that is 2 or 3, performed equally well.

5.2.4 Intra-task strategy flexibility and success

To explore the differences between students who showed intra-task strategy flexibility and students who used a single strategy within the problems on their problem solving performance we used the *t*-criterion for independent samples. The results showed that the difference in performance between students who used various strategies ($\bar{x}$ = 1.17, *SD* = 0.85, *n* = 36) and students who used a single strategy ($\bar{x}$ = 0. 95, *SD* = 0.78, *n* = 73) within problems was not significant ($t = -1.36$, $p = .18$). This equivalence in performance deviates from prediction 6, that is, students who use different strategies would outperform students who use only one strategy in the solution of a problem. This is an interesting finding which motivated us to carry out a qualitative analysis of the responses of the students who showed intra-task strategy flexibility.

The purpose of the qualitative analysis is to understand why the intra-task flexibility did not support the students in finding the correct answer. For this analysis we selected the quiz problem (see Section 4.2), because it had the lowest proportion of correct answers of the three problems when multiple strategies were used. More specifically, out of the 21 students who applied more than one strategy only six came up with the correct answer.

Table 6 shows all the combinations of strategies that were used by the students in the quiz problem. The first thing that stands out is that repeating is the most frequently used strategy. In total, we found that the repeat strategy was applied ten times in combination with other strategies, which is the highest frequency among all problems, as already noted in Table 2. This probably reflects that the students had difficulties in grasping or keeping in mind the complex data structure of the problem.

Table 6
Distribution of the Applied Strategy Combinations in the Quiz problem and Correctness of the Answers

Applied strategy combinations	All (correct and incorrect answer)	Correct answer
	F	*f*
Rep, PoC	7	–
Extr, Sys	4	2
TE, PoC	3	2
Extr, PoC	2	1
Rep, Sys	2	–
Sys, Extr	1	–
Extr, TE, PoC	1	–
Rep, TE, PoC	1	1
Total	21	6

In seven cases the repeat strategy was followed directly by proof-or-check. In all these cases the students came up with an incorrect solution. This is remarkable because proof-or-check either on its own or in combination with other strategies is mostly connected with finding the correct outcome. Figures 2 and 3 show two examples of students who combined the repeat strategy with the proof-or-check strategy.

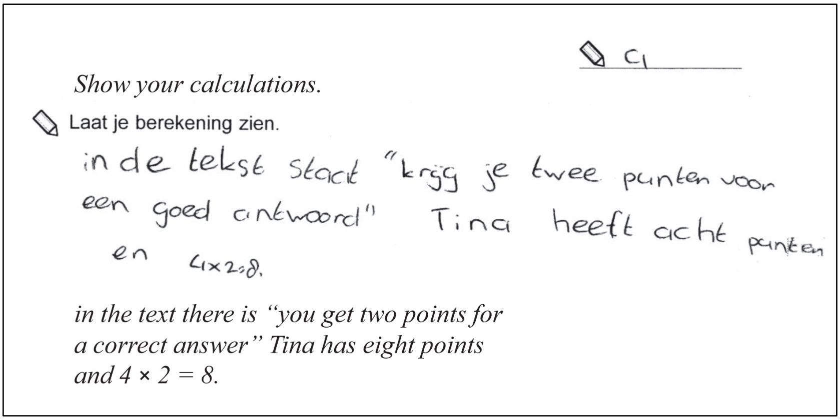


Fig. 2 Example of a combination of repeat strategy with proof-or-check strategy

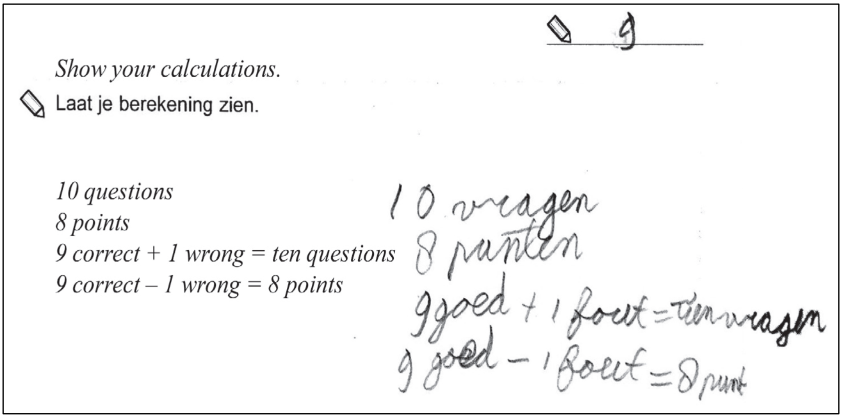


Fig. 3 Example of a combination of repeat strategy with proof-or-check strategy

These examples make clear that when the students repeated the problem information they did it in an incomplete, fragmented way. As a consequence, the students applied the proof-or-check strategy without taking into account all the problem information, therefore their results were incorrect. Specifically, the student whose work is shown in Figure 2 completely disregarded the incorrectly answered quiz questions, while the student in Figure 3 assumed that for a correct answer only one point is added to the score. Therefore, this student found that there were nine correct answers and one incorrect answer.

Among the students' responses, eight test sheets showed the use of the extreme strategy. Four times the extreme strategy was followed by a few systematic trials until an answer was found. The students that applied this combination of strategies in the quiz problem, assumed firstly that all ten answers were correct, resulting in a score of 20 points. Using the ten correct questions as a starting point, the students made a double list with the number of correct questions and the total points. Each step was usually one correct question less, until the score of eight points was attained. What can go wrong in this process though, is that the student may focus only on the points of the correct answers and disregard the penalty points for wrongly answered questions. Figures 4 and 5 depict the work of students who used the combination of extreme strategy and systematic listing strategy. The student in Figure 4 found the correct answer to the problem, while the student in Figure 5 came up with an incorrect answer.

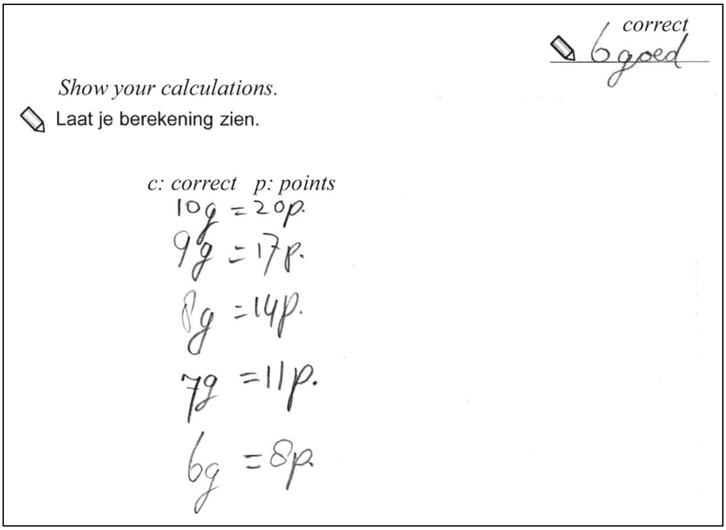


Fig. 4 Example of a combination of extreme strategy with systematic listing strategy

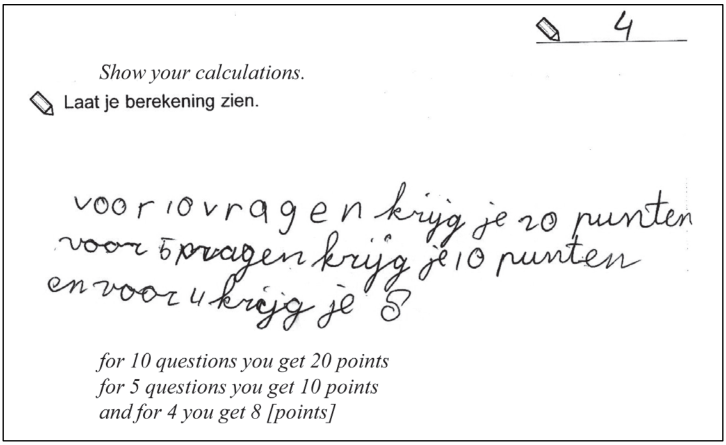


Fig. 5 Example of a combination of extreme strategy with systematic listing strategy

6 Discussion

The present study examined strategy use and strategy flexibility of high mathematical achievers in grade four when solving non-routine problems, which involved the co-variation of different variables. The relationships of strategy use and flexibility with success on these problems were also investigated.

The use of heuristic strategies in students' solutions was poor, despite students' high mathematical competence. This result is in line with previous studies' findings suggesting that heuristics are rarely used by students when confronted with unfamiliar complex problems (De Bock et al. 1998; Schoenfeld 1992; Verschaffel et al., 1999). The marginal place of non-routine problems in the Dutch mathematics textbooks could offer an explanation for this result (Kolovou et al., in press).

However, strategy use in this study was assessed on the basis of what was visible on the work space of the students' test sheets. Since only about half of the students made use of the work space when solving the problems, another explanation could be that students had difficulties in writing down their thinking. The tendency not to write down one's reasoning is a general attitude, as other studies showed similar results in different mathematical tasks (Doorman et al., 2007). For example, students might believe that it is better not to use the paper, because solving the problems mentally indicates a higher level of mathematics. Moreover, students probably have not learned to organize the data and write down the solution steps to support their thought process. This is especially true for high achievers who are not accustomed to use scrap paper when they deal with tasks during regular mathematics class work (Doorman et al., 2007).

Furthermore, our study provided us with some new insights concerning the successfulness of strategies. Findings showed that the trial-and-error strategy, although not very advanced, was the most broadly successful strategy. An explanation that can be given is that when students are not explicitly taught any heuristic strategies, trial-and-error may be the only strategy they can use, as it does not entail high cognitive demands and it is widely used in a variety of mathematical and everyday situations. Thus, students are more experienced and competent in using this strategy rather than other strategies. Systematic listing of possible solutions and proving or checking the answers were the other two strategies that had the potential to lead to success in some problems. It could be interesting for future studies to examine if the pattern between heuristic strategies and problem solving success changes when students receive systematic strategy training in non-routine problem solving.

One of the most important contributions of this study is the introduction and exploration of a new operational conceptualization of strategy flexibility in non-routine problem solving. Two distinct aspects of strategy flexibility were identified and examined: intra-task strategy flexibility and inter-task strategy flexibility. Students were not very often found to show traces of strategy flexibility either between or within the problems. This can be attributed to the problems' novelty and complexity, which may have hindered the flexible change of strategies by the students.

In concern to inter-task strategy flexibility, however, the number of students who changed strategies across the problems was larger than the number of students who persevered with the same strategy. This finding lends support to previous studies' findings which distinguished between solvers who extended or persevered with the use of a particular strategy and solvers who varied their strategy more frequently (Martinsen & Kaufmann, 1991). The larger number of flexible students may be attributed to the participants' high mathematical competence (Shore & Kanevsky, 1993).

As regards intra-task strategy flexibility, the majority of the students who made a solution attempt, used mainly one strategy and failed to consider any alternative or complementary strategies. This finding is in line with the study of Muir and Beswick (2005) which provided evidence for students' lack of any inclination to try an alternative strategy even when frustrated by their weakness to proceed in their solution. This inflexible behavior can be seen as an indication of students' deficiency to reflect on the appropriateness or adequateness of their initially chosen strategy and to use an alternate or complementary strategy that could lead to the correct answer.

The results of the present study showed that students who displayed inter-task strategy flexibility were more successful problem solvers than students who persevered with the same strategy between the problems. Furthermore, higher scores on problem solving performance were found with higher inter-task strategy flexibility. A possible explanation is that students who displayed inter-task strategy flexibility did not only possess more strategies, but could understand the rationale of these strategies, and therefore flexibly modify the procedures so that they were successfully used in different contexts (Baroody, 2003). This result has a practical implication for the teaching of non-routine problem solving in primary school. When solving different non-routine problems, even of a similar structure, it could be useful and effective (and probably less frustrating) for the students of this age to use multiple strategies. Yet more research is needed to find didactical methods to develop inter-task strategy flexibility in non-routine problems and to explore the impact of these types of instruction on problem solving performance.

Surprisingly, students who changed strategies within the problems were equally successful with the students who applied only one strategy for the solution of the problems. This finding suggests that intra-task strategy flexibility did not support the students in reaching a correct answer. A qualitative analysis of the intra-task strategy flexibility showed that comprehending the problem situation intervened with the solution strategies and therefore influenced the correctness of the answer. Specifically, when the students ignored or altered a part of the problem information, no matter how flexible they were in strategy use, they were not able to reach the correct answer. This means that knowledge and flexible use of multiple strategies could not lead to success unless sufficient understanding of the problem was achieved. This finding is in accord with Mayer's (1985) view that the construction of a complete mental representation of a problem is essential for a successful solution. A practical implication for problem solving instruction that may be deduced is that teachers could provide support to their students, so that they can master skills of sense making and organizing the information given in a non-routine problem, before rushing them to make decisions about the problem solving strategies.

Despite the fact that this study gives evidence for a number of conclusions, we need to emphasize that it is only an initial attempt towards the exploration of strategy flexibility in non-routine problem solving. Thus, further research is necessary before we can draw more firm and generalizable conclusions. First, the study included only three non-routine problems of a specific kind. If we want to have more evidence for students' strategy use and flexibility in non-routine problem solving, we need to investigate these aspects of behavior with various types of non-routine problems. A second issue of discussion is the limited number of students involved and their specific characteristics (fourth-grade high achievers from the Netherlands). To find more robust evidence for the findings of this study more students should be involved in the data collection. Future research could also investigate whether these findings about strategy use and flexibility vary with age, ability, and setting, by examining students of different grades, mathematical abilities, and educational systems. A final issue that needs further deliberation is how to measure students' strategy use and flexibility. In our study we focused on what was visible on students' test sheets. Students' inner thinking was not analyzed. In future studies more qualitative techniques could be used to collect data about students' cognitive processes, especially when alternating or maintaining strategies within or across problems.

Notes

1. CITO (Central Institute for the Development of Tests) provides Dutch schools with standardized tests for different subjects and grade levels. One of the CITO Tests is the Student Monitoring Tests for Mathematics. The DLE Test (Didactic Age Equivalent Test) is a different instrument published by Eduforce that teachers can use to measure their students' development in a particular subject.
2. The original versions of these problems have been developed for the World Class Tests. In 2004, Peter Pool and John Trelfall from the Assessment and Evaluation Unit, School of Education, University of Leeds who were involved in the development of these problems asked us to pilot them in the Netherlands.

3. The coding scheme was developed by two of the authors, Marja van den Heuvel-Panhuizen and Angeliki Kolovou, and our Freudenthal Institute colleague Arthur Bakker.
4. This control coding was done by Conny Bodin-Baarends who was involved in the data collection, but did not participate in the development of the coding scheme.

References

Altun, M., & Sezgin-Memnun, D. (2008). Mathematics teacher trainees' skills and opinions on solving non-routine mathematical problems. *Journal of Theory and Practice in Education, 4*(2), 213–238.

Baroody, A.J. (2003). The development of adaptive expertise and flexibility: The integration of conceptual and procedural knowledge. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills* (pp. 1–34). Mahwah, NJ: Lawrence Erlbaum Associates.

Beishuizen, M., Van Putten, C.M., & Van Mulken, F. (1997). Mental arithmetic and strategy use with indirect number problems up to one hundred. *Learning and Instruction*, *7*, 87–106.

Bodin, A., Coutourier, R., & Gras, R. (2000). CHIC : Classification Hiérarchique Implicative et Cohésive-Version sous Windows – CHIC 1.2. Rennes: Association pour la Recherche en Didactique des Mathématiques.

Cai, J. (2003). Singaporean students' mathematical thinking in problem solving and problem posing: An exploratory study. *International Journal of Mathematical Education in Science and Technology*, *34*, 719–737.

Carlson, M., & Bloom, I. (2005). The cyclic nature of problem solving: An emergent multidimensional problem-solving framework. *Educational Studies in Mathematics, 58,* 45–75.

Charles, R., Lester, F., & O'Daffer, Ph. (1992). *How to evaluate progress in problem solving.* Reston, VA: NCTM.

De Bock, D., Verschaffel, L., & Janssens, D. (1998). The predominance of the linear model in secondary school students' solutions of word problems involving length and area of similar plane figures. *Educational Studies in Mathematics, 35*, 65–83.

De Corte, E. (2007). Learning from instruction: The case of mathematics. *Learning Inquiry, 1*, 19–30.

Demetriou, A. (2004). Mind intelligence and development: A cognitive, differential, and developmental theory of intelligence. In A. Demetriou & A. Raftopoulos (Eds.), *Developmental change: Theories, models and measurement* (pp. 21–73). Cambridge, UK: Cambridge University Press.

Doorman, M., Drijvers, P., Dekker, T., van den Heuvel-Panhuizen, M., de Lange, J., & Wijers, M. (2007). Problem solving as a challenge for mathematics education in the Netherlands. *ZDM-The International Journal on Mathematics Education, 39*, 405–418.

Elia, I., Panaoura, A., Gagatsis, A., Gravvani, K., & Spyrou, P. (2008). Exploring different aspects of the understanding of function: Toward a four-facet model. *Canadian Journal of Science, Mathematics and Technology Education, 8*(1), 49–69.

Follmer, R. (2000). Reading, mathematics and problem solving: The effects of direct instruction in the development of fourth grade students' strategic reading and problem solving approaches to textbased, nonroutine mathematical problems, *Unpublished Doctoral Thesis* (Ed.D.). Chester PA: University of Widener.

Freudenthal, H. (1991). *Revisiting mathematics education*. Dordrecht: Reidel.

Gras, R., Suzuki, E., Guillet, F., & Spagnolo, F. (Eds.) (2008). *Studies in Computational Intelligence127: Statistical Implicative Analysis*. Heidelberg: Springer-Verlag.

Jonassen, J. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development, 48*(4), 63–85.

Kaizer, C., & Shore, B. (1995). Strategy flexibility in more and less competent students on mathematical word problems. *Creativity Research Journal, 8*(1), 77–82.

Kantowski, M.G. (1977). Processes involved in mathematical problem solving. *Journal for Research in Mathematics Education, 8*, 163–180.

Kolovou, A., Van den Heuvel-Panhuizen, M., & Bakker, A. (in press). Non-routine problem solving tasks in primary school mathematics textbooks-A needle in a haystack. *Mediterranean Journal for Research Mathematics Education.*

Krems, J.F. (1995). Cognitive flexibility and complex problem solving. In P.A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp. 201–218). Hillsdale, NJ.: Lawrence Erlbaum Associates.

Krutetskii, V.A. (1976). *The psychology of mathematical abilities in school children*. Chicago: University of Chicago Press.

Martinsen, O., & Kaufmann, G. (1991). Effect of imagery, strategy and individual differences in solving insight problems. *Scandinavian Journal of Educational Research, 35*(1), 69-76.

Mayer, R.E. (1985). Implications of cognitive psychology for instruction in mathematical problem solving. In E. A. Silver (Ed.), *Teaching and Learning Mathematical Problem Solving* (pp. 123–145). Hillsdale, NJ: Lawrence Erlbaum.

Muir, T., Beswick, K. (2005). Where did I go wrong? Students' success at various stages of the problem-solving process. Retrieved August 13, 2008, from http://www.merga.net.au/publications/counter.php?pub=pub_conf&id=143.

Pantziara, M., Gagatsis, A., & Elia, I. (in press). Using diagrams as tools for the solution of non-routine mathematical problems. *Educational Studies in Mathematics.*

Pape, S.J., & Wang, C. (2003). Middle school children's strategic behavior: classification and relation to academic achievement and mathematical problem solving. *Instructional Science, 31*, 419–449.

Polya, G. (1957). *How To Solve It; A New Aspect of Mathematical Method* (2nd ed.). Garden City: Doubleday.

Schoenfeld, A.H. (1983). The wild, wild, wild, wild, wild world of problem solving: A review of sorts, *For the Learning of Mathematics, 3*, 40–47.

Schoenfeld, A.H. (1985). *Mathematical problem solving.* San Diego, California: Academic Press.

Schoenfeld, A.H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D. Grouws (Ed.), *Handbook for Research on Mathematics Teaching and Learning* (pp. 334–370). New York: MacMillan.

Shore, B.M., & Kanevsky, L.S. (1993). Thinking processes: Being and becoming gifted. In K.A. Heller, F.J. Monks, & A.H. Passow (Eds.), *International handbook of research on giftedness and talent* (pp. 131–145). Oxford, England: Pergamon.

Silver, E.A. (1997). Fostering creativity through instruction rich in mathematical problem solving and problem posing. *ZDM Zentralblatt für Didaktik der Mathematik, 29*(3), 75–80.

Stacey, K. (1991). The effects on student's problem solving behaviour of long-term teaching through a problem solving approach. In F. Furinghetti (Ed.) *Proceedings of Fifteenth Psychology of Mathematics Education Conference* (Vol. 3, pp. 278–285). Assisi: PME.

Stanic, G, & Kilpatrick, J. (1988). Historical perspectives on problem solving in the mathematics curriculum. In R.I. Charles & E.A. Silver (Eds.), *The Teaching and Assessing of Mathematical Problem Solving* (pp. 1–22). Reston, VA: NCTM.

Torbeyns, J., Desmedt, B., Ghesquière, P., & Verschaffel, L. (in press). Solving Subtractions Adaptively by Means of Indirect Addition: Influence of Task, Subject, and Instructional Factors. *Mediterranean Journal for Research in Mathematics Education.*

Verschaffel, L., De Corte, E., Lasure, S., Van Vaerenbergh, G., Bogaerts, H., & Ratinckx, E. (1999). Learning to solve mathematical application problems: A design experiment with fifth graders. *Mathematical Thinking and Learning, 1*, 195–229.

Verschaffel, L., Luwel, K., Torbeyns, J., & Van Dooren, W. (in press). Conceptualizing, investigating, and enhancing adaptive expertise in elementary mathematics education. *European Journal of Psychology of Education.*

Chapter 3

Non-routine problem solving tasks in primary school mathematics textbooks – a needle in a haystack

Angeliki Kolovou [a], Marja van den Heuvel-Panhuizen [a, b], and Arthur Bakker [a]
[a] *Freudenthal Institute for Science and Mathematics education, Utrecht University, the Netherlands*
[b] *IQB, Humboldt University, Berlin, Germany*

Kolovou, A., Van den Heuvel-Panhuizen, M., & Bakker, A. (2009). Non-routine problem solving tasks in primary school mathematics textbooks - A needle in a haystack. *Mediterranean Journal for Research in Mathematics Education 8*(2), 29–66.

Non-routine problem solving tasks in primary school mathematics textbooks – a needle in a haystack

1 Introduction

1.1 Background of the study and its research question

Problem solving is the heart of mathematics (Halmos, 1980) and is supposed to play a crucial role in mathematics education. The significance of problem solving is evident in many curricula and educational policy documents (Stacey, 2005; and see the Special Issue of ZDM – The International Journal on Mathematics Education edited by Törner, Schoenfeld, and Reiss, 2007). Moreover, much attention is paid to the topic of problem solving in research literature (Lesh & Zawojewski, 2007; Lester & Kehle, 2003; Schoenfeld, 1985; Törner et al., 2007). As a consequence, problem solving is one of the key competences assessed in international assessments studies, such as TIMSS and PISA, which compare students' achievements in mathematics (Dossey, McCrone, & O'Sullivan, 2006).

For the Netherlands, the PISA findings on problem solving were rather disappointing. Dutch 15-year-old students scored relatively low on real-life problem solving, which in the PISA study covers a wide range of disciplines including mathematics, science, literature, and social studies. The students in the Netherlands were placed twelfth of 40 OECD-countries for problem solving, while for mathematics in general, they obtained the fourth position. It is also noteworthy that of all OECD-countries, the difference in scores between mathematics in general and problem solving is the largest in the Netherlands (HKPISA Centre, 2006; PISA-NL-team, 2006).

These PISA findings are more or less in agreement with the results from the Dutch POPO (Problem Solving in Primary School) project that aims at getting a better understanding of the mathematical problem solving performance of Dutch primary school students. The first study carried out in this project investigated the problem solving competences and strategies of the 25% best achievers in mathematics in grade 4 (Van den Heuvel-Panhuizen, Bakker, Kolovou, & Elia, in preparation). Unlike in PISA, problem solving was interpreted in this study as solving non-routine puzzle-like numerical problems. Preliminary results of the POPO study already revealed that the students have considerable difficulties with non-routine problem solving tasks that require higher-order mathematical thinking (Van den Heuvel-Panhuizen & Bodin, 2004; see also Doorman et al., 2007).

In the present study, we attempted to uncover a possible reason for Dutch students' poor performance in problem solving by investigating to what degree the students get the chance to learn problem solving. We explored this so-called "opportunity to learn" (Husen, 1967) by analyzing how much mathematical problem solving Dutch textbooks series provide to teachers and, consequently, to students. The guiding research question for our study was: *What proportion of the number-related tasks in textbooks documents can be qualified as problem solving tasks that require higher-order thinking?* In the aforementioned POPO study, student data were collected from the middle of grade 4, therefore we carried out the textbook analysis on the textbook series documents that are used in the first half year of grade 4.

1.2 Determining role of textbooks in the Netherlands

In the Netherlands, textbook series are published by commercial publishers and a school team can freely decide which textbook series to purchase and use. The textbook authors are free to determine the content and the layout of the textbooks, and are even free in choosing a particular underlying teaching principle. The only requirement for textbooks is that they are in agreement with the core goals published by the Dutch Ministry of Education. Should a textbook series not meet this criterion, the school inspectorate would give schools the advice not to use it. However, the core goals are rather limited in their description and leave much space for different interpretations. Consequently, there is much free space in interpreting these goals. Therefore, textbook series can differ greatly in how they translate these goals into teaching-learning activities and in how they are structured in different kind of documents. As a consequence, by including particular mathematical content and activities and excluding others, textbooks can influence students' opportunity to learn.

Another reason why textbook series affect to a great extent what is taught in Dutch classrooms is that Dutch teachers use their textbook series as a daily guide for organizing their teaching, both with respect to the teaching content and the teaching methods (Van den Heuvel-Panhuizen & Wijers, 2005). According to the last National Assessment of the Educational Achievement (PPON) (Janssen, Van der Schoot, & Hemker, 2005), almost all teachers reported that they follow the textbook and only rarely deviate from it.

The determining role of Dutch textbooks is also clear if we think of the innovation in mathematics education that has taken place in the Netherlands since the beginning of the 1970s and that had its breakthrough in the mid 1980s. This reform is largely attributed to the implementation of the new reform-based textbook series (De Jong, 1986; Van den Heuvel-Panhuizen, 2000).

In sum, we may say that there are enough reasons to look at the Dutch textbooks series when we want to understand why Dutch students have poor results in problem solving. Yet, the determining role of Dutch textbooks on what is taught is not the only reason for this low performance.

1.3 Growing international interest in textbooks

Worldwide, over the last decades, school mathematics textbooks and curriculum materials have received a growing interest. More and more, they have been found to be important factors in influencing the teaching of mathematics and the output of that teaching (Braslavsky & Halil, 2006; Cueto, Ramírez, & León, 2006; Doyle, 1988; Nicol & Crespo, 2006; Schmidt, McKnight, Valverde, Houang, & Wiley, 1997; Stein, Schwan, Henningsen, & Silver, 2000). For that reason, the Third International Mathematics and Science Study (TIMSS) (Schmidt et al., 1997) carried out a thorough analysis of the curriculum guides and textbooks of the 50 participating countries.

For the Dutch curriculum guides and textbooks, these findings of TIMSS (Schmidt et al., 1997) were rather daunting. The analysis revealed that for Population 1 (third and fourth-grade students) the main focus was on procedural skills. Problem solving was almost absent. Later, these results were confirmed by the national analyses of the TIMSS data. These showed that, by the end of grade 4, more than half of the time has been invested on numbers and only 13% on patterns, comparisons and connections (Meelissen & Doornekamp, 2004). However, at the same time the Dutch students in Population 1 showed high mathematics achievement scores in TIMSS (Meelissen & Doornekamp, 2004; Schmidt et al., 1997). The aforementioned facts – high general mathematics scores, low scores on problem solving, and mathematics textbook series with limited attention paid to problem solving – combined with the determining role that Dutch textbooks have in primary school mathematics education brought us to analyze more deeply what the textbooks offer students in terms of tasks that ask for mathematical problem solving.

Before we describe how the textbook analysis was carried out and what results we got from it, we continue with two literature reviews that guided the setup of our analysis. In order to obtain knowledge about how textbooks can contribute to students' ability for mathematical problem solving, we first needed to have a better image of what we mean by problem solving. Based on our experiences in the POPO project our focus is on non-routine, puzzle-like tasks that imply higher-order understanding and application of higher-order skills. At the primary school level, this distinction touches on the difference between plain arithmetic and mathematics. In other words, the tasks we had in mind were tasks where mathematics comes into numerical problem solving. The first literature review deals with this issue. The second review elaborates the concept of opportunity to learn and the ways in which this can be assessed.

2 Literature review

2.1 Problem solving

In the previous section we briefly explained how we interpreted problem solving in this study. Here we embed this interpretation in the existing research literature about problem solving. It will become clear that problem solving is not an unequivocal concept (Törner et al., 2007).

Some authors call every task a problem and use the definition of problem solving as the process from the givens to the goal in which the goal is finding the right and often the one and only solution (Moursund, 1996). In this interpretation, problem solving is seen as doing calculations with numbers that are presented either as bare numbers or in a context. The second, in particular, is often called problem solving, although it might be just solving rather straightforward word problems.

2.1.1 Non-routine character

Other interpretations of problem solving are more plausible. Many researchers have emphasized that in problem solving the path from the givens to the solution is not a straightforward one. Such an interpretation is reflected in TIMSS and PISA. For example, PISA 2003 focuses mainly on real-life problem solving covering a wide range of disciplines. The PISA researchers used the following definition of problem solving: "Problem-Solving is an individual's capacity to use cognitive processes to confront and resolve real, cross-disciplinary situations where the solution path is not immediately obvious and where the literacy domains or curricular areas that might be applicable are not within a single domain of mathematics, science or reading" (OECD, 2003, p. 156). In TIMSS 2003, problem solving is included in the cognitive domain of reasoning. Here, much attention is given to the non-routine character of the problems. "Non-routine problems are problems that are very likely to be unfamiliar to students. They make cognitive demands over and above those needed for solution of routine problems, even when the knowledge and the skills required for their solution have been learned" (Mullis et al., 2003, p. 32). In this interpretation, genuine problem solving is the counterpart of solving routine problems aimed at getting practice in particular methods or techniques and in problem settings that are more familiar to students.

Although at first glance the non-routine characteristic of problem solving problems may appear obvious, it is not clear-cut. The difficulty is that a non-routine task may itself become routine. According to Zhu and Fan (2006, p. 612) the characterization of a problem as a problem largely "depends on the person who is dealing with the situation". Stein et al. (2000) also mentioned prior knowledge and experiences as important factors when deciding what tasks can count as tasks for problem solving. Furthermore, we should be aware of the fact that tasks can have both routine and non-routine aspects (Mamona-Downs & Downs, 2005).

2.1.2 Genuine problems for students

What is essential in the case of real problem solving is that the problems are genuine problems for the students. In the words of Kantowski (1977, p. 163) this means that "[a]n individual is faced with a *problem* when he encounters a question he cannot answer or a situation he is unable to resolve using the knowledge immediately available to him. [….] A problem differs from an exercise in that the problem solver does not have an algorithm that, when applied, will certainly lead to a solution."

2.1.3 Interpreting the problem situation

When the problem on which the students have to work is really a problem for them – in the sense that it is not clear in advance which calculation has to be carried out – the solution process often requires many steps back and forth until the student is able to unravel the complexity of the problem situation. Furthermore, students have to be aware of how the given numbers or quantities relate to one another in order to find a way to the solution (O'Brien & Moss, 2007). This ability to find an underlying pattern in a problem was also recognized by Lesh and Zawojewski (2007) as a crucial aspect of problem solving. They emphasized that problem solving is a goal-directed activity that "requires a more productive way of thinking about the given situation […] The problem solver needs to engage in a process of interpreting the situation, which in mathematics means modeling" (p. 782). Similar thoughts were expressed by Kilpatrick, Swafford, and Findell (2001) who argued that the problem solving competence involves the construction of mental models.

2.1.4 Higher-order thinking

The aforementioned interpretation of problem solving, in which the problems are true problems and the solution strategy is not immediately clear at the moment that the problem is presented, is in line with our interpretation. We want to focus on problem solving as a cognitive activity that requires both an insightful approach to the problem situation and strategic thinking. In other words, there is something more involved in non-routine mathematical problem solving than carrying out a calculation in an appropriate way. Our point of view implies that problem solving is a complex activity that requires higher-order thinking and goes beyond procedural skills.

Several authors have elaborated the distinction between higher and lower types of cognitive engagement of students. More than thirty years ago, Skemp (1976, p. 2) discerned relational and instrumental understanding in which relational understanding is "knowing both what to do and why", whereas instrumental understanding is "rules without reasons". While instrumental understanding suggests memorizing an increasing number of procedures, relational understanding involves building conceptual structures. In a similar way, Stein et al. (2000) based their Task Analysis Guide on the difference between tasks of low-level and higher-level demands. The first category includes memorization tasks and algorithmic tasks unrelated to the underlying meaning, whereas the second requires engagement with conceptual ideas and complex, non-algorithmic thinking. Doing mathematics as a high-level cognitive demand includes tasks where there is no pathway suggested by the task and where the focus is on looking for the underlying mathematical structure. Although problem solving is located more on the side of the high-level demands, Silver (1986) reminds us that problem solving involves elements of both sides. Moreover, as stated by Stein et al. (2000), the cognitive demands of a task can change during a lesson. A task that starts out as challenging might not induce the high-level thinking and reasoning that was intended as the students actually go about working on it. However, according to them, in any case it is clear that challenging tasks appear to be a prerequisite to elicit high-level thinking.

2.1.5 To conclude this review on problem solving

In sum, we can say that although the interpretations differ, there is consensus that genuine problem solving refers to a higher cognitive ability in which a straightforward solution is not available and that mostly requires analyzing and modeling the problem situation. In order to be a true problem for students, it should not be a routine problem. On the other hand, the review makes it clear that the distinction between tasks with a low-level demand and a high-level demand is not fixed; the developmental level and experience of the students also determine whether a task is a true task for problem solving.

2.2 Opportunity to learn

In this section, we review relevant research literature related to procedures and methods that have been used for assessing what mathematical content is taught. The findings of this review are used for developing our textbook analysis instrument.

Many studies have shown that there is a strong correlation between the content that is taught and the achievements of the students (Leimu, 1992; Floden, 2002; Haggarty & Pepin, 2002; Törnroos, 2005; Cueto et al., 2006). Whether primary school students are able to solve non-routine mathematical problems will therefore largely depend on whether they have been taught to solve these kinds of problems. The generative concept behind the correlation between what is taught and what is learned is the so-called “opportunity to learn” (OTL). According to Floden (2002) the most quoted definition of OTL comes from Husen’s report of the First International Mathematics Study (FIMS). This report describes OTL as “whether or not ... students have had the opportunity to study a particular topic or learn how to solve a particular type of problem presented by the test” (Husen, 1967, pp. 162-163).

Although OTL seems to be a clear-cut concept, there are several reasons it is not. According to Schmidt et al. (1997) there is an intricate system of factors that affect the so-called the potential educational experiences. Moreover, cross-national comparisons of textbooks, teachers’ mediation and students’ access to the textbooks have shown that students get significantly different opportunities to learn (Haggarty & Pepin, 2002). However, “having an opportunity to learn is a necessary prerequisite for learning, but a learning opportunity is no guarantee of students really learning” (Törnroos, 2005, p. 325).

Despite the complexity of the concept, several ways of measuring OTL have been developed. Roughly speaking they include using teacher reports, document analysis, and classroom observations.

2.2.1 Questionnaires

The first measurements of OTL were based on questionnaires in which teachers had to indicate whether particular mathematical topics or kinds of problems were taught to students. Such questionnaires were used in the international comparative studies FIMS, SIMS and TIMSS, which were carried out by the International Association for the Evaluation of Educational Achievement (IEA) (Floden, 2002). An example of this approach is the study by Leimu (1992). He gathered OTL ratings from teachers who made an item-by-item judgment concerning exposure to a topic, or teaching of the knowledge and principles necessary for solving a problem, as it applied to the group of students in question. These OTL ratings included emphasis placed by teachers on particular contents and expected student success in those contents. A similar approach was applied in the Dutch version of TIMSS 2003 (Meelissen & Doornekamp, 2004), where OTL was used to evaluate whether the TIMSS test items fit the implemented curriculum. They selected 31 TIMSS test items and asked 129 teachers whether they would include these items in a test that would contain everything their students had been taught up to that moment. However, the interesting thing in Leimu's (1992) approach of OTL was that he also asked students whether they had had an opportunity to learn the contents required for a correct solution of the test items.

2.2.2 Curriculum and textbook analysis

Another approach applied in TIMSS was looking at what content is offered in curricula and – in connection with this – in textbook series. Schmidt et al. (1997, p.4) see the curriculum as "a kind of underlying 'skeleton' that gives characteristic shape and direction to mathematics instruction in educational systems around the world" and that provides "a basic outline of planned and sequenced educational opportunities."

A textbook analysis to measure the OTL was also applied in a study by Törnroos (2005) that examined whether the test items of TIMSS fit the curriculum. Each of the 162 test items for mathematics was judged on the question of whether the textbook contained adequate material to enable the student to answer the item correctly. A scale from 0 to 2 was used, where the codes ranked from 0 (inadequate material) to 2 (fully adequate material). The values 0, 1, and 2 were used to describe the opportunities to learn offered by the textbooks.

The method of curriculum and textbook analysis developed in TIMSS (Schmidt et al., 1997) is a natural extension of the informal analyses of curriculum guides and textbooks in earlier IEA studies. The basis for this analysis was a mathematics framework containing content areas, performance expectations, and discipline perspectives. This last term refers to what kind of ideas textbooks reflect about mathematics.

Apart from having a framework to classify what is in the curriculum and the textbook, a very crucial thing is that one first defines the unit of analysis. According to Schmidt et al. (1997), the first step in the document analysis process was to subdivide each document into smaller units of analysis on which more detailed analyses could take place. In textbooks, the most fundamental unit type was a lesson. Subsequently these units were subdivided into smaller blocks, containing narrative blocks, graphic blocks, exercise and question sets, suggested activities, and worked mathematical examples. After a document had been divided into units and blocks, each block was described by assigning codes based on relevant aspects of the mathematics framework.

A more fine-grained approach in defining the unit of analysis was used by Stein et al. (2000) and Cueto et al. (2006). They both considered a task the smallest unit of an activity in a workbook or a student's notebook. By a task, they meant every question that requires an answer from a student.

2.2.3 Classroom observations

More recently, direct observations of classrooms have been implemented to overcome some of the limitations of the approaches used in the aforementioned studies. For example, questionnaires are rather economical and simple for the purpose of large-scale administration and statistical analysis; however, it is difficult for teachers to describe classroom events and interactions using questionnaires (Hiebert et al., 2003). Furthermore, textbook analysis captures the influence of the written curriculum on learning, but "the influence of curriculum materials on student learning […] cannot be understood without examining the curriculum as designed by teachers and as enacted in the classroom" (Stein, Remillard, & Smith, 2007, p. 321).

Therefore, the TIMSS 1999 video study (Hiebert et al., 2003) as a supplement to the TIMSS 1999 student assessment has sampled eighth-grade mathematics lessons from six countries where students performed better than their peers in the United States on the TIMSS 1995 mathematics assessments. The TIMSS 1999 video study expanded on the earlier TIMSS 1995 video study which included only one country, Japan. In total, 638 mathematics lessons from seven countries (including the 1995 data from Japan) were analyzed in order to describe and compare teaching practices among countries. In particular, the TIMSS 1999 video study examined the structure and the mathematical content of the lessons, and specific instructional practices, all shaping students' learning opportunities. Furthermore, questionnaire items for teachers and students were designed to help understand and interpret the videotaped lessons.

Several codes were developed and applied to the video data regarding different aspects of teaching. "[The] coding of [the] classroom lessons was based on segmenting the lesson into meaningful chunks. This requires identifying a unit of classroom practice that can be identified reliably so that its beginning and end points can be marked" (Stigler, Gallimore, & Hiebert, 2000, p. 92). In the TIMSS 1999 video study, mathematical problems were the primary unit of analysis. Each mathematical problem was coded as addressing a specific topic and a scheme for coding procedural complexity was developed; problems were sorted into low, moderate, and high complexity (Hiebert et al., 2003). In addition, they examined whether problems required reasoning and the mathematical relationships between the problems were coded.

Classroom observations were also used in a study about OTL in Chicago's public schools (Smith, 1998). More specifically, this study addressed the issue of how teachers make use of school time to create learning opportunities for their students, because "the amount and quality of time available for instruction directly shapes school outcomes and student achievement" (p. 3). Data from three years of school and classroom observations from fifteen schools were coded as a series of activities segments. The observation framework included instructional and non-instructional activities, such as activities linked to academic and non-academic learning, classroom management, transition time, and so on. Furthermore, interviews with teachers and administrators were used as an additional source of information. Subsequently, these data were used to calculate the hours of instruction typically delivered to the students by Chicago's public schools.

2.2.4 To conclude this review on opportunity to learn

Generally spoken, three different methods have been used to measure OTL. These methods differ not only in their focus of analysis but also in costs, time, and the reliability of the collected data. Each method has its advantages and limitations. However, the review made clear that all three approaches require a framework for analyzing the content that is taught, and a unit of analysis. Despite the complexity of the learning processes which makes it not easy to say when learning takes place and — as a consequence — when students are offered an opportunity to learn, according to Hiebert and Grouws (2007, p. 379), "opportunity to learn can be a powerful concept that, if traced carefully through to its implications, provides a useful guide to both explain the effects of particular kinds of teaching on particular kinds of learning". In agreement with Hiebert and Grouws we think that opportunity to learn is "more nuanced and complex than simply exposure to subject matter" (p. 379). However, taking into consideration the determining role that textbooks play in the Netherlands, we decided to do a textbook analysis to investigate whether Dutch students encounter problem solving tasks that require higher-order thinking.

3 Method

3.1 Analyzed textbooks

In the textbook analysis that we carried out, we included the textbook documents of the first half year of grade 4 for the six main textbook series that are currently used in Dutch primary schools: *Pluspunt*, *De Wereld in Getallen*, *Rekenrijk, Talrijk*, *Wis en Reken*, and *Alles Telt*. The last National Assessment of the Educational Achievement (PPON) (Janssen et al., 2005) showed that approximately 40% of the Dutch primary schools were using the textbooks series *Pluspunt.* Nearly 20% were using *De Wereld in Getallen* and 15% were using *Rekenrijk*. Other textbook series were used by less than 5% of the schools. The textbook series analyzed are the same as those that were used in the schools participating in the earlier mentioned POPO study (Van den Heuvel-Panhuizen, Bakker, Kolovou, & Elia, in preparation) in which we investigated the problem solving of high-achieving students.

The six textbook series cover the grades 1 to 6 and also include documents for kindergarten classes, which are part of primary school in the Netherlands. Most of the textbook series consist of a two-volume lesson book (e.g. Student book 6A and 6B), additional documents such as workbooks and books with master pages meant for repetition or enrichment, and a teacher guide that explains how to use the textbooks series.

Table 1 gives a detailed overview of the documents for grade 4 (the Dutch “groep 6”) that were included in the analysis (for every textbook series, the first document in the list is the main book). The teacher guides and assessment materials were excluded from the textbook analysis.

Table 1
Textbook series documents included in the textbook analysis

Textbook series	Number of units	Number of pages	Page size
De Wereld in Getallen (WiG)			
Arithmetic book 6A	391	168	210×297 mm (A4)
Arithmetic workbook 6, Worksheets 1-14	20	16	210×297 mm (A4)
Total	*411*	*184*	
Talrijk (TR)			
Arithmetic book D1	329	120	210×220 mm
Workbook D1, D2	152	90	210×220 mm
Master copies book D, Series 1-4	273	128	210×297 mm (A4)
Total	754	338	

(*Table 1 continues*)

(*Table 1 continued*)

Textbook series	Number of units	Number of pages	Page size
Rekenrijk (RR)			
Student book 6A	475	144	218×227 mm
Workbook 6A	102	35	218×227 mm
Master copies book 6A	211	56	210×297 mm (A4)
Total	788	235	
Pluspunt (PP)			
Lesson book 6, Block 1-6	150	145	238×220 mm
Workbook 6, Block 1-6	58	24	210×297 mm (A4)
Book with assignments 6, Block 1-6	152	97	238×220 mm
Extra book 6, Block 1-6	109	97	238×220 mm
Total	469	363	
Wis en Reken (WeR)			
Math book 6.1	165	100	160×230 mm
Workbook 6.1	143	101	160×230 mm
Master copies book 6, Block 1-9	29	28	210×297 mm (A4)
Miscellaneous book 6.1	102	78	160×230 mm
Total	439	307	
Alles Telt (AT)			
Student book 6A	538	115	197×284 mm
Workbook 6, Block 1-3	109	31	210×297 mm (A4)
Master copies book 6, Block 1-3	133	52	210×297 mm (A4)
Total	780	198	

3.2 Textbook analysis instrument

The development of the textbook analysis instrument required in the first place that we identified what we consider the unit of analysis. Secondly, we had to define more precisely what we mean by problem solving tasks.

3.2.1 Unit of analysis

Since each textbook series differs with regard to the format and the number of the pages, we needed to determine a unit of analysis that fits all the six textbook series. This means that we had to decide how small or how large a unit should be. To avoid extremely large counts, we decided not to count every operation in a column (see Figure 1) but take a larger unit size.

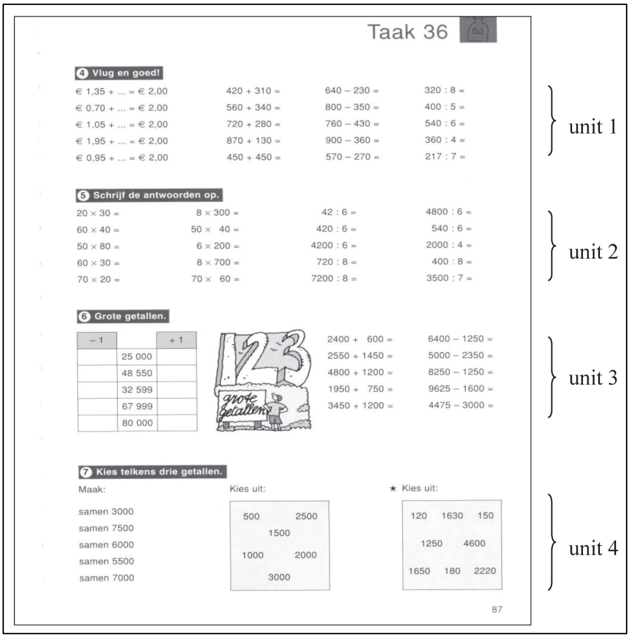


Fig. 1 *De Wereld in Getallen*, Arithmetic book 6A, p.87

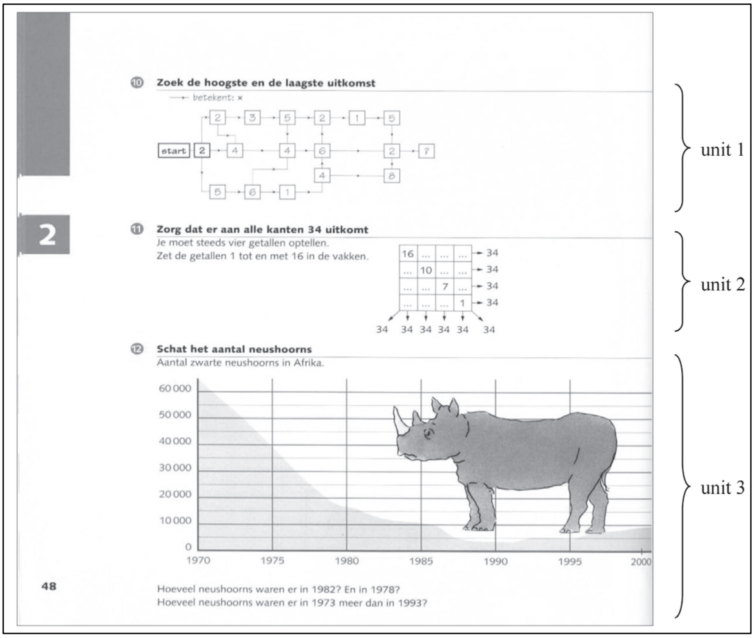


Fig. 2 *Rekenrijk*, *Student* book 6A, p. 48

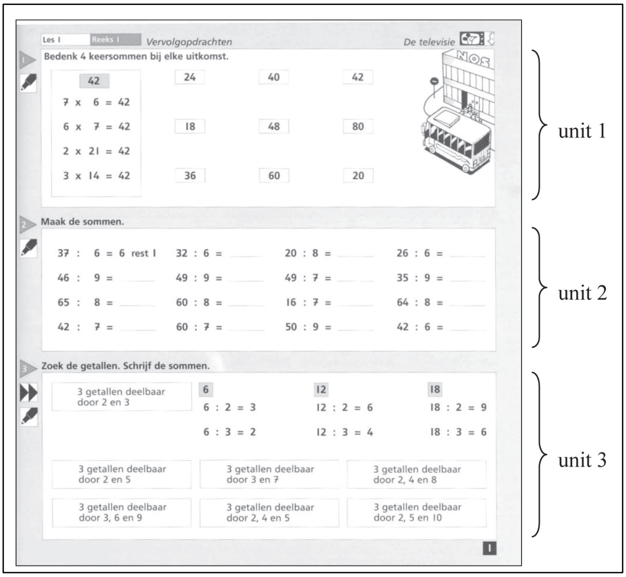


Fig. 3 *Talrijk*, Arithmetic *book* 6A, p. 1

3.2.2 Categories of problem solving tasks

The next step in the design of the textbook analysis instrument was the development of a framework of categories of problem solving tasks to classify the units. To develop a clear-cut definition of problem solving tasks, we needed several rounds. In the first round, we just marked the units that contain tasks that can be considered as non-routine mathematical problems. That means that we were looking for tasks that place a greater cognitive demand on students than tasks that merely require basic computational skills. Since we found extremely few of such genuine non-routine puzzle-like tasks in the textbooks, we decided to make an extra category for what we called "gray-area tasks". Next, we explain our categories more precisely.

The puzzle-like tasks include problems that do not have a straightforward solution, but that require creative thinking, for example, splitting a number into three or four successive numbers. Since fourth-graders do not have any algebraic tools at their disposal, they cannot apply a routine algebraic procedure, but have to tackle such tasks by a problem solving strategy, such as trial-and-error, or systematic listing of possible solutions.

The tasks that fall into the gray-area category are not really puzzles and are not really straightforward either, but can trigger strategic thinking and stimulate non-routine approaches. In other words, such gray-area tasks can provoke and prepare the development of problem solving strategies. Examples of gray-area tasks are problems in which the students have to investigate all possible combinations in which one can throw two dice, problems in which they have to search for a pattern in a series of numbers, and problems like the second task of the three following measurement tasks. These three tasks differ noticeably in cognitive demand and illustrate the difference between straightforward tasks, gray-area tasks, and puzzle-like tasks.

1. You have a soup cup (300 ml). How can you use it to measure 2100 ml of water?
2. You have a soup cup (300 ml), a mug (200 ml) and a glass (250 ml). Show different ways in which you can use these containers to measure 1500 ml of water.
3. You have a 5-liter and a 3-liter jug. How can you take 4 liters of water out of the big bowl using two jugs? You may pour water back into the bowl.

The first task requires the plain application of an algorithm to find that 7 cups make up 2100 ml. The second task can be solved by combining containers that make up the required quantity. Actually, this problem is an example of an "own-construction problem." This means that the students can explore the different ways to reach 1500 ml. However, in case all possible solutions are required, one has to construct a model and tackle the problem systematically. The third task, that is taken from *Alles Telt,* Student book 6A (p. 37), is a real puzzle because the solution is not a straightforward one. It requires building a model of the situation in which one has to find a sequence of steps to set apart an amount of 4 liters of water.

As the next step in developing the textbook analysis tool, we subdivided the two problem solving categories (puzzle-like and gray-area tasks) into more specific types. The puzzle-like tasks were partitioned in context problems and bare number problems. In both sub-categories, the "equations" form a main group. These problems – such as "Fill in the numbers: … − … = 3200. The first number must be the double of the second number" (*Rekenrijk*, Master *copies book 6A*, p. 3) – describe relationships between two or more variables or between quantities. When using algebra, these problems can be solved by means of equations with unknowns. Other sub-categories that have been distinguished are the "switch problems" (within the category of context problems) and the "magic frames" (within the category of bare number problems). The last sub-category, for example, includes tasks in which grids have to be filled with numbers in such a way that horizontally and vertically the totals are the same. The earlier discussed tasks in which a 5-liter and a 3-liter jug have to be used to get 4 liters of water, is an example of a "switch problem" because a quantity of water has to be transferred back and forth between the containers. Another example that can be considered a switch problem is the famous *Towers of Hanoi* problem.

The category of gray-area tasks is subdivided in tasks about numbers and operations, patterns, and combinatorics. The first sub-category includes non-algorithmic tasks with numbers, for example, making number sentences out of a given number of numbers and reasoning about calculation chains. To avoid getting too many sub-categories the gray-area tasks are not subdivided into context problems and bare number problems. The sub-categories include both kinds of problems.

The appendix shows the complete framework of problem solving tasks that we developed. It contains examples of tasks for each of the two categories (puzzle-like and gray-area tasks) and their sub-categories taken from the six analyzed textbook series.

3.2.3 Coding procedure

The framework of problem solving tasks served as the guideline for coding the units in the six textbook series. First, for each textbook series the total number of units was determined, then each of the units was classified according to the categories and subcategories included in the framework. If a unit neither fit the puzzle-like tasks nor the gray-area tasks no classification was given. In case a unit consisted of a sequence of tasks that included both of the two main categories (puzzle-like and gray-area problems), the highest category was coded. Moreover, if the tasks of a unit belonged to more than one sub-category (for example, within the category of puzzle-like problems the unit can contain context problems and bare number problems), then the most prevalent sub-category was coded.

The final coding was done by the first two authors. This was followed by a reliability check by the third author who was not involved in the development of the framework of problem solving tasks. The reliability check was based on a second coding of a part of the main books of the three textbook series. (*De Wereld in Getallen, Arithmetic book 6A; Pluspunt*, *Lesson book 6;* and *Alles Telt, Student book 6A*). In this selection we included all types of problems. This second coding was 96% in agreement with the coding of the first two raters.*

After the coding procedure was completed, for each textbook series in total and for each of the documents that belong to a textbook series, the absolute frequencies of all the categories were determined. Then, the relative frequencies were calculated reflecting what percentage of the total number of units belonged to a particular category.

4 Results

The most important results from the textbook analysis are that the textbooks differ in many aspects and that the majority of the number tasks included in the textbook series are straightforward problems and do not really require problem solving. Before we deal with this main result, we discuss some other differences between the textbook series that were revealed by our analysis.

Table 1 shows that the textbook series differ greatly in size. All the figures in this table belong to the textbook materials that are meant for half a year of teaching in grade 4. The first striking thing to note is that the textbook series do not have the same number of documents. *De Wereld in* Getallen consists of two books, while *Pluspunt* and *Wis en Reken* involve four books. The other three textbook series contain three books. The number of pages is also quite different. *De Wereld in Getallen*, *Alles Telt,* and *Rekenrijk* have approximately 200 pages, while *Pluspunt*, *Talrijk,* and *Wis en Reken* have more than 300 pages. We also found differences in the number of units. *Wis en Reken*, *De Wereld in Getallen*, and *Pluspunt* have between 400 and 500 units while *Alles Telt*, *Talrijk*, and *Rekenrijk* have almost twice that number. Several factors can explain this difference. Apart from differences in the number of books and the number and format of pages, the units in the textbook series do not look alike. *De Wereld in Getallen* has fewer units than *Rekenrijk*, but as can be seen in Figures 1 and 2, the units in the first textbook series contain more tasks than the second. Because of the difference in the number of units we only compared the presence of problem solving tasks in each of the six textbooks in relation to the total number of units in the textbook series.

The results from the coding show that the percentages of puzzle-like tasks – the tasks that require genuine problem solving – vary between the textbook series, but are extremely low (see Table 2 and Figure 4). They range from 2.43% to 0%. *De Wereld in Getallen* contains the highest percentage – 2.43% of the total units include puzzle-like problems tasks. *Rekenrijk* follows with 1.40%, whereas *Talrijk* and *Alles Telt* include less than 1% puzzle-like tasks (0.66% and 0.77% respectively). Two textbook series, *Pluspunt* and *Wis en Reken* do not contain any puzzle-like tasks at all. In these two textbook series all problem solving tasks belong to the gray-area category.

Table 2

Units in Dutch mathematics textbook series that contain problem solving tasks

	De Wereld in Getallen		*Talrijk*		*Pluspunt*		*Rekenrijk*		*Wis en Reken*		*Alles Telt*		Total	
	N	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%
Problem solving tasks														
Puzzle-like tasks	10	19	5	5	0	0	11	17	0	0	6	15	32	1^
Context problems	0		0		0		6		0		6		12	
Bare number problems	10		5		0		5		0		0		20	
Gray-area tasks	43	81	96	95	41	100	54	83	25	100	35	85	294	8^
Numbers and operations	24		84		25		40		22		17		212	
Patterns	3		2		0		1		0		16		22	
Combinatorics	16		10		16		13		3		2		60	
Total problem solving tasks	53	13^	101	13^	41	9^	65	8^	25	6^	41	5^	326	9^
Total units	411		754		469		788		439		780		3641	

^ percentage of the total number of units

If we take the whole category of problem solving tasks (that means puzzle-like problems and gray-area tasks) the frequency is still remarkably low. The textbook series *De Wereld in Getallen* and *Talrijk* hold the highest percentage of problem solving tasks – both 13%, whereas the other textbook series contain less than 10% problem solving tasks. *Alles Telt* includes the smallest percentage – just 5%.

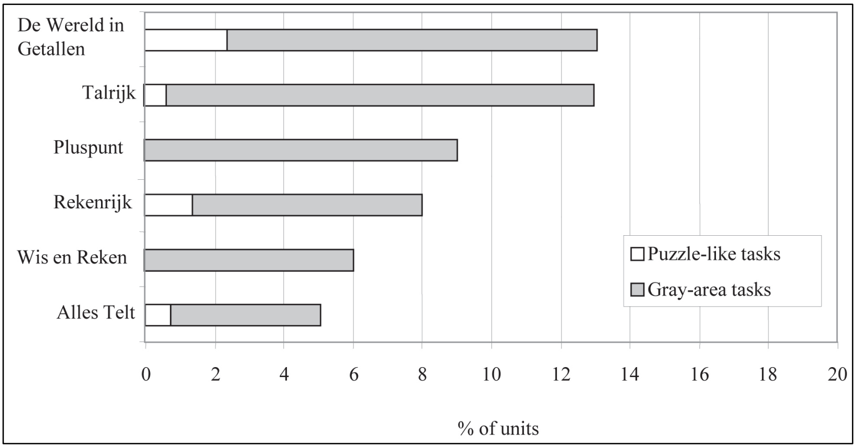


Fig. 4 Percentages of problem solving tasks (puzzle-like tasks and gray-area tasks) per textbook series

The results revealed that problem solving and – in particular – puzzle-like tasks have a marginal place in the six textbook series. Moreover, we found that the problem solving tasks (gray-area tasks and puzzle-like tasks together) are distributed differently over the different documents of the textbook series (see Figure 5). In some textbook series, the problem solving tasks are primarily included in the main book, while in other textbook series they are spread out over all documents. For example, in *De Wereld in Getallen* 96% of the problem solving tasks is in the main book, the *Arithmetic book*. This is different, for example, for *Pluspunt* where the *Book with assignments* and the *Extra book* contain more problem solving tasks (29% and 37% respectively) than the *Lesson book* (22%) which is the main book. In two of the textbook series, a relatively large part of the problem solving tasks are in the *Master copies books*: in *Talrijk* 43% and in *Rekenrijk* 46%. Having the problem solving tasks in these documents (mostly containing enrichment material) does not really guarantee that all students will get the opportunity to work on these tasks. The same is true for the problem solving tasks that are in the main book of *Talrijk*; mostly these tasks are denoted by a special symbol which means that the tasks are meant for the better students.

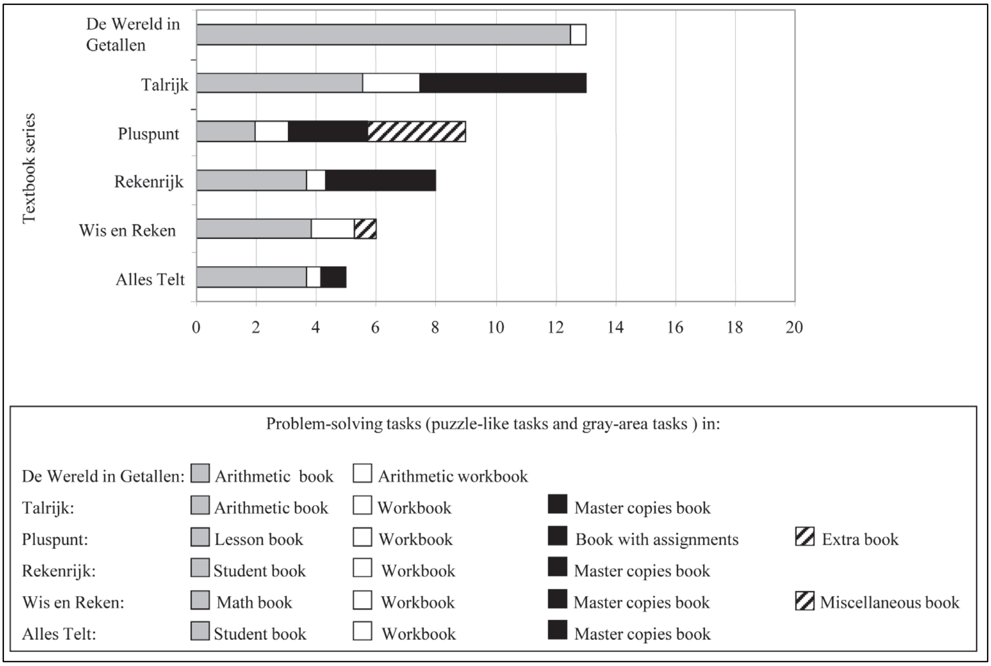


Fig. 5 Percentages of problem solving tasks (puzzle-like tasks and gray-area tasks) per textbook series per document

When we zoomed in on the special group of puzzle-like problems (see Figure 6) we found again that the textbook series differ in how these problems are distributed over the documents. Here, again *De Wereld in Getallen* stands out. This textbook series does not only have the largest proportion of puzzle-like problems, but also has these problems all in the main book, which probably gives the largest chance to students to work on these problems. *Rekenrijk*, on the contrary, has 64% of the puzzle-like problems in the *Master copies book*. This means that more than half of the puzzle-like tasks of *Rekenrijk* are to be found outside the main book. In *Talrijk* 40% of the puzzle-like tasks is also included in the *Master copies book*. To sum up, not only is the number of puzzle-like tasks in the Dutch textbook series very small, but these tasks are often not included in the main book either. In fact, one has to wonder whether the majority of Dutch students encounter any puzzle-like tasks at all.

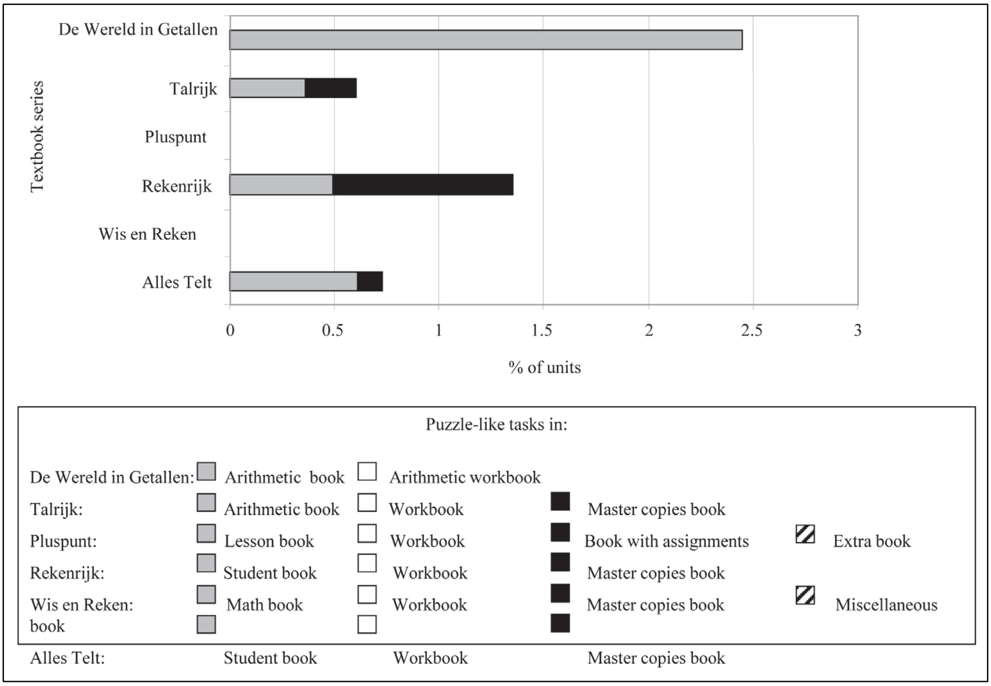


Fig. 6 Percentages of puzzle-like tasks per textbook series per document

5 Discussion

The disappointing performance of Dutch students in problem solving led us to scrutinize the main mathematics textbook series for grade 4. The percentages of problem solving tasks that we found in the six textbook series involved in our study – by which we covered the textbook series used by about 85% of the schools – were correspondingly disappointing. In the textbook series with the highest proportion of puzzle-like tasks, the percentage of these tasks was slightly over 2%. Even when we expanded the strict category of puzzle-like tasks with the gray-area tasks, the highest percentage found was 13%.

Because our study was aimed at investigating what textbooks have to offer to the students, we excluded additional materials that are not part of a textbook series. This might be a limitation of our study in cases where teachers do not stick to their textbook series documents, but also use additional instructional materials. According to the PPON report (Janssen et al., 2005) this is true for a large number of teachers. About two thirds of the teachers reported that they are using additional material for students who need extra support in mathematics. However, the question asked in the PPON study was clearly about additional material for practicing number operations. Therefore, it is not likely that these additional materials contain many puzzle-like problems. Consequently, we will not be far beside the truth when we say that, from the perspective of what is offered to students, non-routine problem solving tasks are rather scarce in Dutch primary school mathematics education.

Another limitation of our study is that we left out of our analysis how teachers interpret what is in the textbooks. According to Gilbert (as cited in Haggarty & Pepin, 2002) one can never conclude with confidence that what results from an analysis of a text is similarly realized in classrooms. Therefore, Gilbert emphasized that textbooks should be analyzed both in terms of their content and structure, and in terms of their use in classrooms. Earlier, Sosniak and Stodolsky (1993, p. 252), argued that "to understand textbook use, it is necessary to consider teachers' thought and action and their relationships, teachers' work within and across subjects, and the full context of teachers' conditions of work".

In our study, we restricted ourselves to analyzing the six textbook series with respect to the presence of problem solving tasks. Taking into account that genuine problem solving – that prepares for algebraic thinking – is not included in the Dutch core goals and is not assessed in the CITO End of Primary School test and the tests of the CITO Monitoring System, we think that the results of our textbook analysis reflect to a large degree what is happening in classrooms. In other words, we can assume that non-routine problem solving gets almost no attention in Dutch primary schools. This, however, contrasts sharply with theoretical and societal claims of the importance of problem solving.

Although our study addressed the situation in the Netherlands, discrepancies between the intended curriculum and the curriculum that is reflected in the textbook series can also be present in other countries, as was, for example, recently revealed by an Australian study on proportional reasoning (Dole & Shield, 2008). In this Australian study, it was explored to what degree proportional reasoning was promoted by mathematics textbooks, and – similar to our study – the researchers found a predominance of calculation procedures with relatively few tasks to support conceptual understanding.

Disclosing possible inconsistencies between what we value as important to teach our students and the instructional materials we use to reach these educational goals, is of crucial importance to improve our teaching. Like Dole and Shield (2008, p. 33) we see textbook analysis as "a potential means to raise awareness of instruction in key topics within the school mathematics curriculum" and consequently as a vital tool for educational progress. To realize this potential, further research is needed in this research domain of textbook analysis, which unfortunately and erroneously has a somewhat outmoded and moldy image, but from which we can learn so much.

---

* After we published this article in the Mediterranean Journal for Research in Mathematics Education, we computed the Cohen's kappa to get a more robust measure of the interrater reliability of coding of the tasks in the textbooks. We found a Cohen's kappa of .94.

References

Braslavsky, C., & Halil, K. (Eds.) (2006). Textbooks and Quality Learning for all: Some Lessons Learned from International Experiences. Genève: International Bureau of Education, UNESCO.

Cueto, S., Ramírez, C., & León, J. (2006). Opportunities to learn and achievement in mathematics in a sample of sixth grade students in Lima, Peru. *Educational Studies in Mathematics*, *62,* 25–55.

De Jong, R. (1986). Wiskobas in methoden [Wiskobas in textbooks]. Utrecht: OW&OC, Utrecht University.

Dole, S., & Shield, M. (2008). The capacity of two Australian eight-grade textbooks for promoting proportional reasoning. *Research in Mathematics Education, 10*, 19–35.

Doorman, M., Drijvers, P., Dekker, T., Van den Heuvel-Panhuizen, M., De Lange, J., & Wijers, M. (2007). Problem solving as a challenge for mathematics education in the Netherlands. *ZDM The International Journal of Mathematics Education, 39*, 405–418.

Dossey, J.A., McCrone, S.A., & O'Sullivan, C. (2006). *Problem Solving in the PISA and TIMSS 2003 Assessments* (NCES 2007-049). Washington, DC: U.S. Department of Education, National Center for Education Statistics. Retrieved November 11, 2007, from http://nces.ed.gov/pubsearch.

Doyle, W. (1988). Work in mathematics classes: The context of students' thinking during instruction. *Educational Psychologist*, *23*, 167–180.

Floden, R.E. (2002). The measurement of opportunity to learn. In A.C. Porter & A. Gamoran (Eds.), *Methodological advances in cross-national surveys of educational achievement* (pp. 231-266). Washington, DC: National Academy Press.

Haggarty, L., & Pepin, B. (2002). An investigation of mathematics textbooks and their use in English, French and German Classrooms: Who gets an opportunity to learn what? *British Educational Research Journal*, *28,* 567–590.

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

Hiebert, J., Gallimore, R., Garnier, H., Givven, K. B., Hollingsworth, H., Jacobs, J., et al. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 Video Study* (NCES 2003–013 Revised). Washington, DC: U.S. Department of Education, National Center for Education Statistics. Retrieved November 11, 2007, from http://nces.ed.gov/pubs2003/2003013.pdf.

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning: a project of the National Council of Teachers of Mathematics* (Vol. 1, pp. 371–404). Charlotte, NC: Information Age Publishing.

HKPISA Centre (n.d.). Pisa *2003 cross-disciplinary problem solving framework.* Hong Kong: Author. Retrieved October 29, 2007, from http://www.fed.cuhk.edu.hk/~hkpisa/sample/sample.htm.

Husen, T. (Ed.). (1967). *International study of achievement in mathematics: A comparison of twelve countries* (Vol. II). New York: John Wiley & Sons.

Janssen, J., Schoot, F. van der, & Hemker, B. (2005). Periodieke peiling van het onderwijsniveau (PPON). Balans (32) van het reken-wiskundeonderwijs aan het einde van de basisschool 4 [Periodical assessment of the educational level. Balance (32) of the mathematics education at the end of the primary school 4]. Arnhem, the Netherlands: Cito Instituut voor toetsontwikkeling.

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

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.

Leimu, K. (1992). Explorations in opportunity-to-learn. Finnish national analyses of IEA/SISS data. *International Journal of Educational Research, 17*(3)*,* 291–317.

Lesh, R., & Zawojewski, J.S. (2007). Problem solving and modeling. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning : a project of the National Council of Teachers of Mathematics* (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.M. Doerr (Eds.), *Beyond constructivism: Models and modeling perspectives on mathematics problem solving, learning and teaching* (pp. 501–518). Mahwah, NJ: Erlbaum.

Mamona-Downs, J., & Downs, M. (2005). The identity of problem solving. *Journal of Mathematical Behavior, 24*, 385–401.

Meelissen, M.R.M., & Doornekamp, B.G. (2004). *Nederland in TIMSS-2003: Leerling prestaties in rekenen/wiskunde en natuuronderwijs in groep 6 van het basisonderwijs* [The Netherlands in TIMSS-2003: Student achievement in mathematics and natural sciences in grade 4 of the primary education]. Enschede, the Netherlands: Universiteit Twente.

Moursund, D. (1996). *Increasing Your Expertise as a Problem Solver: Some Roles of Computers* (2nd edition). Eugene, OR: International Society for Technology in Education.

Mullis, I.V.S., Martin, M.O., Smith, T.A., Garden, R.A., Gregory, K.D., Gonzalez, E.J., et al. (2003). *TIMSS assessment frameworks and specifications 2003* (2nd edition). Chestnut Hill, MA: Boston College.

Nicol, C.C., & Crespo, S.M. (2006). Learning to teach with mathematics textbooks: How preservice teachers interpret and use curriculum materials. *Educational Studies in Mathematics*, *62* , 331–355.

O'Brien, T.C. & Moss, A. (2007). What's basic in mathematics? *Mathematics Teaching*, *202*, May, 26–27.

OECD (2003). *The PISA 2003. Assessment framework- mathematics, reading, science and problem solving knowlegde and skills*. Paris: Author. Retrieved October 29, 2007, from http://www.oecd.org/dataoecd/46/14/33694881.pdf.

PISA-NL-team (2006). *Wiskundige geletterdheid volgens PISA. 1. Analyse - hoe staat de vlag erbij?* [Mathematical literacy according to PISA. 1. Analysis - what is the situation?] Utrecht, the Netherlands: Freudenthal Institute.

Schmidt, W.H., McKnight, C.C.,Valverde, G.A., Houang, R.T., & Wiley, D.E. (1997). *Many visions, many aims: A cross-national investigation of curricular intentions in school mathematics*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Schoenfeld, A.H. (1985). *Mathematical problem solving*. New York: Academic Press.

Silver, E.A. (1986). Using Conceptual and Procedural Knowledge: A Focus on Relationships. In J. Hiebert (Ed.), *Conceptual and Procedural Knowledge: The Case of Mathematics* (pp. 181–198). Hillsdale, NJ: Lawrence Erlbaum.

Skemp R.R. (1976). Relational understanding and instrumental understanding. *Mathematics Teaching, 77*, 20–26.

Smith, B. (1998). It's About Time: Opportunities to Learn in Chicago's Elementary Schools. Chicago, IL: Consortium on Chicago School Research. Retrieved May 7, 2008, from http://ccsr.uchicago.edu/publications/p0f03.pdf.

Stacey, K. (2005). The place of problem solving in contemporary mathematics curriculum documents. *Mathematical Behavior, 24*, 341–350.

Sosniak, L.A., & Stodolsky, S.S. (1993). Teachers and textbooks: Materials use in four fourth-grade classrooms. *The Elementary School Journal, 93*(3), 249–275.

Stein, M.K., Remillard, J., & Smith, M.S. (2007). How curriculum influences student learning. In F.K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning: a project of the National Council of Teachers of Mathematics* (Vol. 1, pp. 319–369). Charlotte, NC: Information Age Publishing.

Stein, M.K., Schwan, S., Henningsen, A., & Silver, E. (2000). *Implementing standards-based mathematics instruction*. New York: Teachers College Press.

Stigler, J.W., Gallimore, R., & Hiebert, J. (2000). Using Video Surveys to Compare Classrooms and Teaching Across Cultures: Examples and Lessons From the TIMSS Video Studies. *Educational Psychologist, 35*, 87–100.

Törner, G., Schoenfeld, A.H., & K.M. Reiss (2007). Problem solving around the world: summing up the state of the art. *ZDM The International Journal on Mathematics Education, 39*, 353.

Törnroos, J. (2005). Mathematical textbooks, opportunity to learn and student achievement. *Studies in Educational Evaluation, 31*(4), 315–327.

Van den Heuvel-Panhuizen, M. (2000). *Mathematics education in the Netherlands: A guided tour*. FI-ICME-9 cd-rom. Utrecht, the Netherlands: Freudenthal Instituut.

Van den Heuvel-Panhuizen, M., Bakker, A., Kolovou, A., & Elia, I. (in preparation). Getting a better understanding of mathematical problem solving by high-achieving primary school students: lessons for and from the Netherlands.

Van den Heuvel-Panhuizen, M., & Bodin-Baarends, C. (2004). All or nothing: Problem solving by high achievers in mathematics. *Journal of the Korea Society of Mathematical Education, 8*(3), 115–121.

Van den Heuvel-Panhuizen M., & Wijers M. (2005). Mathematics standards and curricula in the Netherlands. *ZDM International Journal of Mathematics Education*, *37*, 287–307.

Zhu, Y., & Fan, L. (2006). Focus on the Representation of Problem Types in Intended Curriculum: A Comparison of Selected Mathematics Textbooks from Mainland China and the United States. *International Journal of Science and Mathematics Education, 4*, 609–626.

Appendix

Puzzle-like tasks and gray-area tasks

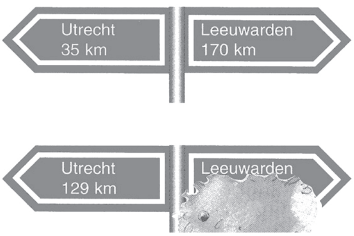
PUZZLE-LIKE TASKS

Context problems

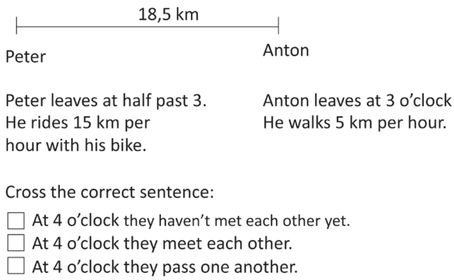
Equations

How many three- and four-wheeled buggies are there in the shop if the total number of wheels is 125? Can you find different possibilities?
(*Alles Telt, Student book 6A, p. 36*)

There is €23400 in a suitcase. How many €200 and €500 banknotes are there if the €200 banknotes are twice as many as the €500 banknotes?
(*Alles Telt, Student book 6A, p. 38*)



Joris cycles from Utrecht to Leeuwarden. On his way he sees this sign.
How many kilometers is it from Utrecht to Leeuwarden?
Some hours later Joris sees this sign. How much more must he cycle?
(*Alles Telt, Student book 6A, p. 51*)



(*Rekenrijk, Master copies book 6A, p. 1*)

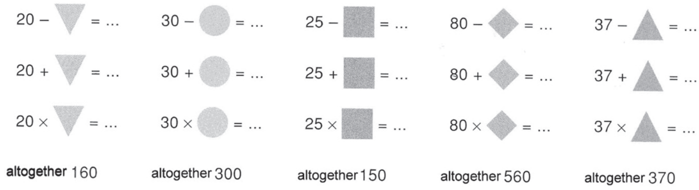
Switch problems

How can you take exactly 4 liters water out of the bowl using a 5 liter- and a 3 liter-jug? You can pour water back to the bowl.
(*Alles Telt, Student book 6A, p. 37*)

Bare number problems

Equations

Fill in the numbers. The first number must be the double of the second number: ……… − ……….. = 3200
(*Rekenrijk, Master copies book 6A, p. 3*)



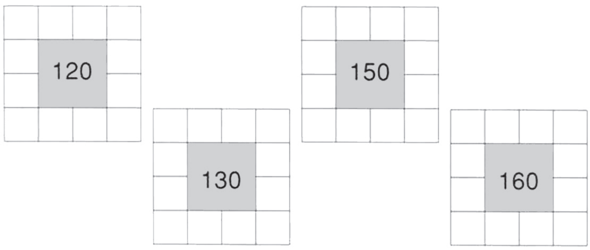
Three times the same number.
(*De Wereld in Getallen, Arithmetic book 6A, p. 36*)

sum (+)	difference(-)	number a	number b
1	2		5
13	3		

(*Rekenrijk, Master copies book 6A, p. 15*)

Find 3 successive numbers that make up the result.
……. + …… + ……. = 270
(*Talrijk, Arithmetic book D1, p. 12*)

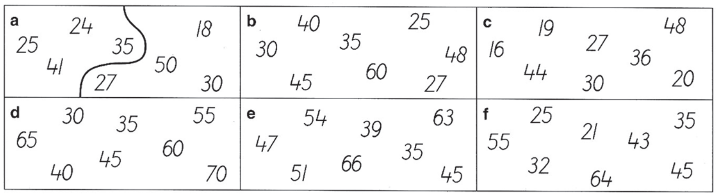
Magic frames



Use each of the numbers 10, 20, 30, 40, 50, 60 two times.
In every small square there is only one number. In the middle of the squares there is the sum of the rows and the columns.
(*De Wereld in Getallen, Arithmetic book 6A, p. 67*)

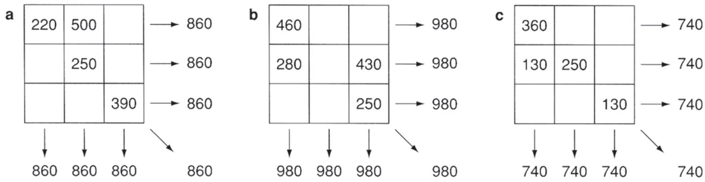
GRAY-AREA TASKS

Numbers and operations



a 24 + 35 + 25 + 41 = 18 + 27 + 50 + 30

Make two groups of numbers that have an equal value.
(*De Wereld in Getallen, Arithmetic book 6A, p. 37*)



(*Pluspunt, Workbook 6, p. 15*)

Try the calculation chain with three different numbers. What strikes you? Explain. Think of a calculation chain yourself.
start → think a number → add 6 → multiply by 2 → subtract 12 → half the number→ done
(*Rekenrijk, Student book 6A, p. 140*)

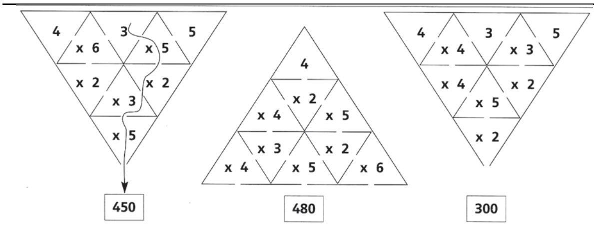
When Jelmer had spent the half of the half of his money, he had the half of €150 left. How much money did he have at the beginning?
(*Rekenrijk, Master copies book 6A, p. 38*)

Find the numbers that are equally distant from 7500:
9750, 6950, 7950, 7050, 6925, 8075, 4050, 10950, 8050, 5250

less than 7500	more than 7500	difference with 7500
…………..	…………..	……………

(*Rekenrijk, Master copies book 6A, p. 15*)

Lodewijk has to pay €56, 55. He pays with three notes of 20 euro and some extra coins and gets back €5. How much money has he paid in total?
(*Rekenrijk, Master copies book 6A, p. 11*)



Find the path with the correct product.
(*Talrijk, Master copies book D, p. 7*)

Cross out 3 digits so that you get the biggest 4-digit number: 9150362
(*Talrijk, Master copies book D, p. 110*)

Do sums. Use 150, 20, 5, +, −, = and make 120, 280, 165, 275.
(*Wis en Reken, Math book 6.1, p. 30*)

Patterns

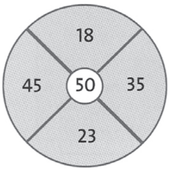
How many blocks from each color do you need for a tower with three floors? Fill in the table.

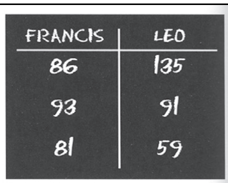
floors	red	blue	yellow
1	1	0	0
2	2	1	0
…	…	…	…

(*Alles Telt, Student book 6A, p. 45*)

Fill in the number line:
112 – 119 – 152 – … – … – … - 232 – … – … – …
(*De Wereld in Getallen, Arithmetic book 6A, p. 102*)

Combinatorics



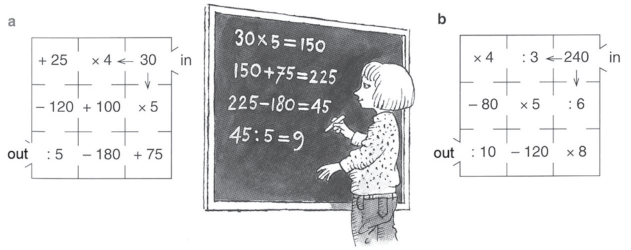


Francis and Leo throw darts. In each turn they throw 3 darts. In the first turn Francis got 86 points. How did he throw the darts? Are there different ways?
(*Alles Telt, Student book 6A, p. 62*)

Pay the exact amount. Try it in at least five ways. Draw the money.
(*De Wereld in Getallen, Arithmetic book 6A, p. 59*)

3 1 6 8
Use all digits. How many different numbers can you make?
(*Pluspunt, Lesson book 6, p. 43*)

How do you pack 68 eggs in boxes of 4, 6, and 10 eggs?
Search for different ways.
(*Pluspunt, Lesson book 6, p. 47*)



In every maze find 6 different routes and results.
(*Pluspunt, Extra book 6, p. 21*)

Joop sold sausages (€1, 50 each), pea soup (€2, 75 per cup) and coffee (€1, 25 per cup) for €880. How many sausages, cups of soup and coffee did he sell? Are different answers possible?
(*Rekenrijk, Student book 6A, p. 14*)

Take a pack of cards and remove all the jokers and the cards between 2 and 6. You want two cards of the same color. How many cards do you have to pull at the most?
(*Rekenrijk, Master copies book 6A, p. 14*)

Use two dice. In what ways can you throw 4, 7, and 10?
(*Rekenrijk, Master copies book 6A, p. 16*)

Draw all the possible sketches of a building that consists of 4 stones.
(*Talrijk, Arithmetic book D1, p. 26*)

Chapter 4

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

Angeliki Kolovou [a], Marja van den Heuvel-Panhuizen [a, b], Arthur Bakker [a], and Iliada Elia [c]

[a] *Freudenthal Institute for Science and Mathematics Education, Utrecht University, the Netherlands*

[b] *IQB, Humboldt University Berlin, Germany*

[c] *Department of Education, University of Cyprus*

Kolovou, A., Van den Heuvel-Panhuizen, M., Bakker, A. & Elia, I. (2008). An ICT environment to assess and support students' mathematical problem-solving performance in non-routine puzzle-like word problems In A. Gagatsis (Ed.), *Research in Mathematics Education* (pp. 175–190). Nicosia, Cyprus: University of Cyprus.

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

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 were 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).

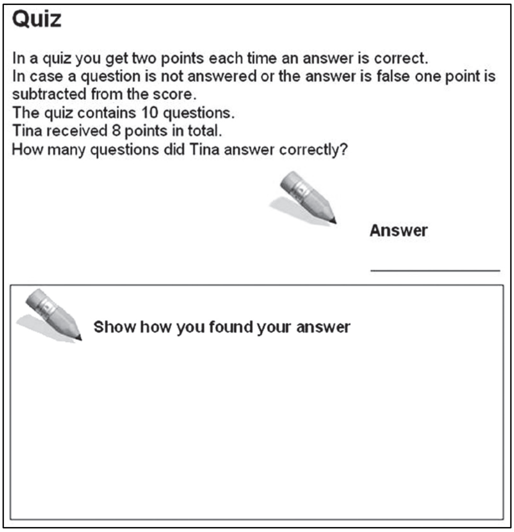


Fig. 1 Problem used in the POPO Study

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 offer 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). 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 pretest-posttest 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 pretest and posttest. 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 Pretest and posttest

The test that was used as pre-test and posttest 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.

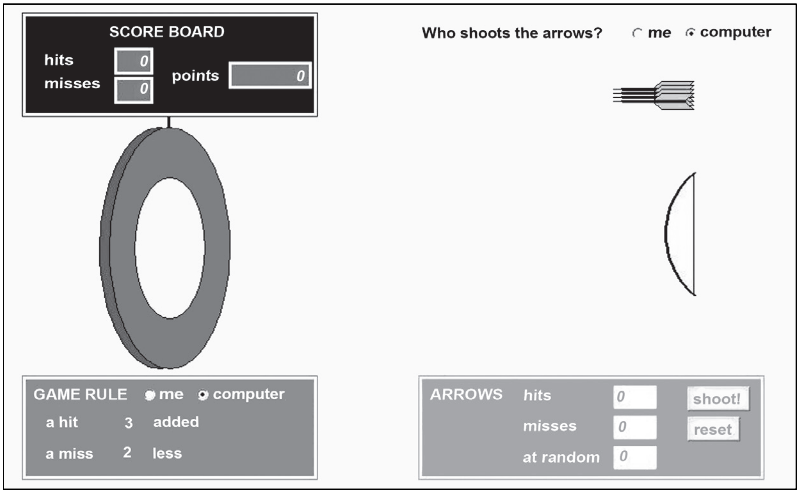


Fig. 2 Screen view of applet in the computer-shooting mode

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.

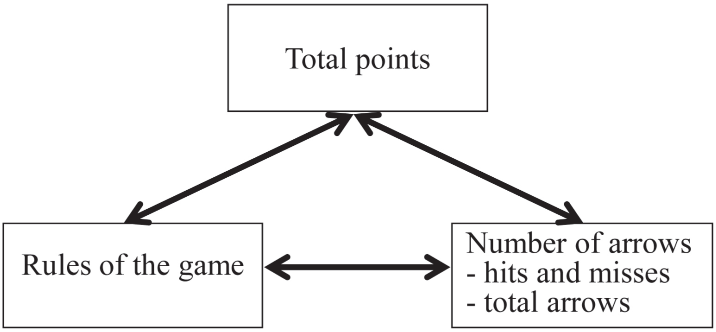


Fig. 3 Variables involved

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) were about the arrows while the rules of the game and the gained points were known. In the third question (C), which consisted of two parts, the rules of the game were 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. 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 took place when answering the second part of Question C. The students found 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 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 pretest and the posttest 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.

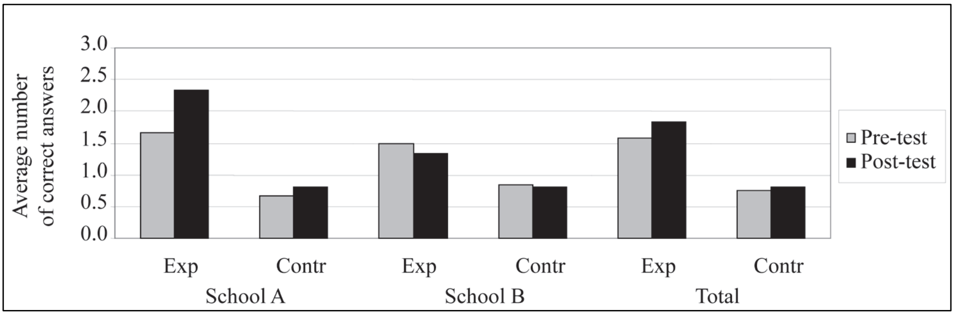


Fig. 4 Average number of correct answers per student in the pre and the posttest in both groups

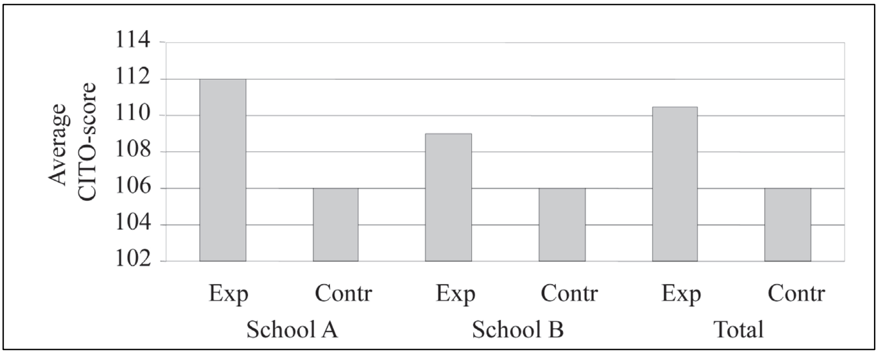


Fig. 5 Average CITO score of the experimental and control group

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 was a considerable improvement in the scores of the posttest. 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 the difference there 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. 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.

Note

1. The applet is programmed by Huub Nilwik.

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 April 20, 2008, from http://www.nottingham.ac.uk/education/MARS/papers/whither/index.htm.

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: Unique contributions of computers to innovative mathematics education. *Journal of Mathematical Behavior, 19*, 9−47.

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

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

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*, 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: Kluwer Academic Publishers.

Kantowski, M.G. (1977). Processes Involved in Mathematical Problem Solving. *Journal for Research in Mathematics Education, 8*, 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. (1966). The Use of Computers in Education. *Scientific American*, *215*, 206–220.

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 in Mathematics 66*, 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.

Chapter 5

Online game-generated feedback as a way to support early algebraic reasoning

Angeliki Kolovou and Marja van den Heuvel-Panhuizen
Freudenthal Institute of Science and Mathematics Education, Utrecht University,the Netherlands

© Kolovou, A., & Van den Heuvel-Panhuizen, M. (2010). Online game-generated feedback as a way to support early algebraic reasoning. *International J. Continuing Engineering Education and Life-Long Learning 20*(2), 224–238.

Online game-generated feedback as a way to support early algebraic reasoning

1 Introduction

Developing algebraic reasoning is an important goal in mathematics education (Katz, 2007). Algebra is considered as foundational for all mathematics and science and a subject that allows further access to technical professions. In fact it is seen as a gateway to later achievement. Despite the fact that algebra comprises a substantial part of the mathematics curriculum in secondary school, many students encounter considerable problems in learning algebra. For example, the equal sign (Booth, 1984) and different conceptions of variable (Malisani & Spagnolo, 2009) have often been mentioned as sources of difficulties for beginning algebra students.

While some researchers have attributed these obstacles to a cognitive gap between arithmetic and algebra (Herscovics & Linchevski, 1994), others have ascribed them to the late and abrupt introduction of school algebra (Kaput, 2007). In the primary grades the focus in mathematics curricula is on developing arithmetic and computational fluency, which in secondary school is followed by an isolated and superficial procedural approach to algebra (Kaput, 2007). This is an undesirable situation, because it does not offer students a continuing learning pathway. Therefore, several researchers (Kaput, 2007; Smith & Thompson, 2007) have supported the integration of algebra into the early mathematics curriculum. Introducing algebraic reasoning in the early grades prepares students for algebra courses in middle and high school and is essential for accomplishing coherence, depth, and power in school mathematics (Kaput, 2007).

In the Netherlands, early algebra is not included in the key goals of the primary school mathematics curricula, and students hardly ever deal with (non-routine) problems that could offer them an entry point in learning algebra. Previous research has shown that the mathematics textbook series contain only a small percentage of non-routine tasks (Kolovou, Van den Heuvel-Panuizen, & Bakker, 2009). Consequently, it is no surprise that even the more able Dutch primary school students encounter difficulties in solving non-routine problems (Van den Heuvel-Panhuizen & Bodin-Baarends, 2004). Similar results were reported in the PISA study. Although Dutch students attain high scores in mathematics, their performance in problem solving is significantly lower (PISA-NL-team, 2006). Naturally, this situation raises much concern and needs improvement.

The POPO project was set up to investigate ways to improve the problem solving competence of primary school students and to offer them opportunities to prepare for the learning of algebra in secondary school. The focus of this project is on non-routine contextual number problems with interrelated values. These problems can be solved by setting up and solving equations with unknowns. In contrast with secondary school students who have the option of applying formal algebra, primary school students may solve these problems through informal algebraic reasoning. To offer students opportunities to gain experiences with this kind of reasoning, we provided them with an online computer game. The students played the game at home and got instant feedback on the outcome of their actions. In this article we describe how the game-generated feedback supported students' learning in solving early algebra problems.

2 Teaching algebra in the elementary grades

The introduction of algebra into the elementary grades has been approached from various, often overlapping perspectives, such as generalizing arithmetic, introducing variables and covariation in word problems, and focusing on functions. For example, Carraher and Schliemann (2007) argue that early algebra entails a shift away from computations on particular numbers and measures, and towards reasoning about relations among sets of numbers and measures. In line with these definitions of early algebra, students should be engaged "in specially designed activities, so that they can begin to note, articulate, and represent the general patterns they see among variables" (Schliemann et al., 2003, p.128). Notably, rich problem contexts pose the challenge of generating abstract knowledge about mathematical objects and structure from experience and reasoning in specific situations (Carraher & Schliemann, 2007).

Information and communication technology (ICT) is recognised as a tool and environment that can play a crucial role in the development of early algebraic reasoning. Researchers have suggested that Logo programming environments (Clements, 2000) and spreadsheets (Sutherland & Rojano, 1993) support the understanding of variables and functional relations. As such, computer software makes the transition from arithmetic to algebraic reasoning possible (Clements, 2000). Recently, research on exploiting ICT in education has shifted its attention to the potential of computer games to enhance students' mathematical learning (Lee & Chen, 2009). Computer games create engaging learning experiences and present content in meaningful and relevant contexts, which makes the learning environment more powerful.

Notwithstanding the potential of computer games, research that reports on the use of games for the learning of (early) algebra is scarce. Nevertheless, some encouraging results have been achieved. A study by Lach and Sakshaug (2004), which included some online games, showed that students' abilities in solving problems involving spatial sense and algebraic reasoning improved by playing games. Another study carried out by Figueira-Sampaio, dos Santos, and Carrijo (2009) investigated sixth graders who used a computer tool – consisting of a virtual balance – for solving first degree equations. The study showed that the use of the tool enhanced students' motivation, involvement, cooperation, discussion, and reflection.

3 The role of feedback

Feedback is considered to have a very powerful effect on learning and achievement (Bransford, Brow, & Cocking, 2000; Hattie & Timperley, 2007). What is meant by feedback can differ, but it is generally understood as the information provided by an agent (teacher, peer, book, parent, self, experience) regarding some aspects of one's performance (Hattie & Timperley, 2007; Kluger & DeNisi, 1996). Characteristic of this information is that it allows the comparison and evaluation of one's actual performance relative to a standard of performance (Kluger & DeNisi, 1996). Furthermore, according to Shute (2008) feedback can include two types of information, namely verification and elaboration. Verification can occur explicitly (indicating the correctness of a response) or implicitly (showing the result of a student's action, for example within a simulation). In the case of elaboration, the feedback includes, among other things, discussing particular errors, providing worked examples, or giving guidance.

Hattie and Timperley (2007) argue that the effect of feedback depends on the level(s) at which the feedback is directed. For example, feedback at the level of the task can be powerful, especially when it gives students information about erroneous hypotheses, but it relies heavily on the students' ability to interpret the information. Other forms of effective feedback are feedback about the processing of the task and feedback about self-regulation; the latter form of feedback is related – among other things – to the students' ability to create internal feedback and to self-assess, which both characterise the more effective learners. Internal feedback, which is generated as learners monitor their engagement in tasks, affects the way externally provided feedback is perceived and utilised (Butler & Winne, 1995). Since self-regulation plays a fundamental role in problem solving, feedback that influences students' self-regulation might be essential for successful problem solving.

However, the effects of feedback on performance are by no means straightforward. Kluger and DeNisi (1996) showed that these effects are influenced by several variables, such as the nature of the feedback cues (e.g., whether they are positive or negative, what the content is, and what the frequency is), the task characteristics (e.g., novelty, complexity), and situational variables (e.g., goal setting). Furthermore, the mediating role of cognitive engagement on feedback effects has been emphasised (Butler & Winne, 1995).

4 Computer-assisted feedback

As technologies change rapidly, new opportunities for providing and using feedback for learning emerge. Bransford et al. (2000) suggest five ways that new technologies can contribute in establishing effective learning environments. Offering opportunities for feedback is one of these ways. In particular, technology provides increasing opportunities for learners to receive feedback from tutoring systems, teachers, and peers, to reflect on their own learning processes, and to revise and improve their understanding and reasoning.

Regarding problem solving, several studies (Bottino, Ferlino, Ott, & Tavella, 2007; Clements, 2000; Lee and Hollebrands, 2006) have pointed out the importance of computer-assisted feedback. For example, feedback can support the students in error comprehension and allows immediate verification of the correctness of their actions (Bottino et al., 2007) Furthermore, technology tools can provide feedback for reflection, so that students can take more control of their problem solving (Lee & Hollebrands, 2006). This type of feedback differs from the feedback in which students' problem solving is guided by hints and prompts (Lee & Hollebrands, 2006). This distinction resembles the idea of Shute (2008) who juxtaposed verification feedback and elaboration feedback.

Actually, a similar distinction is made by Nathan (1998). He distinguishes between feedback that is delivered by an intelligent tutoring system (ITS) and feedback that is generated from an unintelligent tutoring system (unITS). An ITS compares the student's response with the performance of an expert. Nathan calls this type of feedback knowledge-based. The diagnosis of the students' problem solving and suggestions for improvement are presented directly to the students. On the contrary, a unITS just shows the consequences of students' actions and the students have to interpret this information. Therefore, this type of feedback, called situation-based feedback, requires that the students use their own knowledge and perform self-assessment and error detection. The students then debug their solutions until an acceptable solution is found. Experimental results showed that students who had worked in unITS that incorporate the situation-based type of feedback performed significantly higher in solving algebra story problems.

In general, the role of computer feedback to support students' learning has been investigated when the computer was used in school. Little research has been done in a setting – for instance, at home – where no teacher is present to guide students' learning processes. As a consequence, hardly anything is known about how students regulate their problem solving processes based only on the feedback they receive from a computer.

5 Research design

5.1 Research question

The present study addressed the following research question: Can feedback resulting from playing an online game at home contribute to students' ability in solving early algebra problems?

5.2 Method

To answer these research questions, we first collected data with a paper-and-pencil test on early algebra consisting of seven non-routine contextual number problems with interrelated values. During and after doing this test, students received no feedback, neither on their answers, nor on their problem solving processes. After one week, each of the students got a unique account and a password in order to log in to an online environment and play with a computer game. They also got a set of problems to solve with the aid of the computer game at home. Because we wanted to create a situation similar to regular home computer use, the students could decide themselves whether or not they went online and how they played the game. While playing the game, the students were automatically informed about the consequences of their actions and their activities were recorded by special monitoring software. After four weeks the paper-and-pencil test was administered again.

5.2.1 Participants

Five grade 6 classes from five schools in a big city in the Netherlands participated in the study. Of the 123 students who were invited to go to the online environment, 96 students (46 boys and 50 girls) logged in. Their average age was 11.8 years (*SD* = .55) and their general mathematical ability as measured by the CITO E5 test was 116.0 (*SD* = 10.1). This score corresponds to level B, which ranges from 111 to 117 points and includes students whose score lies between the 50th and 75th percentile. The students who logged in did not differ from the total group of students in respect with the CITO E5 test score (*M* = 115.8, *SD* = 10.4), the students' age (*M* = 11.8, *SD* = .53) and the boys-girls ratio (62 boys and 61 girls).

5.2.2 Paper-and-pencil test

The paper-and-pencil test consisted of seven non-routine contextual number problems with interrelated values. The following two problems were included in the test:

> *'Quiz' problem*: In a quiz you get two points for each correct answer. If a question is not answered or the answer is wrong, one point is subtracted from your score. The quiz contains ten questions. Tina received eight points in total. How many questions did Tina answer correctly?
>
> *'Hit the target' problem*: You get two points if your arrow hits the white part in the middle of the target. If you do not hit it, one point is taken away from your score. Anne shot ten arrows. She has in total eight points. How often did Anne shoot in the target?

These problems can be solved by a variety of approaches, ranging from trial-and-error to a more sophisticated algebraic approach. However, the latter approach is unlikely because primary school students do not possess any algebraic tools. Instead, they can apply informal, context-connected solutions. These strategies and the accompanying notations can provide important entry points for learning algebra.

5.2.3 The computer game and the ICT environment

The environment that was developed to give students experience in dealing with interrelated varying values includes an archery game, called 'Hit the target'[1]. The default game screen (see Figure 1) displays the following objects: a target, a pile of arrows and a bow, a score board, a frame where the students can fill in the game rule, and a frame where the students can fill in the number of hits, misses, and randomly shot arrows.

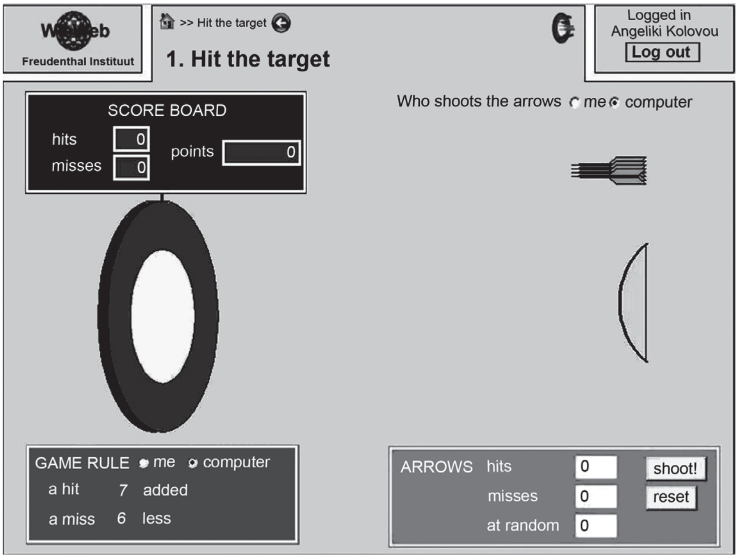


Fig. 1 Screen view of game in the computer-shooting mode

The students have the opportunity to define the game settings. For example, they can choose between user-defined or computer-defined game rule. In the user-defined game rule they can fill in the points that are added or subtracted in case of a hit or a miss. In the computer-defined game rule, the computer sets the rule randomly. The students can also determine the shooting mode by choosing between the user-shooting mode and the computer-shooting mode. In the user-shooting mode the students can shoot arrows one by one by dragging them to the bow and monitor their score after each shot. In the computer-shooting mode the arrows are shot at once according to the entered values and the scoreboard updates rapidly to inform the students about their score. This dynamic game feature provides feedback to the students. As a matter of fact, the game does not diagnose the students' answers; instead, the feedback includes information that the students can use to compare the values on the scoreboard with the intended values. In other words, the game provides situation-based feedback, which the students must interpret and can use to modify their solution if necessary.

The ICT environment was connected to the so-called digital mathematics environment (DME)[2]. This software traces students' actions while working online. The log data of each student consisted of a list of events (i.e., shooting actions) carried out by the student in the online environment. Furthermore, the date, time, and duration of the online work were registered. Figure 2 shows how a student solved a problem by several attempts.

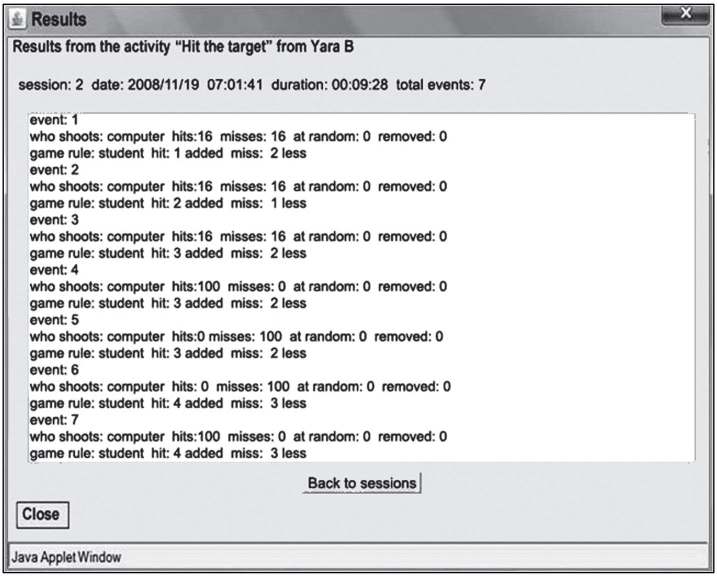


Fig. 2 Screen view of the log data generated by the DWO

The students had a three-week period to work in the online environment. Each week they received a number of problems; 14 in total (eight problems of which six consisted of two parts). These problems were given to the students to guide their work with the game and varied from finding the pair of hits and misses that produces a particular score, to generating a general rule by systematizing all solutions. For some problems more than one correct answer was possible, while others had only one correct answer. For example, the following problems were given to the students:

Problem 3a What is the game rule to get 15 points in total with 15 hits and 15 misses?
Problem 3b Are there other game rules to get 15 hits, 15 misses and 15 points?
Problem 4a What is the game rule to get 16 points in total with 16 hits and 16 misses?
Problem 4b Are there other game rules to get 16 hits, 16 misses, and 16 points?
Problem 8 For every hit you gain 2 points and for every miss 1 point is taken away from your score. You have 10 arrows in total. How many hits and misses do you have to shoot to get 5 points in total?

When the students were verifying their solution to a problem they generally compared the number of hits, misses, and the total score on the scoreboard with the values in the problem statement.

The aim of this computer-based intervention was to make the shift from the students' focus on particular values to a focus on relations between values. This means that – in spite of the fact that the numbers involved in a problem vary – the students may become aware of the invariant structure of the problem. Comprehending this invariant structure is an important aspect of algebraic reasoning. In this way, algebra as understanding patterns, relations, and functions (NCTM, 2000) can be taught in the early grades.

6 Results

In this section we first discuss the results of the *Quiz* and the *Hit the target* problem that were included in the paper-and-pencil test. Then, we relate these results to the results of five online problems that we consider as key problems in supporting the students' problem solving.

6.1 Pre and posttest

Of the 96 students who were logged in, 85 did both the pretest and the posttest. As shown is Table 1, in the pretest one third of the students came up with a correct answer in both problems. In the posttest, however, half of the students answered both problems correctly. Table 1 also displays the strategies that students used to come to a solution (correct or wrong). Some students applied more than one strategy. The use of trial-and-error and systematic trialing was rather limited and generally did not change between pre and posttest. The contrary was the case for the proof-or-check strategy, which was applied by a considerable percentage of students (41% to 51%). This strategy implies that students take their answer and perform a calculation with it to test whether the result fits the given number values in the problem. The frequency of applying this strategy was larger in the posttest than in the pretest. The above pattern in strategy use applied with minor differences to both problems.

Table 1
Results from the *Quiz* and the *Hit the target* problem

	Quiz		*Hit the target*	
	% students in pretest (N = 85)	% students in posttest (N = 85)	% students in pretest (N = 85)	% students in posttest (N = 85)
Correct answers	32	53	34	54
Strategies for all answers				
Trial-and-error	9	8	5	4
Systematic trialing	13	14	13	8
Proof-or-check	45	51	41	51

Furthermore, in the posttest most strategies were more successful than in the pretest (see Tables 2 and 3). In the *Quiz* problem 54% of the cases of strategy use were related to the correct answer, while in the posttest this occurred in 71% of the cases. In the *Hit the target* problem these percentages were respectively 56% and 74%.

Table 2
Relation of strategy use and successfulness in the *Quiz* problem

Strategies	Number of cases in which a strategy was applied in pretest			Number of cases in which a strategy was applied in posttest		
Trial-and-error	8	Answer-C	4	7	Answer-C	7
		Answer-W	4		Answer-W	0
Systematic trialing	11	Answer-C	6	12	Answer-C	9
		Answer-W	5		Answer-W	3
Proof-or-check	38	Answer-C	21	43	Answer-C	28
		Answer-W	17		Answer-W	15
Total cases	57	Answer-C	31 (54%)	62	Answer-C	44 (71%)
		Answer-W	26 (46%)		Answer-W	18 (29%)

Table 3
Relation of strategy use and successfulness in the *Hit the target* problem

Strategies	Number of cases in which a strategy was applied in pretest			Number of cases in which a strategy was applied in posttest		
Trial-and-error	4	Answer-C	2	3	Answer-C	3
		Answer-W	2		Answer-W	0
Systematic trialing	11	Answer-C	6	7	Answer-C	4
		Answer-W	5		Answer-W	3
Proof-or-check	35	Answer-C	20	43	Answer-C	32
		Answer-W	15		Answer-W	11
Total of cases	50	Answer-C	28 (56%)	53	Answer-C	39 (74%)
		Answer-W	22 (44%)		Answer-W	14 (26%)

6.2 Online working

Because the game play was optional, the 92 students who were logged in and did the pretest, did not process all the online problems. For example, the five online problems that we consider as key problems were only processed each by 30% to 50% of these students. Table 4 also shows that in general the involvement in the online problems was higher for the students who solved the *quiz* problem correctly in the pretest. Another finding is that 29% of the students showed an improvement on this problem between the pre and posttest.

To find out how the students developed while working online, we traced the online activities of the 11 students who worked on all the five online problems. Table 5 presents the types of solution strategies that the students applied to solve the problems. We distinguished the following types of solution strategies: one attempt (O), more than one attempt (M), complete information use (C) (which implies a correct solution), incomplete information use (I) (which implies a wrong solution), trial-and-error (TandE), and systematic trialing (SYS).

Table 5 shows that the students very often (in 70% of the cases) found a correct solution in the first attempt and that they did not apply the same strategy in every problem.

Table 4
Results from the logged-in students in the *Quiz* problem and their participation in the online game

Number (%) of logged in students				Number of students (%) who worked on the problems in the online game				
Pretest		Posttest		P 3a	P 3b	P 4a	P 4b	P 8
Answer-C	30	Answer-C	27 (90)	15 (50)	12 (40)	20 (67)	18 (60)	12 (40)
		Answer-W	0					
		Missing	3 (10)					
Answer-W	62	Answer-C	18 (29)	31 (50)	23 (37)	22 (35)	15 (24)	16 (26)
		Answer-W	40 (65)					
		Missing	4 (6)					
Total	92		92	46 (50)	35 (38)	42 (46)	33 (36)	28 (30)

Notes: Answer-C = correct answer; Answer-W = wrong answer

Table 5
Types of solution strategies applied by the 11 students in the five online problems and their scores on the *Quiz* problem in the pretest and posttest

Students	Solution types					Score *Quiz* problem	
	P 3a	P 3b	P 4a	P 4b	P 8	Pretest	Posttest
St 1	OC	OC	MCTandE	MCTandE	OI	0	0
St 2	OC	MCTandE	OC	OC	MCSYS	1	1
St 3	MCSYS	OC	MCTandE	OC	MCSYS	1	1
St 4	MCTandE	OC	OC	OC	OC	1	1
St 5	OC	OC	OC	OC	OC	1	1
St 6	OC	OC	OC	OC	OC	–	0
St 7	MCTandE	MCTandE	OC	OC	OC	1	1
St 8	OC	MCTandE	MCTandE	OC	MCTandE	0	1
St 9	OC	OC	MCTandE	OC	OC	0	1
St 10	MCTandE	OC	OC	OC	OC	0	1
St 11	OC	OC	OC	MCTandE	OC	0	0
Total						5	8

Notes: 1 = correct; 0 = wrong

For example, student 1 (see Figure 3) found a correct solution in Problems 3a and 3b with one attempt, applied a trial-and-error strategy in Problems 4a and 4b, and gave a wrong answer to Problem 8 in one attempt. Evidently, for this student the online game was not very helpful. Both in the pretest and the posttest she could not solve the *Quiz* problem.

Student 3 (see Figure 4) found the correct answer to the *Quiz* problem in the pretest and posttest. Her online working showed a smooth pattern of first applying multiple attempts in Problems 3a and 4a followed by an one-attempt answer in Problems 3b and 4b.

Student 8 (see Figure 5) is one of the students who clearly gained from the online game. She solved Problem 3a in one attempt and applied multiple attempts to solve Problems 3b and 4a. Then, she solved Problem 4b in one attempt and answered Problem 8 by applying a trial-and-error strategy (see Figure 6). The comparison between the pretest and posttest solution of the *quiz* problem also shows a clear gain. In the posttest she systematically tried a number of combinations of hits and misses (see Figure 6), while in the pretest she gave a wrong answer without providing any solution information.

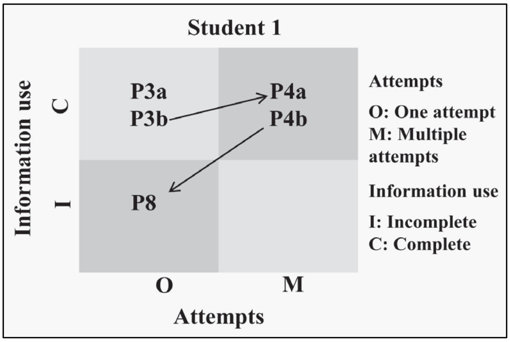


Fig. 3 Solution strategies student 1

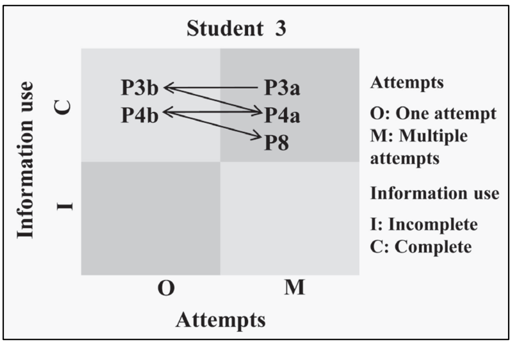


Fig. 4 Solution strategies student 3

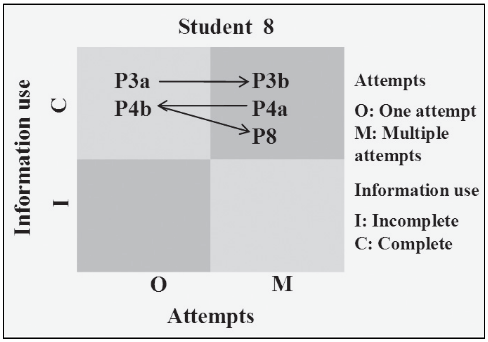


Fig. 5 Solution strategies student 8

Problem	Event	# Hits	# Misses	Rule Hit	Rule Miss	Points
P 4a	1	16	16	+1	–0	16
P 4b	1	16	16	+2	+1	48
	2	16	16	+2	–1	16
P 8	1	7	3	+2	–1	11
	2	5	5	+2	–1	5

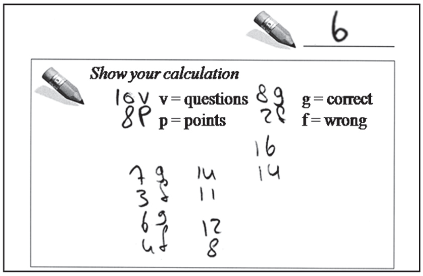


Fig. 6 Solution history in online Problems 4a, 4b, and 8 (left) and *Quiz* problem strategy (right) by student 8 in posttest

The work of student 3 is another example in which the influence of the online game is recognizable. In the posttest, this student used the model of the archery game to solve the *quiz* problem (see Figure 7): first, she stated the rule ‘correct: two points; wrong: –1 point’ and then she made three systematic attempts, the first with eight correct questions, the second with seven correct questions, and the last with six correct questions.

Problem	Event	# Hits	# Misses	Rule Hit	Rule Miss	Points
P 8	1	3	4	+2	–1	2
	2	3	6	+2	–1	0
	3	5	4	+2	–1	6
	4	5	5	+2	–1	5

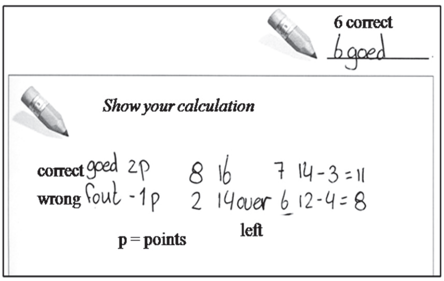


Fig. 7 Solution history in online Problem 8 (left) and Quiz problem strategy (right) by student 3 in posttest

To compare problem solving in situations with and without feedback, we present the results of all the students that worked on the online Problem 8 and who did the *Quiz* problem in the pre and posttest. These were 26 in total. Table 3 displays their results for the *Quiz* problem in the pretest and the posttest and in the online Problem 8.

Table 6
Solution categories in the *Quiz* problem (pretest and posttest) and in online Problem 8

Solution categories	Students %		
	Quiz problem pretest (*n* = 26)	P8 (*n* = 26)	*Quiz* problem posttest (*n*= 26)
Answer-C	46	88	69
Complete information use	46	88	69
Multiple attempts		19	
Trial-and-error	15	15	15
Systematic trialing	15	4	23
Verifying answer	50		58
Unknown	3		

As can be seen in Table 6, the students had more difficulties with the *Quiz* problem in the paper-and-pencil tests than with the similar online Problem 8. The percentages of correct answers – which actually correspond with the percentages of complete use of information – differ substantially. In the online Problem 8 almost all the students found the correct solution. Obviously, for some students this was a fruitful experience; the posttest result of the *quiz* problem was higher than the pretest result (46% compared to 69%).

Twenty-one students (81%) solved Problem 8 with one attempt, which in most of the cases (for 18 of the 21 students) meant that they shot five hits and five misses in the first attempt. Five students (19%) made multiple attempts and employed either trial-and-error or systematic trialing. In comparison to the pretest, the percentage of trial-and-error remained the same and the systematic trialing increased in the posttest.

Another observation that came to the fore when analyzing the solution processes of the students was that verifying does not straightforwardly lead to the correct answer. In case students do not use the complete problem information either when solving the problem or when verifying the answer, the verification procedure does not help students in detecting their error. For example, to solve online Problem 8, a student might try out three hits and one miss, leading to a score of five points, which is the intended amount of points. Then, the student might assume that this is the correct answer, while ignoring the total number of arrows, which are ten. However, this actually happened less often in the online environment than in the paper-and-pencil test: 12% compared to 54% (pretest) and 31% (posttest) of the students gave a wrong solution due to incomplete information use.

7 Conclusions and discussion

This study is about an online archery game played at home, which directly provided students with the results of their shooting actions. The aim of the study was to investigate the potential of this game to support primary school students in solving early algebra problems.

Based on our results we can claim that there is some evidence that the game-generated or situation-based feedback – as Nathan (1998) called it – contributed to students' ability in solving early algebra problems. Firstly, the students in our study performed better in the online environment than in the paper-and-pencil condition where no feedback was provided. Apparently, during game playing the students were confronted with all the parameters involved in the problem, which helped them to use all the problem information and encouraged them to detect and rectify their errors. Secondly, the students' online working had a positive effect on their performance in the paper-and-pencil posttest. This result is in agreement with Nathan's (1998) suggestion that situation-based feedback can foster higher-order thinking skills and internal learning control producing learning gains. These gains were observable in the increase in performance level between the pretest and the posttest. Moreover, we found some students enhanced their internal learning control. By playing the game, these students were inclined to verify their answers in the posttest. In fact, this means that the game-generated feedback might have changed the students' attitude and induced student-generated feedback. It is worth noting that a short online intervention that not only took place outside school, but which was also voluntary for the students, turned out to be sufficient to make this happen.

Of course, these findings should be treated with caution. To begin with, we could not control whether the students worked on the online problems by themselves or with the help of a peer or a family member. Nevertheless, because working online was voluntary and the tasks were low-stakes, we believe that the students did not feel the need for assistance. Actually, the diversity in the students' participation (not all students did all online problems) is also an indication that the students worked on their own. If parents had been involved it would have been more likely that the students would have accomplished all the presented problems.

More importantly, we carried out a limited experiment with a small number of students and we only measured a few student characteristics. As a matter of fact, besides the game itself, many other factors could have been of influence on the gain in performance. Therefore, more research is necessary to further explore the potential of computer games with inherent feedback in supporting students' problem solving competence in the domain of early algebra. On top of that, the monitoring software that we attached to the online environment

– and provided us with detailed information about students' problem solving processes – can be used to investigate new possibilities for teacher-generated feedback in connection with this game-generated feedback.

Notes

1. Our colleague Huub Nilwik programmed this game.
2. Our colleague Peter Boon developed the DME.

References

Booth, L. (1984). *Algebra: Children's Strategies and Errors*. Windsor: NFER-Nelson.

Bottino, R.M., Ferlino, L., Ott, M., & Tavella, M. (2007). Developing strategic and reasoning abilities with computer games at primary school level, *Computers & Education, 49*, 1272–1286.

Bransford, J.D., Brown, A.L., & Cocking, R. (2000). *How People Learn.* Washington: National Academy Press.

Butler, D.L., & Winne, P.H. (1995). Feedback and self-regulated learning: a theoretical synthesis. *Review of Educational Research*, *65*, 245–274.

Carraher, D.W., & Schliemann, A.D. (2007). Early Algebra and Algebraic Reasoning. In F. Lester (Ed.). *Second Handbook of Research on Mathematics Teaching and Learning* (Vol. 2, pp. 669–705). Charlotte. NC: Information Age Publishing.

Clements, D.H. (2000). From exercises and tasks to problems and projects: unique contributions of computers to innovative mathematics education. *Journal of Mathematical Behavior, 19*, 9–47.

Figueira-Sampaio, A.S., dos Santos, E.E.F., & Carrijo, G.A.C. (2009). A constructivist computational tool to assist in learning primary school mathematical equations. *Computers & Education*, *53*, 484–492.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*, 81–112.

Herscovics, N., & Linchevski, L. (1994). A cognitive gap between arithmetic and algebra, *Educational Studies in Mathematics*, *27*, 59–78.

Kaput, J. (2007). What is algebra? What is algebraic reasoning? In J. Kaput, D. Carraher, & M. Blanton (Eds.), *Algebra in the Early Grades* (pp. 5–18). New York: Lawrence Erlbaum.

Katz, V.J. (Ed.). (2007). *Algebra: Gateway to a technological future*. Washington, DC: Mathematical Association of America.

Kluger, A.N., & DeNisi, A. (1996). The effects of feedback on performance: a historical review, a meta-analysis, and a preliminary feedback intervention, *Psychological Bulletin, 119*, 254–84.

Kolovou, A., Van den Heuvel-Panhuizen, M., & Bakker, A. (2009). Non-routine problem solving tasks in primary school mathematics textbooks-A needle in a haystack, *Mediterranean Journal for Research in Mathematics Education, 8*(2), 31–67.

Lach, T., & Sakshaug, L. (2004). The role of playing games in developing algebraic reasoning, spatial sense and problem solving, *Focus on Learning Problems in Mathematics*, *26*(1), 34–42.

Lee, C.Y., & Chen, M.P. (2009). A computer game as a context for non-routine mathematical problem solving: the effects of type of question prompt and level of prior knowledge, *Computers & Education, 52*, 530–542.

Lee, H.S., & Hollebrands, K.F. (2006). Students' use of technological features while solving a mathematics problem. *Journal of Mathematical Behavior, 25*, 252–266.

Malisani, E., & Spagnolo, F. (2009). From arithmetical thought to algebraic thought: the role of the 'variable', *Educational Studies in Mathematics, 71*, 19–41.

Nathan, M.J. (1998). Knowledge and situational feedback in a learning environment for algebra story problem solving, *Interactive Learning Environments, 5*, 135–159.

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

PISA-NL-team (2006). *Wiskundige Geletterdheid Volgens PISA. 1. Analyse – Hoe Staat de Vlag Erbij?* [Mathematical Literacy According to PISA. 1. Analysis – What is the Situation?]. Utrecht, The Netherlands, Freudenthal Institute.

Schliemann, A.D., Carraher, D.W., Brizuela, B.M., Earnest, D., Goodrow, A., Lara-Roth, S., et al. (2003). Algebra in elementary school. In N. Pateman, B. Dougherty, & J. Zilliox (Eds.), *Proceedings of the 27th PME International Conference* (Vol. 4, pp. 127–134). Honolulu, HI: University of Hawaii.

Shute, V.J. (2008). Focus on formative feedback, *Review of Educational Research*, *78*, 153–189.

Smith, J.P., & Thompson, P.W. (2007). Quantitative reasoning and the development of algebraic reasoning. In J. Kaput, D. Carraher, M. Blanton, (Eds.), *Algebra in the Early Grades* (pp. 95–132). New York: Erlbaum.

Sutherland, R., & Rojano, T. (1993). A spreadsheet approach to solving algebra problems, *Journal of Mathematical Behavior, 12*, 353–383.

Van den Heuvel-Panhuizen, M., & Bodin-Baarends, C. (2004). All or nothing: Problem solving by high achievers in mathematics. *Journal of the Korea Society of Mathematical Education, 8*(3), 115-121.

Chapter 6

Can an intervention including an online game to be played at home contribute to students' early algebra performance?

Angeliki Kolovou [a], Marja van den Heuvel-Panhuizen [a], and Olaf Köller [b]

[a] *Freudenthal Institute for Science and Mathematics Education, Utrecht University, the Netherlands*

[b] *IPN, Leibniz Institute for Science and Mathematics Education, University of Kiel, Germany*

Kolovou A., Van den Heuvel-Panhuizen M., & Köller O. (submitted). Can an intervention including an online game to be played at home contribute to students' early algebra performance?

Can an intervention including an online game to be played at home contribute to students' early algebra performance?

1 Introduction

The great significance attached to algebra as an educational goal (see, e.g., Schoenfeld, 1995; Katz, 2007; Kilpatrick, Swafford, & Findell, 2001; NCTM, 2000) goes together with considerable concern about students' difficulties with this mathematical domain. A considerable number of students in secondary education show poor performance and drop out of algebra classes (Kaput, 2007). Among other things, researchers attribute the difficulties in learning algebra to the limitations of mathematics education in primary school (Schliemann et al., 2003). Actually, the problems with algebra make explicit that students have already misconceptions in arithmetic, which remain unnoticed until algebra is introduced (Booth, 1984; Schliemann, Carraher, Brizuela, & Jones, 1998). One of these misconceptions is that the equal sign is only interpreted as a command to carry out an operation (Booth, 1984; Knuth, Stephens, McNeil, & Alibali, 2006). Another source of difficulties for beginning algebra students is that in primary school the focus is on working with particular numbers and quantities rather than on reasoning about relations among sets of values (Schliemann et al., 1998).

While some researchers have attributed these obstacles to a cognitive gap between arithmetic and algebra (Herscovics & Linchevski, 1994), others have ascribed them to the late and abrupt introduction of school algebra (Kaput, 2007). In the primary grades the focus is on developing arithmetic and computational fluency, which in secondary school is followed by an isolated and superficial procedural approach to algebra (Kaput, 2007). To bring in more coherence in school mathematics, several researchers have supported the integration of algebraic reasoning into the primary school mathematics curriculum (Carraher, Schliemann, Brizuella & Enrnest, 2006; Kaput, 2007; Schoenfeld, 1995).

However, the suggested inclusion of algebra in primary school mathematics does not imply adding traditional algebra to the primary school curriculum; rather, it means providing entry points to algebra through treating existing number topics in a deeper and more connected way (Kaput, Carraher, & Blanton, 2007). These entry points can include, for instance, a collection of problems that offer students opportunities for reasoning algebraically by applying informal methods, i.e. solving problems without using letters to represent unknown values. This approach to teaching and learning mathematics in the primary school is mostly referred to as 'early algebra' (Katz, 2007). In fact, early algebra builds on students' proficiency in arithmetic and develops it further (Kilpatrick, Swafford, & Findell, 2001). Although grounded in arithmetic, students' informal methods can be the stepping stones to developing algebraic reasoning. For example, exploring numerical patterns can help students move towards understanding functional relations (NCTM, 2000). Rich problem contexts can play an indispensable role herein, as experience and reasoning in particular situations may support students in generating abstract knowledge (Carraher & Schliemann, 2007).

The present study investigates how primary school students' can be provided with help to develop algebraic reasoning. Our focus is on their ability to solve contextual number problems with covariating values. In contrast with secondary school students who have the option of applying formal algebra, which means setting up and solving equations with unknowns, primary school students may solve these problems through context-connected informal reasoning.

To meet this challenging goal of supporting early algebraic reasoning we use the benefits that new technologies offer. In this way, we seek to extend the work of other researchers who have successfully implemented paper-and-pencil activities for teaching algebra to young students, such as function tables (Schliemman, Carraher, & Brizuela, 2001) function machines (Warren, Cooper, & Lamb, 2006), pattern activities (Moss & Beatty, 2006), and solving word problems (Van Amerom, 2003). In particular, we want to explore the potential of computer games, as research has shown that they can offer engaging learning experiences to students, encourage active learning, and support the learning of complex subject matter (e.g., Ke, 2008). Therefore, we designed a dynamic computer game – meaning that the game has an interactive character, contains animations, and provides instant feedback – to support students in understanding covariation and functional relations. Our assumption is that such an environment can contribute to the development of thinking processes related to varying situations. According to Lew (2004), this dynamic thinking constitutes an essential part of algebra. A dynamic computer environment can prompt students to focus on the ways that output values vary as input values change, so that they can more easily identify relations between covariating values. This is a unique advantage of computer activities over paper-and-pencil for introducing young children to the concept of function.

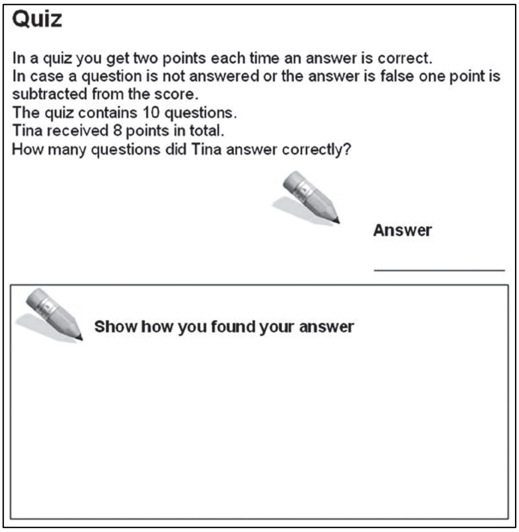
Another means that is often brought into action to remedy educational deficits is homework. Several studies have come up with evidence that homework can enhance achievement, but there are many factors that influence its effectiveness (e.g., Trautwein, 2007). Homework is usually employed to practice basic skills learned at school, whereas there is hardly any knowledge about the role of homework in learning complex competences, such as algebraic reasoning.

The present study aims at acquiring knowledge about innovative educational approaches to the teaching of algebra in the primary grades by investigating whether an online game played at home can function as a learning environment for eliciting early algebraic problem solving.

2 Theoretical background

2.1 Teaching and learning of early algebra

According to Carraher et al. (2006) early algebra involves a shift from working with particular numbers and measures towards working with relations among sets of numbers and measures. In fact, they see early algebra as departing from arithmetic reasoning with a focus on generalizing and reasoning about functions. To stimulate algebraic thinking and enable the move from arithmetic to algebra students should participate “in specially designed activities, so that they can begin to note, articulate, and represent the general patterns they see among variables” (Schliemann et al., 2003, p. 128). Despite the fact that secondary school students have difficulties with algebra, which might be caused by shortcomings in the primary mathematics curriculum, several experiments have shown that primary school students have the potential to handle fundamental algebraic concepts such as the notion of variable, function, and covariation (Carraher et al., 2006), can reason algebraically by recognizing patterns and making generalizations (Blanton & Kaput, 2005), and are capable of algebraic problem solving (Schliemann et al., 1998). Following these researchers, our approach to early algebra focuses on supporting primary school students to reason about relationships between numbers and quantities and to express general rules between sets of values. An example of a problem that can elicit such reasoning is the *Quiz*[1] problem.


Quiz

In a quiz you get two points each time an answer is correct.
In case a question is not answered or the answer is false one point is subtracted from the score.
The quiz contains 10 questions.
Tina received 8 points in total.
How many questions did Tina answer correctly?

Answer

Show how you found your answer

Fig. 1 *Quiz* problem.

The *Quiz* problem can be solved by a formal algebraic approach, which is solving through a system of two linear equations with two unknowns. However, this problem can also be solved by applying an arithmetic method. Then, students may reason informally about the relationships between the numbers in the problem. For example, they can choose a value to represent the number of correct answers, perform the calculations that are described in the problem, and check whether the guess leads to the intended result, i.e. the score of 8 points. In fact, before receiving instruction in algebra, students have only arithmetic approaches at their disposal. Although such approaches cannot be labeled as formal algebraic procedures, Johanning (2004) suggests that they can be seen as a way in which students make sense of algebraic situations.

However, it is important to be aware of possible gender differences in students' performances in this type of problems, since it is known that boys outperform girls in higher-order number skills. For example, Winkelmann, Van den Heuvel-Panhuizen, and Robitzsch (2008) found that boys attained higher scores than girls on items that include calculations with a missing term or with an unknown start and scored higher in mathematical reasoning, problem solving and modeling, and in problems which ask for backward reasoning. Results of gender studies with secondary school and college students are mixed. On the one hand, several studies have revealed that female students perform better than male students in algebra tasks. For example, in TIMSS (Mullis, Martin, & Foy, 2008) girls scored higher than boys in the content domain of algebra in general, and in the study by Ryan and Chiu (2001) this was the case in the subdomain of algebraic operations. On the other hand, research indicates that male students perform better in higher-order thinking algebra items (Ryan & Chiu, 2001; Mendes-Barnett & Ercikan, 2006). Furthermore, in the PISA 2003 study, boys generally outperformed girls in the content scale 'Change and relationships' (OECD, 2003).

2.2 Role of ICT in teaching and learning mathematics and algebra in particular

Several research studies have emphasized the role of the use of Information and Communication Technology (ICT) in the teaching and learning of mathematics. There is evidence that ICT has an overall positive effect on mathematics achievement. This is shown by Li and Ma (2010), who carried out a meta-analysis of 46 studies on the effect of computer technology on students' learning of mathematics. In general, the positive influence of ICT is accounted for by several factors. For example, digital technologies offer access to external representations and reactive feedback (Zbiek, Heid, Blume, & Dick, 2007) and support active engagement and interactive learning (Roschelle, Pea, Hoadley, Gordin, & Means, 2000).

Although Li and Ma (2010) did not find gender differences in the effect of ICT, other researchers have come to different conclusions. According to Imhof, Vollmeyer, and Beierlein (2007) male students are more successful than girls in computer applications. Moreover, several researchers have reported gender differences in computer attitudes. For example, boys are more self-confident with computers and score higher in pleasure and perceived relevance of computer use (Mumtaz, 2001). In a study by Oosterwegel, Littleton, and Light (2004), however, both boys and girls held positive attitudes towards computers.

In addition to the important role that ICT plays in all domains of mathematics, using ICT is considered to be significant in the teaching and learning of algebra (see, e.g., Stacey, Chick, & Kendal, 2004). Logo environments (Hoyles & Sutherland, 1989) and spreadsheets (Lannin, 2005; Sutherland & Rojano, 1993) have been successfully used for developing algebraic concepts. Furthermore, Cuoco (1995) showed that dynamic computer environments can support the development of increasingly sophisticated concepts of functions, for example from the action concept of function (i.e., a set of isolated calculations) to a process concept of function (i.e., a network of calculations). Resnick, Eisenberg, Berg, Mikhak, and Willow (2000) also suggested that technology tools might be more appropriate for developing concepts related to dynamic processes. Through this dynamic-oriented approach, as they call it, the concept of function might become accessible to primary school students. However, further research in the use of technology to enhance the learning of algebra is needed (Katz, 2007).

Among the ICT environments, computer games have attracted special interest from educators, researchers, and policymakers (McFarlane, Sparrowhawk, & Heald, 2002). On the basis of an extensive literature review Mitchell and Savill-Smith (2004) asserted that computer games have a significant impact on students' cognitive skills. They concluded that computer games are engaging and can embed mathematical concepts that may be hard to grasp with concrete materials. Sedighian and Sedighian (1996) found that computer-based mathematical game environments support meaningful learning and provide successful and challenging experiences. Furthermore, Squire (2003) – in his reflection on the history of games in educational research – refers to the state of flow that students can achieve when playing games (see also Hsu & Lu, 2004). In such situations, students engage in complex, goal directed activities not because there are external rewards but for the excitement of game playing. However, apart from a few studies that show positive effects on students' performance (Kebritchi, Hirumi, & Bai, 2010; Redfield, Gaither & Redfield, 2009), less is known about whether games can contribute to the learning of algebra.

2.3 Homework

Homework is defined as "tasks assigned to students by schoolteachers that are meant to be carried out during non-school hours" (Cooper, 1989, p.7). Although it is usually situated at the students' house, homework has also a school-based part where teachers explain, control and discuss it. Homework is recognized as a form of self-regulated learning (Trautwein, Lüdtke, Schnyder, & Niggli, 2006). In general, homework is considered to be a potentially powerful instructional device (Trautwein, Lüdtke, Kastens, & Köller, 2006). However, several variables influence the effect of homework on achievement. For example, effort put into homework, such as number of tasks completed and percentage of tasks attempted, has been found to be positively related to achievement, whereas time spent on homework and achievement have been found to be negatively related or unrelated (Trautwein, Lüdtke, Schnyder, et al., 2006). Furthermore, students' characteristics, such as cognitive abilities and gender, affect the relation between homework and achievement (Trautwein, Lüdtke, Schnyder, et al., 2006). Cognitive abilities predict homework expectancy, which is positively related to motivation and homework behavior (i.e., effort, time, and learning strategies). Regarding gender differences, girls show more compliance, which is positively related to homework effort (Trautwein, Lüdtke, Schnyder, et al., 2006) and spend more time than boys on homework (Trautwein, 2007). Furthermore, girls report a higher frequency of homework completion and find homework less boring than boys.

In comparison to regular paper-and-pencil homework and the commonly used computer programs that are mostly focused on routine skills, dynamic and interactive ICT environments allow the development of more complex skills such as problem solving (Clements, 2000). Moreover, since game playing is the most prevalent type of home computer activity (Moseley, Mearns, & Tse, 2001), computer games with a dynamic and interactive character might play an important role in broadening the scope of competences being practiced at home and raising the quality of homework. However, like in general computer use gender differences might also apply to home computer use and game playing in particular. According to Colley and Comber (2003) boys have a marked preference for playing games, which is associated with a greater out-of-school computer use by boys.

2.4 Focus of the study

The goal of this study was to investigate the effects of an intervention, in which students got the opportunity to work at home with an online game, on their performance in early algebra problems that involve covarying variables. Moreover, we wanted to know whether the effect on the performance differs between girls and boys. The present article addresses what we found with respect to the students ability to solve these early algebra problems as a result of the intervention; elsewhere (Kolovou, & Van den Heuvel-Panuizen, 2010) we described the nature of the early algebraic reasoning that was elicited while working in this online environment.

For the purpose of the study, an online ICT environment was developed that included an archery game, in which the students gain points by shooting arrows, and a tool for monitoring the students' working in this environment. The game was intended to offer students experience with covarying variables; the number of arrows, the points assigned to each shot, and the total score are interrelated so that a change in one of these values affects the other values as well. The students could use the game to solve problems at their own pace as part of their homework. In addition, the students received classroom instruction to get introduced to the online environment and, later, to share their experiences with the game playing.

The monitoring tool enabled the registration of the online activity and the logged-in time of every student. The frequency of the online activity was determined by the number of events (shooting actions) carried out in the online environment. Other indicators of effort that were used were the number of problems that the students worked on and the number and percentage of focused events (i.e., shooting actions intended to answer a given problem).

3 Research questions and hypotheses

In the light of the above, we addressed three research questions. The *first research question* was whether an intervention including an online game enhanced the students' performance in early algebra. Because prior studies have shown positive effects of computer technologies on mathematics learning, we expected that an intervention program offering students the opportunity to play a computer game would contribute positively to the students' achieved performance in early algebra (Hypothesis 1).

The *second research question* zoomed in on the effect of the effort indicators of the online work (the time being logged in, number of carried out events, number and percentage of focused events, number of worked problems) on the achieved performance in early algebra problem solving. With respect to the efforts indicators our expectations were varying. Concerning the influence of login time we were torn between two assumptions. On the one hand, a larger amount of time invested on the online work might bring about a stronger positive effect. On the other hand, the relation between logged-in time and performance might be negative; the more able students might need to spend less time in the online environment. Because of this contradiction, no hypotheses were made for the logged-in time. Regarding (a) the number of carried out events, (b) the number of focused events, (c) the percentage of focused events, and (d) the number of worked problems we expected them to be positively related with the achieved problem solving performance in early algebra (Hypotheses 2a, b, c, and d).

The *third research question* addressed the previous two questions from the gender perspective. To begin with, we investigated whether boys and girls differed in their achieved performance in early algebra as a result of the intervention. Research has shown that, on the one hand, boys use the computers more successfully and might, therefore, benefit more than girls from the intervention. Moreover, several studies have suggested that boys outperform girls in more complex problem solving tasks. On the other hand, girls are more willing to accomplish homework assignments, which is positively associated with performance. Because of these somewhat opposing findings, we did not make predictions about the role of gender on the effect of the intervention on the early algebra performance.

Additionally, we investigated whether gender was related to effort indicators. According to previous research outcomes, girls might spend less time than boys on computer activities, but they are also inclined to spend more time on their homework tasks. Because of these conflicting empirical results, we did not formulate a hypothesis about differences in logged in time. Furthermore, in accordance with empirical findings that girls show more effort in homework, we expected that girls (a) carried out more events, (b) performed more focused events, (c) carried out a higher percentage of focused events, and (d) accomplished a larger number of worked problems in the online environment (Hypotheses 3a, b, c, and d).

4 Method

4.1 Design – Procedure

A pretest-posttest-control-group experiment was set up to investigate the influence of the intervention on students' early algebra performance. Compared to earlier research, this study introduces two innovative elements: Firstly, working with the computer game did not take place in school but at home (outsourcing education) and, secondly, special software was used to collect data about the students' online computer activity. This software provides access to the students' performance from a distance (tele-assessment).

In total the duration of the experiment was about six weeks. The intervention consisted of three periods of one week in which students could play the online game. Each period started with a class instruction lasting 30–45 minutes. During the first class instruction one of the authors introduced the online environment and the game features with the help of an interactive whiteboard or a digital projector. Each student received a unique internet account and a worksheet with a set of problems and the request to visit the online environment at home and solve these problems. In the second and third class instruction, the solutions of the problems were discussed for about 15 minutes and a new set of problems was provided as homework to the students. During the discussions the students presented their answers and emphasis was put on the underlying structure of the problems. Apart from discussing the students' solutions the class instruction was also meant to encourage students to keep up with the online work at home.

The online activity was introduced as a game and was not part of the students' compulsory homework. As a matter of fact, the students were asked to help researchers collect information about how primary school students utilize the game. The teachers had a minor role in the intervention; they only had to stimulate their students to go online and play the game. At the beginning of the experiment both the students and their parents were informed about the registration of the online work.

Before and after the intervention a pretest and a posttest were administered by the researcher who gave the class instruction. The testing was not only done in the experimental group but also in the control group. However, the students in the control group did not receive the intervention.

4.2 Sample

Our sample consisted of 236 sixth-graders. In the sampling process considerable attention was paid to recruiting students from various socio-economic backgrounds. Due to the rather noticeable phenomenon of ethnic segregation in Dutch primary schools, which can be partly attributed to segregated city districts, we contacted schools from different districts in a large city in the Netherlands. Initially, we approached schools from five city districts that cover a wide socio-economic range in student population. Each school was contacted by email and phone and the sixth-grade class was invited to participate in the intervention. After we had found five schools that agreed to be involved in the study, we looked for five matching control classes within the same five districts. Through this procedure we found students from in total ten schools to take part in the study. All participating students were experiencing a similar type of mathematics education. Their schools were using comparable mathematics textbooks which only contain a minimum of early algebra type of problems (Kolovou, Van den Heuvel-Panhuizen, & Bakker, 2009).

In general, Dutch students are quite familiar with computer technology. In fact, 96% of the primary school students in Netherlands have internet access at home (Mullis et al., 2008). Particularly, computer games are very popular among them. This is also true for the students of this study. According to their teachers, the students of the participating classes were familiar with playing educational computer games at school and at home. Therefore, we assumed that a novel effect of technology (Li & Ma, 2010) would not be likely for the students of this study.

4.3 Instruments

4.3.1 The online environment

The environment that was developed to give students experience in dealing with covarying variables includes a dynamic game called 'Hit the target'[2]. The default screen of the game displays five features: The target, the pile of arrows and the bow, the score board, the board that displays the game rule, and the board that shows the number of arrows (hits and misses) (see Figure 2). In this game the students can set the shooting mode and the game rule mode. The shooting mode includes two options: Shooting arrows one by one by dragging them to the bow (user shooting) or entering a number of hits and misses and shooting them at once (computer shooting). The game rule mode also includes two options: The students can set the game rule by filling in the points added or subtracted in case of a hit or a miss (user defined) or the computer sets the rule randomly (computer defined). The features of the game are dynamically linked. During the shooting the values on the scoreboard update rapidly to inform students about their score. The same happens when removing arrows from the target. In this way, students may comprehend that the arrows, the score and the game rule are related to each other so that a modification in the value of one of these variables has a direct effect on the other variables. Moreover, the game offers instant visual feedback by showing on the score board the number of hits and misses and the number of points resulting from a shooting action.

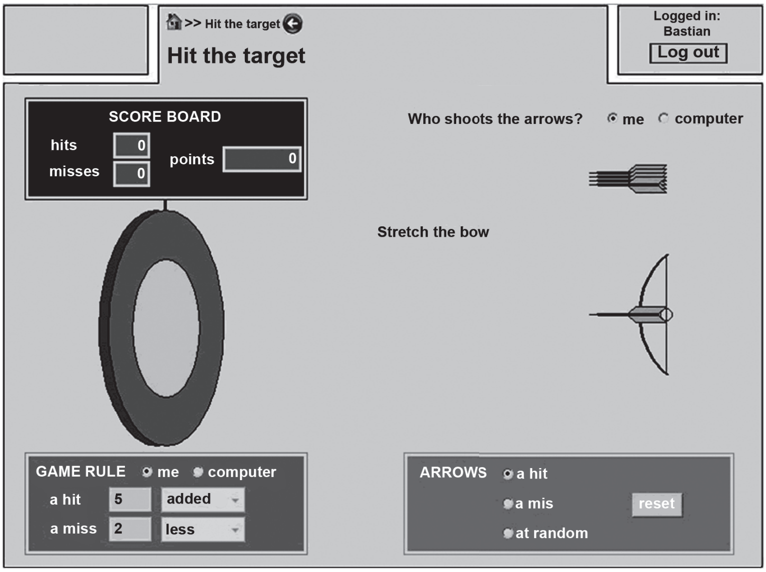


Fig. 2 Screen view of game in the computer shooting-mode.

For working in the online environment each student of the experimental group received a unique account and eight problems (six of them consisting of two parts) split in three sets meant for the three periods of the intervention. The problems varied from finding the number of arrows that were hits and the number of arrows that were misses that produce together a particular score, to generating a general solution by systematizing all possible answers. Figure 3 shows a selection of the problems that were given to the students in the three periods of the intervention.

In period 1
Problem 1: For every hit you gain 3 points and for every miss 1 point is taken away from your score. How many hits and misses do you have to shoot to get 15 points in total?
Problem 3a: What is the game rule to get 15 points in total with 15 hits and 15 misses?
Problem 3b: Are there other game rules to get 15 hits, 15 misses, and 15 points?

In period 2
Problem 4a: What is the game rule to get 16 points in total with 16 hits and 16 misses?
Problem 4b: Are there other game rules to get 16 hits, 16 misses, and 16 points?
Problem 5a: What is the game rule to get 100 points in total with 100 hits and 100 misses?
Problem 5b: Are there other game rules to get 100 hits, 100 misses, and 100 points? Can you explain your answer?

In period 3
Problem 8a: For every hit you gain 2 points and for every miss 1 point is taken away from your score. You have 10 arrows in total. How many hits and misses do you have to shoot to get 5 points in total?
Problem 8b: Are there any other solutions possible?

Fig. 3 Selection of problems for working in the online environment

In this series two types of problems are included: Problems that can be translated into a linear equation of the type $ax + by = c$, thus having an infinite number of solutions, and problems that can be translated into a system of two linear equations with two unknowns with only one solution: $ax + by = c$ and $x + y = d$. However, both types of problems are presented in such a way that students can solve them at a level that fits their understanding and in an environment that offers them the opportunity to use the game for checking their answers. The idea behind this series of problems is that they support students in detecting and generalizing relationships between sets of numbers. Moreover, these problems confront students with situations in which multiple conditions should be taken into account.

In *Problem 1* the students have to find the number of hits and the number of misses that result in a total score of 15 points when 3 points are added for a hit and 1 point is subtracted for a miss. Several solutions are correct as long as they result in 15 points.

In *Problem 3a*, instead of looking for the number of hits and misses, students have to find a game rule that produces the given result (15 hits; 15 misses; 15 points). The students can solve this problem by just guessing-and-checking, or by systematically trying out solutions, or by finding the solution directly by reasoning.

In *Problem 3b* the students have to come up with other game rules that lead to the same result as in *Problem 3a*. By finding other game rules, students might discover a general solution that describes how the points for a hit and the points for a miss are related to get a total number of points that equals the number of hits and the number of misses. They can realize that the points added for a hit and the points subtracted for a miss should result together in one point. This means that every time you shoot one hit and one miss you also get one point. This general solution can be used to find game rules for any situation in which the numbers of hits, misses, and points are equal (n hits; n misses; n points). For students who arrive at this general solution, *Problem 4a* and *Problem 4b* are easy to solve. Those who do not discover the relationship can still solve the problems by guessing-and-checking. Hereafter, *Problem 5a* and *Problem 5b* bring in large numbers, which help students realize that the general solution applies regardless of the size of the numbers. In this way, students are induced to think about the relationships between the numbers rather than just doing operations with the numbers involved in the problem.

A new step towards algebraic problems is made in *Problem 8*. Actually this problem is similar to *Problem 1* but it introduces a new condition. In addition to the game rule and the total score, the sum of the hits and misses is known as well. Mathematically this means that this problem cannot be translated in one equation with an infinite number of solutions, but that it needs to be translated into a system of two linear equations which has only one solution. For primary school students who do not have algebraic procedures available, solving *Problem 8* means that they have to take into account all the relevant information and make sure that not only the total score is 5, but also that the total number of hits and misses is 10. If, for instance, a student comes up with the solution 6 hits and 7 misses, then, on the one hand, the total score is 5 points, but, on the other hand, the total number of hits and misses is 12 and not 10. In that case, the student's solution does not satisfy all the criteria of the problem.

The students were asked to solve these problems while playing the game and to write down their answers on a worksheet, which was meant for the students' use only. During the class instruction the students' answers were discussed and emphasis was placed on determining rules and regularities. To keep track of the students' steps while working online, the game was connected to the so-called Digital Mathematics Environment (DME)[3]. This software creates a log that traces students' actions in the game. The log data consist of a list with all the actions carried out by each student, divided into sessions (i.e., every time that the student logged in the online environment) and events (i.e., every time that the student clicked on the shoot button).

4.3.2 Tests for assessing students' mathematical abilities

In this study we used two tests to measure students' mathematical abilities: A test for assessing ability in solving early algebra problems (which was used as a pretest and a posttest) and a test for assessing general mathematical ability (which was used as a pretest only). For early algebra problem solving we developed a paper-and-pencil test containing seven contextual number problems with interrelated values.

Students' responses to the problems were coded according to the correctness of the answer. For each student the proportion of correct items was computed as an individual raw test score. The pretest results indicated that, despite the low number of items, the internal consistency was satisfying (Cronbach's alpha = .79). This score was found after excluding one problem from the test due to poor psychometric properties ($M = .17$, item-total-correlation $r = .37$).

The general mathematical ability of the students was measured by the End Grade 5 CITO-LOVS Mathematics Test (CITO E5 test). The CITO-LOVS Mathematics Test is a series of standardized tests meant for monitoring students' mathematics performance in Dutch primary school and assesses students' knowledge and understanding of number-related and measurement concepts. The items were scaled through an IRT model and the reliability coefficient of the CITO E5 test (version 2002) was .90 (Janssen & Engelen, 2002). The correlation of the pretest with the CITO E5 test was significant with a moderate strength of $r = .65$ ($p < .001$) suggesting that the standardized test for general mathematical ability and the test that assesses problem solving in early algebra to some extent measure different abilities.

4.4 Statistical issues

4.4.1 Hierarchical Data

In the present study students came from 10 classes. In the literature on multilevel analyses (e.g., Hox, 2002) it is argued that due to strong intra-class correlations of achievement measures (i.e, the average correlation between variables measured on pupils coming from the same class) the effective sample size of class-based samples is much lower than the number of real participants in the study. This typically leads to an underestimation of standard errors and consequently to an inflation of type-II errors (rejection of true null-hypotheses). Thus, a researcher is likely to come up with spurious significant results. Multi-level analyses, however, take the nested character of the data into account and provide unbiased standard errors. The intra-class correlations of our math tests were above .20 so students shared considerable amount of variance in their achievement scores. Therefore, we decided to conduct analyses with the HLM 6.0 package (Raudenbush, Bryk, & Congdon, 2004).

4.4.2 Missing Data

Missing data is a common problem in repeated measurement studies. One recommended solution to handle this problem is multiple imputation (e.g., Graham, 2009; Schafer & Graham, 2002). According to Schafer and Graham (2002) in multiple imputation each missing value is replaced by a list of $m > 1$ simulated values producing m plausible alternative versions of the complete data. Each of the m data sets is analyzed in the same way and the results, which may vary, are combined by simple arithmetic to obtain overall estimates and standard errors that reflect missing-data uncertainty as well as sample variation. Out of the 236 students involved in the study, 11 were absent during the pretest and 25 other students were absent during the posttest. Moreover, the CITO score of 17 students was not provided by their schools and 16 students took a different version of the CITO test, which was not comparable with the version that the rest of the students had taken. To test for patterns in missing data, we conducted a missing value analysis. Little's *MCAR test* (1998) clearly indicated that data were missing completely at random ($\chi^2 = 8.24$, $p = .41$). Therefore, we decided to carry out data imputation by applying the SPSS multiple imputation module to create $m = 10$ complete data sets. This is in agreement with the recent recommendations of Graham, Olchowski and Gilreath (2007). All information available was used to obtain a good imputation model. All subsequent analyses were then conducted 10 times and the results were combined in accordance with Rubin (1987).

5 Results

5.1 Initial differences between the groups

Before imputing missing values, we examined whether there were differences between the experimental and control group in age, gender composition, general mathematical ability and pretest performance. With respect to age and gender both groups were rather similar (see Table 1). Also, for the mean pretest scores of the two groups on the mathematical problem solving test there was not a significant difference ($t = -1.19$, *n.s.*). However, this was not the case for the students' general mathematical ability as measured by the CITO E5 test. Here we found a significant difference ($t = -3.36$, $p = .001$). Therefore, we decided to include the CITO E5 scores in our analyses.

Table 1
Descriptive statistics for experimental and control group before intervention

Group	N	Age		Gender		Pretest score		CITO E5 test		
		M	SD	Boys	Girls	M	SD	N	M	SD
Experimental	119	11.82	.53	61	58	.37	.35	115	115.8	10.4
Control	106	11.87	.52	48	58	.32	.30	88	111.1	9.3
Total	225	11.85	.52	109	116	.34	.33	203*	113.8	10.2

*These students are a subset of the 225 students with complete data in the pretest

The mean pretest scores of the boys and girls on the problem solving test in early algebra (see Table 2) differed significantly ($t = -4.20$, $p < .001$) in favor of boys. The same was true regarding their general mathematical ability as measured by the CITO E5 test ($t = -3.23$, $p < .01$).

Table 2
Descriptive statistics for boys and girls before intervention

Gender	*N*	Age		Pretest score		CITO E5 test		
		M	*SD*	*M*	*SD*	*N*	*M*	*SD*
Boys	109	11.90	.57	.43	.34	111	116.2	10.9
Girls	116	11.79	.48	.26	.29	92	111.7	9.1
Total	225	11.85	.52	.34	.33	203*	113.8	10.2

*These students are a subset of the 225 students with complete data in the pretest

5.2 Effect of the intervention

5.2.1 General effect of the intervention

To answer the first research question we compared the students' performance in the experimental versus the control group. The experimental group received an intervention consisting of class instruction in which they were requested to play the online game at home. However, because playing the game was voluntary not all students went online. Nevertheless, we considered the students who did not go online to be part of the experimental group. The main goal of the study was to figure out whether the intervention contributed to students' performance in early algebra problems. The fact that some students did not go online is an inherent feature of the intervention. This situation may also occur when teachers carry out this intervention in their educational practice in the future. In fact, the research question we aim to answer here is whether such an intervention, in any degree of implementation, has an effect or not.

Table 3 displays the pre- and posttest scores (average scores form 10 imputed data sets) in the control and experimental group, which is further split into students who only followed the class instruction and did not log in (i.e., the only-class-instruction group) and students who followed the class instruction and did log in (i.e., the logged-in group). No change over time occurred in the control group, whereas students in the experimental group showed significant increase with medium effect size (d = .57) in their achievement.

Table 3
Means and standard deviations of pretest and posttest scores for experimental and control group

Group	*N*	Pretest scores		Posttest scores	
		M	*SD*	*M*	*SD*
Experimental (*E1+E2*)	123	.35	.35	.48	.37
E1: Only-class-instruction	27	.27	.30	.40	.37
E2: Logged-in	96	.38	.36	.51	.38
Control	113	.31	.30	.33	.34

Multi-level regression analyses including pretest scores, the CITO E5 scores, gender, and group (0 = control group, 1 = experimental group) revealed that the effect of the intervention was indeed significant (*B* = .086, *p* = .016 see Table 4), i.e., in the posttest students in the experimental group outperformed their peers in the control group. Note, that the intervention effect is significant even when we include the pretest scores and the CITO E5 test as predictors, which results in a conservative estimation of the effect. Furthermore, there is a tendency for boys gaining more than girls from the intervention even after controlling for CITO E5 and pretest score. However, an additional analysis revealed that there was no significant Gender *x* Group interaction effect (*B* = –.004, *p* = .981). This finding clearly suggests that the positive main effect of the intervention was the same for girls and boys.

Table 4
Multi-level regression model predicting posttest scores by CITO E5, gender, pretest, and group (experimental vs. control): Unstandardized regression coefficients, SE, and their p-values from HLM

	B	*SE*	*p*
CITO E5	.012	.002	< .001
Gender	.044	.024	.072
Pretest score	.542	.047	< .001
Group	.086	.036	.016
R^2 (explained variance)	.63		

5.2.2 Specific effect of the intervention

To investigate whether the effect of the intervention (see Table 4) remained stable for the 96 students (78%) in the experimental group who followed the instruction given in class and were logged in at least once (i.e., the logged-in group), we compared the posttest performances of the logged-in group and the control group. Since the number of students per class became somewhat small, we decided to run here ordinary regression analyses instead of multi-level analyses. Multiple regression analyses including the CITO E5 and pretest score revealed that the students in the logged-in group, similar to the students in the complete experimental group, performed significantly better than the students in the control group (*B* = −.084, *p* = .017).

To further disentangle the effect of the online working from the effect of the class instruction we compared the performance of the students who followed the class instruction and were logged in and the students who only followed the class instruction (27 students). By comparing their pretest performance we found that the students who went online and followed the class instruction did not differ significantly from the students that were only present in the classroom sessions ($t = -1.50$, *n.s.*). With respect to the posttest performance, multiple regression analyses including CITO E5, pretest score, and group (0 = *only-class-instruction group*, 1 = *logged-in group*) revealed that the students who went online and followed the class instruction did not significantly differ from the students who only followed the class instruction ($B = .026$, *n.s.*).

5.2.3 Characteristics of online work as predictor of the achieved performance

To investigate the influence of the way of online working on the problem solving performance in early algebra we focused on the group of students who went online and investigated the role of logged-in time, number of carried out events, number of focused events, percentage of focused events, and number of worked problems. A correlation analysis showed that the characteristics of the online work were strongly related (see Table 5). The lower correlations of logged-in time with the other online characteristics might indicate that students were not actively engaged in the online activity for part of the time being logged in.

Table 5
Correlations among indicators of effort

	Characteristics	*1*	*2*	*3*	*4*	*5*
1	Time	–	$.49^{**}$	$.44^{**}$	$.13^{**}$	$.22^{*}$
2	Number of events		–	$.88^{**}$	$.40^{**}$	$.73^{**}$
3	Number of focused events			–	$.68^{**}$	$.87^{**}$
4	Percentage of focused events				–	$.77^{**}$
5	Number of worked problems					–

$^{**}p < .01$, $^{*}p < .05$

The tolerance values (.65 to .095) and the variance inflation factors (VIF) (1.54 to 10.51) of these characteristics revealed that multicollinearity would have been problematic in further data analysis (see Miles & Shevlin, 2001). Therefore, we performed a principal component analysis (PCA) in order to group the variables into a smaller number of uncorrelated factors. Only the first principal component displayed an Eigenvalue greater than 1 (Kaiser, 1960), accounting for 67% of the total variance. The factor loadings of the variables were all above .5. Consequently, all characteristics were collapsed into a single factor that we consider as an indicator of effort. This composite variable was used in our subsequent analyses.

To answer the research question about the influence of the online work, we conducted a regression analysis on the posttest scores with CITO E5 score, pretest score, and effort as predictor variables for the group of the students who went online (see Table 7). However, the effect of effort in online work was not significant (*B* = .039, *n.s.*).

5.2.4 Gender-specific effect of the intervention

With respect to gender, regression analyses revealed, that when controlling for the CITO E5 score and the pretest score, boys did show slightly higher problem solving performance in early algebra than girls (see Table 4). In addition to investigating their performance, we also looked into how the logged-in girls and boys used the online environment. The comparison of the means of the effort indicators revealed a number of significant differences (see Table 6).

In general, the girls put more effort into the online work than the boys did. Compared to the boys, the girls carried out a significantly higher number of events and a significantly higher number of focused events. However, we did not find significant gender differences with respect to logged in time, percentage of focused events, and number of worked problems.

Table 6
Means, standard deviations, results of t-tests, and effect sizes (d) of indicators of effort by boys and girls in the group of the logged in students

	M	*SD*	*t*	*p*	*d*
Logged in time (min)					
Boys	22.16	29.95	1.81	.075	.36
Girls	41.48	68.87			
Number of events					
Boys	10.52	6.45	2.84	.003	.57
Girls	15.94	11.72			
Number of focused events					
Boys	5.48	5.08	2.44	.009	.49
Girls	9.32	9.77			
Percentage of focused events					
Boys	0.45	0.33	.43	.335	.09
Girls	0.48	0.35			
Number of worked problems					
Boys	4.02	3.45	1.06	.148	.22
Girls	4.86	4.29			

To answer the research question about the gender-specific influence of the effort in online work we conducted a regression analysis on the posttest scores with CITO E5, pretest score, effort, gender and the interaction of effort and gender as predicting variables (see Table 7). Due to the relatively small sample size, we again ignored the nested character of the data.

Table 7
Regression models predicting posttest scores of students who were logged in (N = 96) by CITO E5, gender, pretest, effort, and interaction of effort and gender: Unstandardized regression coefficients, SE, and their p-values

	B	*SE*	*p*
CITO E5	.013	.003	< .001
Gender	.095	.049	.099
Pretest score	.538	.088	< .001
Effort	.039	.041	.339
Effort *x* Gender	–.043	.055	.436
R^2(explained variance)	.71		

The results of this regression analysis indicate that, after controlling for effort, the slight effect of gender on the problem solving performance remained stable. Effort and the interaction between effort and gender were not significant.

6 Discussion

This study investigated the effect of an intervention including an online game on primary school students' early algebra performance and the gender differences related to this effect. In the analysis of our data special attention was paid to the role of effort. After summarizing and discussing our findings we bring up educational implications and conclude our paper with addressing the limitations of our study and suggestions for further research.

6.1 The effect of the intervention

As predicted (Hypothesis 1), students benefitted from the intervention. In the experimental group, posttest scores were significantly higher (effect size d = .54) than pretest scores, with no change observed in the control group. Although this result agrees with previous findings (Roschelle et al., 2000; Mitchell & Savill-Smith, 2004), the outcome of our study, based on a short and voluntary intervention, is nevertheless noteworthy. Similarly, controlling for differences in the pretest and the general mathematics achievement, the performance of the students who were logged in and followed the class instruction was significantly higher than the performance of the control group students.

However, the performance of the logged-in students and the students who only attended the class instruction did not differ significantly. This last observation raises two issues. First, the number of students who did not log in was very low, which might imply that this result is not very reliable. Nevertheless, the direction of the effect reveals a small advantage – although not significant – for the students who were logged in and followed the class instruction. With a larger sample we could perhaps detect a significant difference in the posttest scores for the students who were logged in. Second, although it is intuitively appealing to believe that an intervention is effective only when it is implemented exactly as intended by the developers, the relationship between treatment integrity (i.e., the degree to which an intervention is implemented as intended) and treatment outcomes is not that straightforward (Sanetti & Kratochwill, 2009). In general, higher levels of treatment integrity result in better outcomes, but at the same time low levels of treatment integrity may have negative as well as neutral or even positive effects on treatment outcomes (Sanetti & Kratochwill, 2009). Our study was intended to be carried out partially in the class and partially at home. Therefore, the complete intervention included both parts, but since the at-home part was voluntary, one should expect deviations from the intended treatment. In addition, in a real setting there might be factors that influence the outcome in a way that is not anticipated. For example, in the present study some students might have talked with each other about the game or the tasks, which also influenced their posttest performance.

6.2 The effect of effort

Compared to the self-reported questionnaires that are usually used in homework studies, our study was based on more precise data about students' activity at home. The computer-generated log files enabled an in-depth analysis of students' online work. This resulted in an unexpected outcome.

One might expect effort – measured as a composite variable consisting of login time, number of events (Hypothesis 2a), number of focused events (Hypothesis 2b), percentage of focused events (Hypothesis 2c), and number of worked problems (Hypothesis 2d) – as the most important variable determining performance gain. Yet, this was not the case. Effort did not predict students' posttest performance. A possible explanation for this result is that students' benefit from the game may lie in an *aha*-experience occurring after a short period of playing the game. Probably, it was not the large amount of effort spent online that was important, but the fact that any effort was spent. In other words, gain might be more related to experience with the game as such than to repetitive experience with it. This observation is in agreement with what Durlak and DuPre (2008) suggested; there may be a critical threshold of intervention implementation, above which increased implementation does not meaningfully influence outcomes.

Moreover, we should also keep in mind that, in order to solve the problems, students may not exclusively rely on the game but they may solve a problem partly in their head.

Therefore, the effort of the online work might not completely capture the students' cognitive effort that is related to the gain in performance.

6.3 Role of gender

When testing the effect of gender, we found that boys and girls both benefitted from being in the experimental group. However, existing pretest differences in favor of boys became somewhat larger over time both in the experimental and control group.

The finding that there were already gender differences in the pretest scores matches prior research (Winkelmann et al., 2008) showing boys are more able at this type of problem solving. In the same line is the small gender effect on change in achievement suggesting, that independent of the group the differences in favor of boys increased over time.

In agreement with studies showing that girls are generally more compliant than boys in doing school assignments at home (Trautwein, Lüdtke, Schnyder et al., 2006), we found in our study that girls outperformed boys in two effort indicators. In particular, we observed significant differences regarding number of events (Hypothesis 3a) and number of focused events (Hypothesis 3b). Percentage of focused events (Hypothesis 3c) and number of worked problems (Hypothesis 3d) did not differ between boys and girls. However, we also found that girls' achievement gains in the treatment group were the same as for boys. This finding indicates the complexity of the factor effort.

An explanation for boys' higher scores despite their relatively lower effort could be that boys did not need that much experience with the online game. Boys tend to use computer applications more effectively than girls (Imhof et al., 2007) and are more frequent users of computer games (Colley & Comber, 2003). Probably, boys stopped using the game earlier than girls, as soon as they understood its aim. Furthermore, boys might have profited more from the game because they were more intellectually challenged by it. Boys may be more active users of games because they see them as something they can conquer.

6.4 Educational implications

The findings of our study indicate that computers are not only suitable for practicing basic skills, but can also be used as a tool to improve students' performance on early algebra problems. Moreover, *outsourcing education* by giving students the opportunity to play an online game at home seems to be a promising idea. Given appropriate tasks, home computing may create an effective learning environment supporting and extending school learning.

The results of the study also suggest that a voluntary homework assignment can contribute to positive effects on learning. Therefore, teachers can be more comfortable with letting students take responsibility for completing homework tasks. However, teachers should be aware of possible disparities between boys and girls in the educational benefits drawn from computer use as implied by the different effects of effort for boys and girls in our study.

6.5 Limitations of the study and suggestions for further research

Some limitations should be kept in mind when interpreting the results of the study. The effect of the intervention including an online environment was examined with only one game and the operationalization of early algebra focused on one type of problems. Furthermore, early algebra competency was measured by a test consisting of a limited number of items. Another shortcoming might be that the items were too difficult for the students, which might have obscured the effects of the intervention. However, making the items easier would not have been appropriate, because lowering the cognitive demand of the items would undermine their algebraic character. Moreover, the low performances of this group of Dutch students are in agreement with results from a Dutch study on non-routine problem solving with high-achieving fourth-grade students (Elia, Van den Heuvel, & Kolovou, 2009) and the finding in TIMSS 2007, that the Dutch students' performance in an early algebra task was below the international benchmark (Mullis et al., 2008).

Additionally, the duration of the intervention was quite short. A longer intervention might yield stronger effects and follow-ups might be suitable for investigating the long-term effectiveness of the treatment. Although the significant results of this short intervention indicate the treatment's power, an alternative explanation might be that the novel character of the intervention, which included a computer game to be played at home, influenced the results (Li & Ma, 2010).

Another reason for caution is that our sample selection was not optimal; we used a matching procedure on the basis of a number of schools that consented to participate in the study. Furthermore, the voluntary character of participation might imply that the students who used the online game were more motivated, which might have played a role in the effectiveness of the intervention. Yet, the results did not reveal a difference in performance for the logged-in students.

Overall, the promising results of our study encourage us to continue this line of research and further pursue the development of early algebraic reasoning through computer games. Especially, the potential of voluntary training programs at home to support the learning of mathematics should be further investigated. Notwithstanding the potential of voluntary participation, the effect of compulsory participation in class should be explored as well. In fact, students might be able to gain more from the intervention if the online activity is done by all students at school under the teacher's supervision. Further investigations might also shed light on the intriguing finding that boys tend to profit more than girls despite the fact that they do not invest as much effort as girls.

Notes

1. This paper-and-pencil problem was developed for the purposes of the World Class Tests. In 2004, Peter Pool and John Trelfall (Assessment and Evaluation Unit, School of Education, University of Leeds) asked us to pilot a series of items in the Netherlands.
2. The game 'Hit the target' was developed by the second author and programmed by our colleague Huub Nilwik at the Freudenthal Institute of Utrecht University.
3. The Digital Mathematics Environment (DME) was developed by our colleague Peter Boon at the Freudenthal Institute of Utrecht University.

Acknowledgments

The authors would like to thank Hendrik Winkelmann for his contribution in this study and Anne Teppo for her insightful comments on the draft of this article.

References

Blanton, M., & Kaput, J. (2005). Characterizing a classroom practice that promotes algebraic reasoning. *Journal for Research in Mathematics Education*, *36*, 412–446.

Booth, L. (1984). *Algebra: Children's Strategies and Errors*. Windsor, UK: NFER-Nelson.

Carraher, D.W., & Schliemann, A.D. (2007). Early Algebra and Algebraic Reasoning. In F. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (Vol. 2, pp. 669–705). Charlotte, NC: Information Age Publishing.

Carraher, D.W., Schlieman, A.D., Brizuella, B.M., & Enrnest, D. (2006). Arithmetic and algebra in early mathematics education. *Journal for Research in Mathematics Education, 37*, 87–115.

Clements, D.H. (2000). From exercises and tasks to problems and projects Unique contributions of computers to innovative mathematics education. *Journal of Mathematical Behavior, 19*, 9–47.

Colley, A., & Comber, C. (2003) Age and gender differences in computer use and attitudes among secondary school students: what has changed? *Educational Research, 45*, 155–165.

Cooper, H. (1989). *Homework.* White Plains, NY: Longman.

Cuoco, A. (1995). Computational media to support the learning and use of functions. In A. diSessa, C. Hoyles, & R. Noss (Eds.), *Computers and exploratory learning* (pp. 79–108). Berlin: Springer.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, *41*, 327–350.

Elia, I., Van den Heuvel-Panhuizen, M., & Kolovou, A. (2009). Exploring strategy use and strategy flexibility in non-routine problem solving by primary school high achievers in mathematics. *ZDM The International Journal on Mathematics Education*, *41*, 605–618.

Graham, J.W. (2009). Missing Data Analysis: Making It Work in the Real World. *Annual Review of Psychology*, *60*, 549–576.

Graham, J.W., Olchowski, A.E., & Gilreath, T.D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science*, *8*(3), 206–13.

Herscovics, N., & Linchevski, L. (1994). A cognitive gap between arithmetic and algebra. *Educational Studies in Mathematics*, *27*, 59–78.

Hoyles, C., & Sutherland, R. (1989). *Logo Mathematics in the Classroom*. London: Routledge.

Hox, J. (2002). Multilevel Analysis: Techniques and Applications. Mahwah, N J: Erlbaum.

Hsu, C.L., & Lu, H.P. (2004). Why do people play online games? An extended TAM with social influences and flow experience. Information and Management, *41*(7), 853–868.

Imhof, M., Vollmeyer, R., & Beierlein, C. (2007). Computer use and the gender gap: The issue of access use motivation and performance. *Computers in Human Behavior*, *23*, 2823–2837.

Janssen, J., & Engelen, R. (2002). Verantwoording van de toetsen Rekenen-Wiskunde 2002 [Account of the Mathematics tests 2002]. Arnhem: Citogroep.

Johanning, D. I. (2004). Supporting the development of algebraic thinking in middle school: A closer look at students' informal strategies. *Journal of Mathematical Behavior*, *23*, 371–388.

Kaiser, H. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*, 141–151.

Kaput, J. (2007). What is algebra? What is algebraic reasoning? In J. Kaput, D. Carraher, & M. Blanton (Eds.), *Algebra in the Early Grades* (pp. 5–18). New York: Lawrence Erlbaum.

Kaput. J., Carraher. D.W., & Blanton. M. (Eds.). (2007). *Algebra in the early grades*. New York: Lawrence Erlbaum.

Katz, V.J. (Ed.). (2007). *Algebra: Gateway to a technological future*. Washington, DC: Mathematical Association of America.

Ke, F. (2008). A case study of computer gaming for math: Engaged learning from gameplay? *Computers & Education*, *51*, 1609–1620.

Kebritchi, M., Hirumi, A., & Bai, H. (2010). The effects of modern mathematics computer games on mathematics achievement and class motivation. *Computers & Education*, *55*, 427–443.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.

Knuth, E.J., Stephans, A.C., McNeil, N.M., & Alibali, M.W. (2006). Does understanding the equal sign matter? Evidence from solving equations. *Journal for Research in Mathematics Education*, *37*, 297–312.

Kolovou, A. & Van den Heuvel-Panhuizen, M. (2010). Online game-generated feedback as a way to support early algebraic reasoning. *Int. J. Continuing Engineering Education and Life-Long Learning*, *20*(2), 224–238.

Kolovou, A., Van den Heuvel-Panhuizen, M., & Bakker, A. (2009). Non-routine problem solving tasks in primary school mathematics textbooks-A needle in a haystack. *Mediterranean Journal for Research in Mathematics Education, 8*, 31–67.

Lannin, J. (2005). Generalization and justification: The challenge of introducing algebraic reasoning through patterning activities. *Mathematical Thinking and Learning*, *7*, 231–258.

Lew, H.C. (2004). Developing algebraic thinking in early grades: Case study of Korean elementary school mathematics. *The Mathematics Educator*, *8*(1), 88–106.

Li, Q., & Ma, X. (2010). A Meta-analysis of the Effects of Computer Technology on School Students' Mathematics Learning. *Educational Psychology Review*, *22*, 215–243.

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*, 1198–1202.

McFarlane, A., Sparrowhawk, A., & Heald, Y. (2002). *Report on the educational use of games*. Retrieved August 15, 2010, from http://www.teem.org.uk/publications/teem_gamesined_full.pdf

Mendes-Barnett, S., & Ercikan, K. (2006). Examining sources of gender DIF in mathematics assessments using a confirmatory multidimensional model approach. *Applied Measurement in Education, 19*(4), 289–304.

Miles, J. & Shevlin, M. (2001). *Applying Regression and Correlation*. London: Sage.

Mitchell, A., & Savill-Smith, C. (2004). *The use of computer and video games for learning: A review of the literature*. London: Learning and Skills Development Agency.

Moseley, D., Mearns, N., & Tse, H. (2001). Using computers at home and in the primary school: where is the value added? *Educational and Child Psychology, 18*(3), 31–46.

Moss, J., & Beatty, R. (2006). Knowledge building in mathematics: Supporting collaborative learning in pattern problems. *International Journal of Computer-Supported Collaborative Learning*, 1, 441–466.

Mullis, I.V.S., Martin, M.O., & Foy, P. (2008). TIMSS 2007 International Mathematics Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mumtaz, S. (2001). Children's enjoyment and perception of computer use in the home and the school. *Computers & Education, 36*, 347–362.

National Council of Teachers of Mathematics (NCTM) (2000). *Principles and standards for school mathematics.* Reston, VA: NCTM.

OECD (2003). *The PISA 2003. Assessment framework- mathematics, reading, science and problem solving knowledge and skills*. Paris: Author.

Oosterwegel, A., Littleton, K., & Light, P. (2004). Understanding computer-related attitudes through an idiographic analysis of gender- and self-representations. *Learning and Instruction, 14*, 215–233.

Raudenbush, S.W., Bryk, A.S, & Congdon, R. (2004). HLM 6 for Windows [Computer software]. Lincolnwood, IL: Scientific Software International, Inc.

Redfield, C., Gaither, D., Redfield, N. (2009). COTS Computer Game Effectiveness. In R. Ferdig (Ed.), *Handbook of Research on Effective Electronic Gaming in Education* (pp. 277–294). Hershey, PA: Information Science Reference/IGI Global.

Resnick, M., Eisenberg, M., Berg, R., Mikhak, B., & Willow, D. (2000). Learning with Digital Manipulatives: New Frameworks to Help Elementary-School Students Explore "Advanced" Mathematical and Scientific Concepts. Proposal to the National Science Foundation.

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.

Rubin, D.B. (1987) Multiple Imputation for Nonresponse in Surveys. NJ: J. Wiley & Sons.

Ryan, K.E., & Chiu, S. (2001). An examination of item context effects, DIF, and gender DIF. *Applied Measurement in Education, 14*(1), 73–90.

Sanetti, L.H., & Kratochwill, T.R. (2009). Toward developing a science of treatment integrity: Introduction to the special series. School Psychology Review, *28*(4), 445–459.

Schafer, J.L., & Graham, J.W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147–177.

Schliemann. A.D., Carraher. D.W., & Brizuela. B.M. (2001). When tables become function tables. In M. Van den Heuvel-Panhuizen (Ed.), *Proceedings of the 25th conference of the international group for the psychology of mathematics education* (Vol. 4, pp. 145–152). Utrecht, The Netherlands: Freudenthal Institute.

Schliemann, A.D., Carraher, D.W., Brizuela, B.M., Earnest, D., Goodrow, A., Lara-Roth, S., et al. (2003). Algebra in elementary school. In N. Pateman, B. Dougherty, & J. Zilliox (Eds.), *Proceedings of the 27th PME International Conference* (Vol. 4, pp. 127–134). Honolulu, HI: University of Hawaii.

Schliemann, A.D., Carraher, D.W., Brizuela, B.M., & Jones, W. (1998). *Solving algebra problems before algebra instruction.* Paper presented at the Second Meeting of the Early Algebra Research Group, University of Massachusetts at Dartmouth & Tufts University, Medford, MA.

Schoenfeld, A. (1995). Report of working group 1. In: Lacampagne, C., Blair, W. and Kaput, J., (Eds.), *The algebra initiative colloquium*, (Vol. 2, pp. 11–18). Washington, DC: US Department of Education, Office of Educational Research and Improvement, and National Institute on Student Achievement, Curriculum, and Assessment.

Sedighian, K., & Sedighian, A. (1996). Can educational computer games help educators learn about the psychology of learning mathematics in children? In E.Jakubowski, D.Watkins, & H.Biske (Eds.), *Proceedings of the 18th annual meeting of the North America Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 573–578). Columbus, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.

Squire, K.D. (2003) Video games in education. *International Journal of Intelligent Games and Simulation 2*(1), 49–62. Retrieved August 15, 2010 from http://website.education.wisc.edu/kdsquire/tenure-files/39-squire-IJIS.pdf

Stacey, K., Chick, H., & Kendal, M. (Eds.). (2004). *The future of teaching and learning of algebra*. *The 12th ICMI study*. Boston: Kluwer.

Sutherland, R. & Rojano, T., (1993) 'A spreadsheet approach to solving algebra problems', *Journal of Mathematical Behavior, 12*, 353–383.

Sutherland, R., & Rojano, T. (1993). A spreadsheet approach to solving algebra problems. *Journal of Mathematical Behavior*, *12*, 353–383.

Trautwein, U. (2007). The homework–achievement relation reconsidered: Differentiating homework time, homework frequency, and homework effort. *Learning and Instruction, 17*, 372–388.

Trautwein, U., Lüdtke, O., Kastens, C., & Köller, O. (2006). Effort on homework in grades 5 through 9: Development, motivational antecedents, and the association with effort on classwork. *Child Development*, *77*, 1094–1111.

Trautwein, U., Lüdtke, O., Schnyder, I., & Niggli, A. (2006). Predicting homework effort: Support for a domain-specific multilevel homework model. *Journal of Educational Psychology*, *98,* 438–456.

Van Amerom, B. (2003). Focusing on informal strategies when linking arithmetic to early algebra. *Educational studies in Mathematics*, *54*, 63–75.

Warren, E.A., Cooper, T.J., & Lamb, J.T. (2006). Investigating functional thinking in the elementary classroom: Foundations of early algebraic reasoning. *Journal of Mathematical Behavior*, *25*, 208–223.

Winkelmann, H., van den Heuvel-Panhuizen, M., & Robitzsch, A. (2008). Gender differences in the mathematics achievements of German primary school students: results from a German large-scale study. *ZDM: The International Journal on Mathematics Education*, *40*, 601–616.

Zbiek, R., Heid, M., Blume, G., & Dick, T. (2007). Research on technology in mathematics education: The perspective of constructs. In F. Lester (Ed.), *Handbook of research on mathematics teaching and learning* (Vol. 2, pp. 1169–1207). Charlotte, NC: Information Age Publishing.

Chapter 7
An online game as a learning environment for early algebraic problem solving by students in upper primary school

Angeliki Kolovou [a], Marja van den Heuvel-Panhuizen [a], and Alexander Robitzsch [b]

[a] *Freudenthal Institute for Science and Mathematics Education, Utrecht University, the Netherlands*
[b] *Federal Institute for Education Research, Innovation and Development of the Austrian School System, Salzburg, Austria*

Kolovou, A., Van den Heuvel-Panhuizen, M., &. Robitzsch A. (submitted). An online game as a learning environment for early algebraic problem solving by students in upper primary school.

An online game as a learning environment for early algebraic problem solving by students in upper primary school

1 Introduction

In primary school mathematics the focus is mainly on developing numeracy and calculation skills. However, several researchers (Goldenberg, Shteingold, & Feurzeig, 2003; Harel, 2008) have pointed out that mathematics also involves ways of thinking like seeking and exploring patterns, making conjectures and performing experiments, and applying heuristics to solve nonstandard problems, which should also be a vital part of primary school instruction. Algebraic activities could offer students in primary school opportunities for developing more sophisticated thinking skills (Kaput, Carraher, & Blanton, 2007). Providing students with such a foundation of algebra in primary school could be a step towards making this mathematical domain accessible to all students and can prepare them for the learning of algebra in the later grades (Kaput, 2007). Moreover, the integration of algebra in the primary grades is also essential for adding coherence, depth, and power to the mathematics curriculum (Kaput, 2007).

However, the inclusion of algebra in primary school does not imply adding traditional algebra to the primary school curriculum; rather, it means providing entry points to algebra through treating existing number topics from the primary school curriculum in a deeper and more connected way (Kaput, et al. 2007). In fact, algebra builds on students' proficiency in arithmetic and develops it further (Kilpatrick, Swafford, & Findell, 2001). Along these lines, the National Council of Teachers of Mathematics (NCTM) suggested that algebra is a strand that permeates all levels of schooling from prekindergarten through grade 12 (NCTM, 2000). For example, the concept of function can build on experiences with numerical patterns in the primary school (NCTM, 2000). Rich problem contexts play an indispensable role in bringing about these experiences, as reasoning in particular situations may support students in generating abstract knowledge (Carraher & Schliemann, 2007). Function tables (Schliemann, Carraher, & Brizuela, 2001) or function machines (Warren, Cooper, & Lamb, 2006) and patterning activities (Moss & Beatty, 2006) have been successfully implemented to support this reasoning in primary school students, but new technologies might bring in further improvements for stimulating the development of algebraic concepts. Computers can carry out calculations and diminish the routine workload so that students can be more focused on exploring the relationships between the quantities in a problem. Furthermore, computer tools can surpass the constraints of paper-and-pencil activities by creating dynamic environments that provide instant feedback (Roschelle, Pea, Hoadley, Gordin, & Means, 2000).

The focus of this study is on solving contextual problems in which students have to deal with covariating values. These problems cannot be solved by carrying out calculations with the known values leading directly to the unknown value; rather, they require solution strategies that capture the relations between the known and unknown values (Van Dooren, Verschaffel, & Onghena, 2002). Not having being taught algebra yet, primary school students can only solve these problems by applying context-connected, informal reasoning. Although grounded in arithmetic, these solution methods can be the stepping stones to developing algebraic reasoning. In fact, dealing with such problems might induce reasoning about relationships between quantities, which is an essential aspect of algebraic problem solving (Humberstone & Reeve, 2008). For the sake of conciseness we call these problems *early algebra problems*. To create opportunities for students to experience covariation we developed a dynamic ICT environment including an archery game.

The purpose of the study was to investigate the informal strategies students used to solve early algebra problems and whether students' performance was influenced by working in this ICT environment. Furthermore, we wanted to know whether the applied strategies and the assumed effect of the ICT environment on the performance were different for students in grades 4, 5, and 6.

2.Theoretical background

2.1 Algebra in the primary school

The development of algebraic thinking in primary school has been approached from diverse perspectives, including exploring patterns and functional relations, constructing pictorial equations to solve problems, or even investigating relations between quantities at a formal level before dealing with numbers (Cai, 2004). Kieran (2004) emphasized that algebraic thinking in the primary grades should encompass general mathematical activities, such as problem solving, modeling, generalizing, and analyzing relationships. Although students can be engaged in such activities without using symbolic expressions, these activities can foster ways of thinking that are crucial for success in formal algebra (Kieran, 2004).

Other researchers have emphasized the role of problem solving in learning algebra as well. Van Amerom (2003) suggested that contextual problems form a link between arithmetic and algebraic reasoning; students can acquire algebraic concepts in an informal way through solving contextual problems involving numbers which are related in multiple ways. Similarly, Nathan and Koedinger (2000) pointed out that also in problems with an unknown start, the context can elicit the application of robust informal strategies, such as guess-and-check, that support the verification of the solution accuracy and error reduction. Students' informal approaches can be actually seen as ways in which they try to deal with algebraic situations without having to rely upon formal algebra (Johanning, 2004).

2.2 Variation theory

From a variation theory perspective a necessary condition for learning is the possibility to experience variation in critical aspects of the object of learning (Marton & Tsui, 2004; Runesson, 2006). The same phenomenon can be experienced or understood in a range of different ways depending on the features or aspects of a situation that are simultaneously discerned. For learning to occur, the learner should be able to distinguish and focus on critical aspects of variation in the phenomenon under study and become simultaneously aware of the possible values that these aspects can take (Marton & Tsui, 2004). Although experiencing patterns of variation is significant for learning mathematics in general, it is especially relevant to the teaching and learning of algebra (Al-Murani, 2006), since the ability to generalize from particular instances implies that one can distinguish between what changes and what remains invariant. According to Watson and Mason (2006), in comparison to unstructured sets of tasks, tasks that display constrained variation are likely to result in progress, provided that learners are working within mathematically supportive learning environments.

Moreover, variation theory is an analytical tool for describing, understanding, and designing learning (Runesson, 2006). In particular, it can be used for identifying different ways of understanding a phenomenon and designing learning situations by creating patterns of variation and invariance in relation to critical aspects of the object of learning (Runesson, 2006).

2.3 Role of computers in teaching and learning of mathematics

According to the National Council of Teachers of Mathematics (NCTM, 2000), technology influences not only how mathematics is taught and learned but also what is taught and when it is taught. Technological tools enable the quick and accurate execution of routine procedures, which allows more time for conceptualizing and modeling (NCTM, 2000). Digital tools provide access to external representations and feedback (Zbiek, Heid, Blume, & Dick, 2007) and support active engagement and interactive learning (Roschelle et al., 2000). Li and Ma (2010) in their meta-analysis of the effects of computer technology on mathematics achievement ascertained an overall positive impact. For example, simulations engage students in active exploration and discovery learning (Lou, Abrami, & Apollonia, 2001) and stimulate the formulation of strategies to deal with complex mathematical systems (Crown, 2003). Computer tools, such as spreadsheets (Lannin, 2005) and logo programming environments (Hoyles & Sutherland, 1989), have been also employed for the learning of algebra. Nathan (1998) showed that a dynamic computer environment offering animation-based feedback can support students in algebra problem solving. In this environment students were able to construct and test models of the relations between the quantities in the problem. Dynamic computer environments might be especially suitable for integrating dynamic processes, such as covariation.

More recently, Kebritchi, Hirumi, and Bai (2010) showed that computer games used for practicing algebraic concepts have a positive impact on mathematics achievement. These recent findings about games and the unique opportunities that technology provides for the development of dynamic thinking indicate that further research on the role of computer-based games in the development of early algebraic reasoning is needed.

3 Purpose of the study

In this study we provided students in upper primary grades (10- to 12-years-olds) with an online learning environment including a game by means of which they could experience situations with covariating variables. In particular, we sought to answer the following research questions:

> *1. How do fourth to sixth graders utilize an interactive online environment including a dynamic game to solve early algebra problems?*
> *2. Is students' working in this environment related to their performance in a written test on early algebra problem solving?*

4 Method

4.1 Participants

In total 318 students from grades 4, 5 and 6, were invited from five schools situated in five different city districts of a major Dutch city. The student population in these schools came from diverse socio-economic backgrounds.

4.2 Design

An experiment was set up with an intervention including an online computer activity that did not take place at school but at home (what we call 'outsourcing education'). Monitoring software was used to collect data on the students' online working. This software provides access to the students' performance from a distance (what we call 'tele-assessment'). Before and after the intervention a paper-and-pencil pretest and a posttest were administered.

To work online, each student received a unique login account and three sets of problems (eight in total) with the request to solve them in the online environment and write down their answers on a worksheet. They also received a short introduction to the online environment including instructions on how to log in and a demonstration of its features. The students had approximately a week at their disposal to work on each set of problems. At the end of each week, the students' presented their answers in a whole class discussion led by one of the authors.

The online activity was introduced as a game and was not part of the students' compulsory homework. The teachers had a minimum role in the intervention; they only stimulated their students to go online and play the game. At the beginning of the experiment both the students and their parents were informed about the registration of the online working.

4.3 Pretest and posttest on early algebraic problem solving

The paper-and-pencil pretest on early algebra problem solving included seven contextual problems (see Table 1).

Table 1
Items in pretest

Item name	Short description of item
Airplane seats	Total sits: 108; Occupied seats: twice as many as empty seats. How many occupied seats?
Age	Angela: 15 years old; Johan: 3 years old. In how many years is Angela two times as old as Johan?
Heads and feet	Rabbits and chickens: 34 heads and 92 legs altogether. How many are the rabbits?
Quiz	Correct answer: gets 2 points; Wrong answer: 1 point is subtracted; Questions: 10; Score: 8 points. How many correct questions?
Coins	18 five-cent and ten-cent coins; Value: €1, 50. How many 5-cent coins?
Pages	Total pages of book: 75; reading begins on Monday; Tuesday: 5 pages more than Monday; Wednesday: 5 pages more than Tuesday. How many pages read on Wednesday?
Hit the target	Hit: 2 points added; Miss: 1 point taken away; Arrows: 10; Score: 8 points. How many hits?

The posttest comprised the seven pretest items and six new items. Each new item corresponded in mathematical structure to an item in the pretest. However, three of the new items were presented in a different context and three as bare number problems. For example, the new context version of both the *Quiz* and the *Hit the target* item was the following item:

> *Throwing the ball* You get 3 points if you throw the ball in the basket; if you miss it, 1 point is taken away from your score. Marja threw 10 shots. She has 6 points in total. How many times did Marja shoot in the basket?

The bare-number version of the *Heads and feet* item is shown Figure 1.

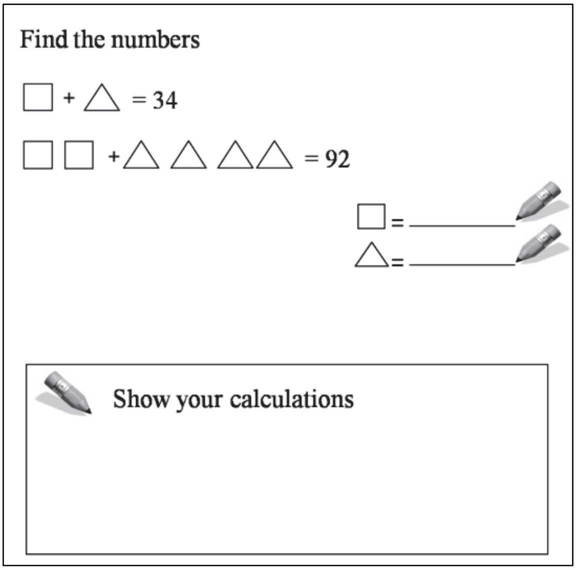


Fig. 1 Bare-number version of the *Heads and feet* item

All items can be solved by a formal algebraic approach, that it to say, setting up and solving a system of two linear equations with two unknowns, or by an arithmetic method in which students have to reason informally about the relationships between the quantities. For instance, they can pick a value to represent the unknown quantity, perform the calculations that are described in the problem, and check whether their guess leads to the intended result. Since formal algebra is introduced in the Netherlands in the first year of the secondary school, we expected that the students of our study had only arithmetic approaches at their disposal. The Cronbach's alpha was high both for the seven items in the pretest ($\alpha = .82$) and for the 13 items of the posttest ($\alpha = .92$).

4.4 The online environment

4.4.1 The game

The environment developed to offer students experience in dealing with covariating variables includes a dynamic java applet called *Hit the target*[1] (Figure 2), which is an interactive simulation of an archery game. The environment was piloted in a study with 24 fourth-grade students (Kolovou, Van den Heuvel-Panhuizen, Bakker, & Elia, 2008). The default screen of the game displays five features: The target, the pile of arrows and the bow, the score board, the board that displays the game rule, and the board that displays the number of hits and misses. In this game the students can set the shooting mode and the game rule mode. The shooting mode includes two options: shooting arrows one by one by dragging them to the bow (i.e., user shooting) or entering a number of hits, misses and random arrows and shooting them at once (i.e., computer shooting). When shooting a random arrow the student does not know in advance whether the arrow is going to be a hit or a miss. The game rule mode also includes two options: The students can set the game rule by filling in the points added or subtracted in case of a hit or a miss (i.e., user-defined game rule) or the computer sets the rule randomly (i.e., computer-defined game rule). The features of the game are dynamically linked. During the shooting or when removing arrows from the target the values on the scoreboard update rapidly to inform students about their score. In this way, students might realize that the number of hits and misses, the score and the game rule are dynamically related to each other. Moreover, the game offers instant visual feedback by showing the consequences of the students' actions.

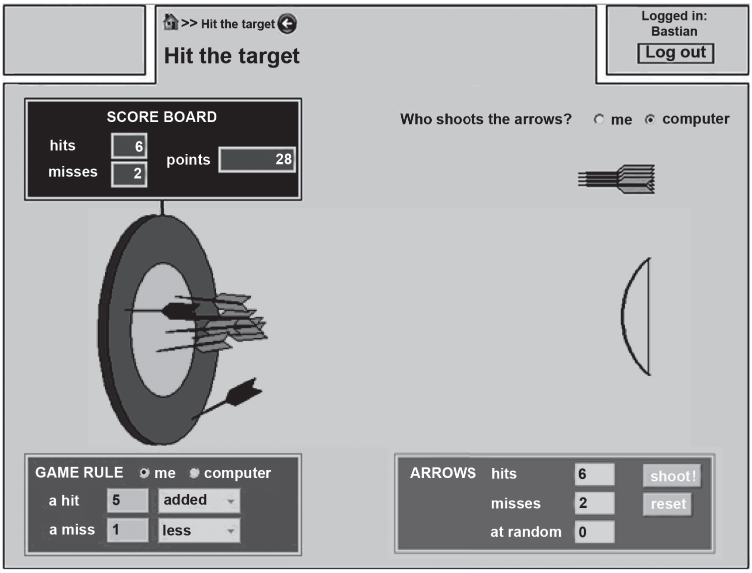


Fig. 2 Screen view of game in the computer shooting-mode

4.4.2 The online problems

The problems that the students had to solve with the game varied from finding how many hits and how many misses produce a given score to generating a general solution. A selection of the problems is shown in Figure 3.

In week 1
Problem 1: For every hit you gain 3 points and for every miss 1 point is taken away from your score. How many hits and misses do you have to shoot to get 15 points in total?
Problem 3a: What is the game rule to get 15 points in total with 15 hits and 15 misses?
Problem 3b: Are there other game rules to get 15 hits, 15 misses, and 15 points?

In week 2
Problem 4a: What is the game rule to get 16 points in total with 16 hits and 16 misses?
Problem 4b: Are there other game rules to get 16 hits, 16 misses, and 16 points?
Problem 5a: What is the game rule to get 100 points in total with 100 hits and 100 misses?
Problem 5b: Are there other game rules to get 100 hits, 100 misses, and 100 points? Can you explain your answer?

In week 3
Problem 8a: For every hit you gain 2 points and for every miss 1 point is taken away from your score. You have 10 arrows in total. How many hits and misses do you have to shoot to get 5 points in total?
Problem 8b: Are there any other solutions possible?

Fig. 3 Selection of problems given to the students during the intervention

Similarly to the items in the pre and posttest, all these problems can be solved by an algebraic procedure. In particular, the algebraic solution of the online problems includes a linear equation with two unknowns or a system of two linear equations with two unknowns. Although the students in our study had not been taught such approaches, in the online environment they had the opportunity to solve these problems informally by trying out solutions and receiving instant feedback. By doing this they might detect patterns and make generalizations. Variation in the numbers involved in the problems might help students grasp the invariant structure of a problem. In other words, students might be able to see the general pattern, despite the fact that the particular numbers vary. Since the computer can perform calculations very quickly – there is little time between one result and the next – students are able to experience the effects of the variation of the values.

Furthermore, the problems also varied with respect to the given and unknown variables. For example, in *Problem 1* they had to determine the number of hits and misses based on the score and the game rule, while in *Problem 3* the score and the number hits and misses were given and the students had to find the game rule. This variation in the given and unknown variables in the dynamic computer environment might also support the development of relational thinking.

4.4.3 The monitoring software

The game was linked to the so-called Digital Mathematics Environment (DME)[2], which is software that keeps track of students' actions while working online. With an account students can log in to the DME at any computer and resume their work. Students' actions while working online are also accessible through the DME logs for teachers or researchers. In the present study the log data consisted of a list of the actions carried out by each student in the online environment structured into sessions and events (see Figure 4). A session consists of the activity from the moment that the student logs in to the online environment until the student logs out and an event is a single shooting action performed when the student clicks on the shoot button.

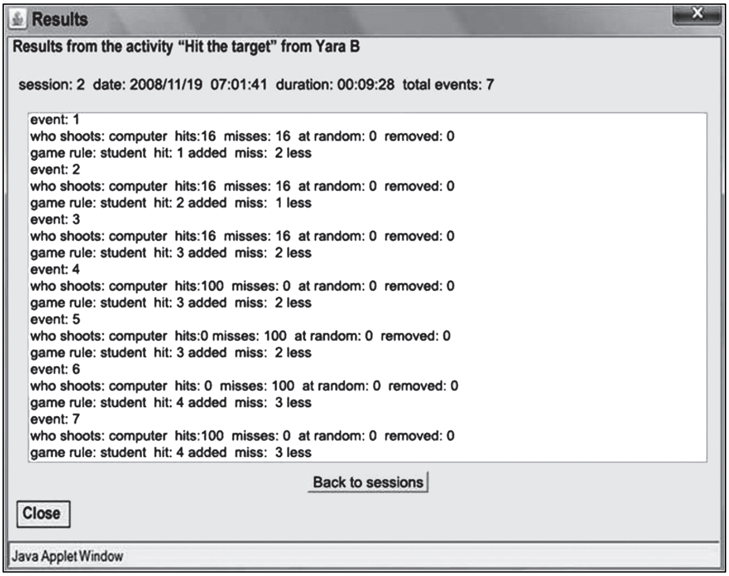


Fig. 4. Screen view of the log data generated by the DME

4.5 Analysis

Our data consisted of students' responses to the paper-and-pencil pre and posttest and the students' log files. The students' responses to the pre and posttest were coded according to the correctness of the answer. For each student the proportion of correct items was computed as an individual raw test score.

Students' log files were analyzed with respect to the online problems they attempted to answer and the strategies they used to solve these problems. We developed a coding scheme in several rounds. The final version of the scheme comprises 12 strategies (see Table 2).

Table 2
Coding scheme for online working

Variable Cluster	Description	Example (*Problem #*)
I. Worked problem		
P1, P2, P3a, P3b, P4a, P4b, P5a, P5b, P6a, P6b, P7a, P7b, P8a, P8b	The student (i) gave a correct answer (by performing an event or a series of consecutive events that lead to the correct answer), (ii) gave a partially correct answer, or (iii) tried to find an answer to a problem.	The student shot 15 hits and 15 misses and applied consecutively the game rules: +5 –2, +4 –2, +3 –2 (*Problem 3a*).
II. Strategy		
Altering/Ignoring information (Alter)	The student arrived at a result, but only part of the problem information was used.	The student shot 15 hits and 0 misses and assigned 1 point to each hit (*Problem 3a*).
Solving an analogous problem (Analogous)	The student substituted the numbers of a problem with smaller numbers.	Instead of shooting 100 hits and 100 misses the student shot 10 hits and 10 misses (*Problem 5a*).
Cancelling out (Cancelout)	The partial (negative) score of the misses cancels out the partial (positive) score of the hits. The total score becomes 0.	The student shot 15 hits and 15 misses and applied the game rule +1 −1 (*Problem 3a*).
Erroneously derived rule (Erroneous)	Based on a correct answer to a problem the student applied an erroneously derived rule to provide more answers to this problem.	Based on the correct game rule +2 –1 the student applied the game rule +4 –2 (i.e., ratio of *points per hit* to *points per miss* is 2) (*Problem 3b*).
Using extreme values (Extreme)	The student assigned the value 0 to one of the variables so that the other variable could get the maximum value.	The student used the game rule +1 –0 (1 point added for hit, 0 points less for a miss) (*Problem 3a*).

(*Table 2 continues*)

(*Table 2 continued*)

Variable Cluster	Description	Example (*Problem #*)
II. Strategy		
Applying a general rule (General)	The student applied a general solution.	The student applied the general rule where the sum of points per hit and points per miss is 1 e.g., +100 −99 (*Problem 3b*).
Repeating (Repeat)	The student repeated a (correct) answer to a problem to provide more answers to this problem.	The student repeated a correct answer to *Problem 3a* (e.g., +3 –2) in order to solve *Problem 3b*.
Reversing solution (Reverse)	The student reversed a correct answer to a problem to provide more answers to this problem.	The student first applied the game rule +2 –1 and then the rule −1+2 (*Problem 3a*).
Splitting the problem (Split**)**	The student answered a problem in two steps and added the partial scores to calculate (mentally) the total score.	The student shot first 100 hits and 0 misses, then 0 hits and 100 misses, and added the two partial results (*Problem 5a*).
Systematic trialing (Sys)	The student adjusted systematically the values until a (correct) answer was found.	The student applied consecutively the game rules: +6 −3, +6 −4, +6 −5 (*Problem 3a*).
Transposing values (Transpose)	The student exchanged the values of the arrows and the points.	The student shot 2 hits and 1 miss and used the game rule +100 −100 (*Problem 5a*).
Trial-and-error (TE)	After one or more trials the student came up with a (correct) answer.	The student applied several game rules until she came up with the intended result.

The first author assigned codes to the students' online activity according to the coding scheme. A single event was taken as the unit of analysis. To determine the interrater reliability one additional rater coded a subsection of the login data (N = 310 ratings). The agreement between the two raters was sufficient. The lowest percentage agreement of all the variables, with a value of .87, was obtained for the problem variable that indicates the variables had larger agreement (M = .98). To correct for agreement by chance Cohen's kappa was calculated. It ranged from .61 to 1.00 with a mean of .84 indicating sufficient reliability given the fact that some strategies were very rarely employed.

Subsequently, for each student the frequency and percentage of focused events and the frequency of worked problems were calculated. A focused event is an event meant to solve an online problem and the percentage of focused events is the proportion of the focused events to the total events of a student. A worked problem is an online problem that the student tried to solve with the game.

In order to analyze the relations between the online working and the students' performance we applied a number of parametric tests: t-tests and analysis of covariance. We chose parametric tests because the assumptions of normality and homoscedasticity of variances were approximately met. Moreover, because of sufficiently large samples (N of about 70 to 80 per grade) the central limit theorem ensures normally distributed regression coefficients. However, to validate the results from the parametric tests, we complemented the parametric version of a paired t-test with the nonparametric Wilcoxon signed rank sum test.

5 Results

5.1 Descriptive statistics

Since the online activity was not compulsory the students could decide whether they logged in or not and how often they worked in the online environment. In total, 253 students (80% of the participants) logged in at least once. In particular, the frequencies of the logged-in students from grades 4, 5, and 6 were 74%, 88%, and 78% respectively. The students who did not log in had the same pretest performance as the students who logged in ($t(309) = -1.72$, $p > .05$). Table 3 displays the means and standard deviations of the performance in the pretest and the posttest, and the characteristics of the online working that indicate the students' engagement in the online computer activity. The posttest performance is given for the seven joint items in the pre and the posttest as well as for the total of the13 posttest items. The characteristics of the online working include: login time, number of events, number of focused events, percentage of focused events and number of problems students worked on in each grade.

Table 3
Pre and posttest performance and characteristics of students' online work

	Grade 4 (*N*=75)		Grade 5 (*N*=73)		Grade 6 (*N*=84)		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Pretest	.06	.14	.23	.26	.38	.30	.23	.29
Posttest 7 items	.11	.20	.30	.33	.48	.36	.30	.34
Posttest 13 items	.09	.15	.28	.29	.45	.35	.28	.32
Login time (minutes)	35.70	41.63	25.75	18.69	32.22	54.46	31.31	42.15
N events	17.04	14.77	17.81	13.12	13.34	9.90	15.87	12.61
N focused events	5.68	7.61	9.60	8.15	7.48	8.07	7.57	8.03
% focused events	26.7	26.9	50.6	31.1	46.2	33.7	41.4	32.5
N worked problems	3.10	3.82	4.92	3.71	4.46	3.91	4.18	3.86

In general, the problems were quite difficult for the students, especially for the fourth graders. The gain score calculated by the difference of the score in the seven joint items in the post and pretest was significant across all grades with moderate effect sizes grade 4: $t(74) = 2.46$, $p < .05$, $d = .16$, grade 5: $t(72) = 2.93$, $p < .01$, $d = .24$, grade 6: $t(83) = 5.19$, $p < .001$, $d = .43$. The value of *d* was calculated using the standard deviation of the pretest of the whole sample. The same assertions with respect to the hypothesis test decisions were obtained using the Wilcoxon signed rank test statistic *W* (grade 4: $W = 2.59$, $p < .01$, grade 5: $W = 2.63$, $p < .01$, grade 6: $W = 4.63$, $p < .001$).

The standard deviations of the characteristics of the online working show a considerable variation in the way the students utilized the online environment. The students of grade 5 exhibited the highest number of events, focused events, and worked problems and the highest percentage of focused events, despite the fact they had the lowest login time.

5.2 Strategy use

In order to solve the problems, students spontaneously applied various strategies. Figure 5 shows the percentage of students who used a particular strategy per grade.

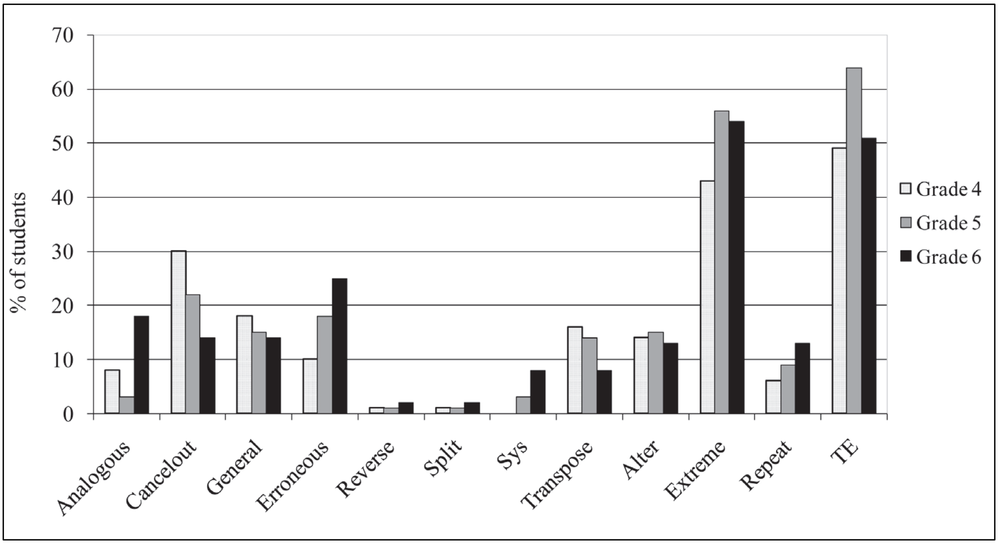


Fig. 5 Strategy use in grades 4, 5, and 6

The mostly applied strategy was trial-and-error (49%-64% of the students). Systematic trials were, however, performed by a few students, with the highest frequency (8%) in grade 6. A high percentage of students (43%-56%) used an extreme value to find a solution. Solving an analogous problem, splitting a problem into two sub-problems and using the transposing strategy were less frequently applied. The reason for the low percentages of these strategies is that they were particularly evoked when students tried to solve *Problem 5.* Since the maximum number of arrows to be shot at once in the game is restricted to 150, the students cannot shoot 100 hits and 100 misses at once; which means that they should come up with another way to solve the problem.

Besides differences in the frequency of the strategies we also found that the online environment triggered the application of qualitatively different strategies. In addition to strategies like Alter, Extreme, Repeat, and trial-and-error, which implied a focus on providing the (correct) solution, we also detected sophisticated strategies, such as solving an analogous problem, cancelling out, generalizing, applying an erroneously derived rule, reversing, systematic trialing, and transposing, which indicate exploration of relations and structure. Table 4 displays how often a sophisticated strategy was used per grade. The frequencies in the three grades were quite similar. The high positive value of skewness indicates that there were a few students in the sample who frequently used sophisticated strategies. Also the large standard deviation shows a substantial variability in the application of these strategies.

Table 4
Frequency of the application of sophisticated strategies per grade

	M	*SD*	Min.	Max.	Skewness
Grade 4	1.84	3.08	0	16	2.49
Grade 5	1.82	2.88	0	14	2.23
Grade 6	1.89	3.02	0	18	2.76

In order to analyze the influence of the online working on the posttest performance, we took two perspectives. The first perspective was more exploratory with the focus on the effect of students' strategy profiles on their performance. The second perspective was confirmatory in nature in the sense that we categorized the students into levels of online working.

5.3 Profiles of strategy use and their effect on performance

Initially, we determined for every type of strategy applied by a student, the ratio of the number of events related to this strategy to the student's total number of events. These ratios formed together the student's individual strategy profile. Subsequently, we performed a cluster analysis on these profiles and explored the effect of the clusters on the gain in performance.

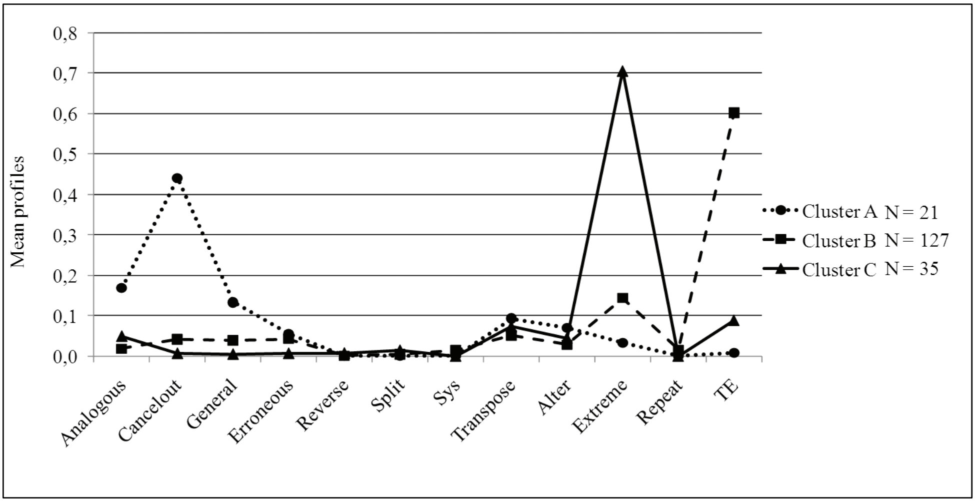


Fig. 6 Clusters of students' profiles of strategy use

We performed a *K*-means cluster analysis in the whole sample. Each of the clusters is characterized by aggregating the individual profiles resulting in a mean profile per cluster. A three-cluster solution (Figure 6) gave the best interpretable mean profiles. On the one hand, this approach gave a clear peak for the dominant strategy in each cluster. On the other hand, having more than three clusters resulted in too few students per cluster. Next, we assessed the stability of the cluster solution by a random split of the sample. In the two sub samples approximately the same mean profiles appeared, i.e. the dominant strategy turned out to be the same in the corresponding mean profile. The correlation (and in parentheses the Spearman rank correlation r_S) of the mean vectors in both samples amounts to $r = .72$ ($r_S = .82$) for Cluster A, $r = .99$ ($r_S = .78$) for Cluster B and $r = .99$ ($r_S = .76$) for Cluster C, which indicate sufficient stability of the cluster mean profile.

Within the three-cluster solution, in Cluster A the most dominant strategy was cancelling out followed by solving an analogous problem and applying a general rule. Cluster B was the cluster to which most of the students belonged. This cluster is characterized by the use of the trial-and-error strategy, while Cluster C is dominated by the Extreme strategy. Because the cluster analysis based on all grades already resulted in small sample sizes for the Clusters A and C, we did not repeat the cluster analysis per grade.

Then, we examined the effect of the clusters on the gain score for the seven joint items. In the whole sample the effects of Cluster B and Cluster C on the increase in performance were significant with effect sizes $d = .32$ ($M = .09$, $t(116) = 4.56$, $p < .001$, $W = 4.24$) and $d = .49$ ($M = .14$, $t(30) = 3.37$, $p < 0.01$, $W = 2.87$) respectively. For Cluster A we found an effect of $d = .19$ but this was not significant ($M = .06$, $t(17) = 1.60$, $p = .130$, $W = 1.55$). This means that the student's individual strategy profiles in Clusters B and C were more powerful than those in Cluster A.

5.4 Levels of online working

We assumed that not only the applied strategies would influence students' problem solving performance, but also the characteristics of students' online working. Therefore, we distinguished three levels of online working. Level 1, performing free playing, includes students that put little or no effort in answering the given problems; in particular, the students who performed less than three focused events or tried to answer less than three problems. Level 2, (just) looking for the answer, included students who mainly (just) tried to answer the problems. This means that they exhibited an activity equal or beyond the threshold of three focused events and three worked problems and that their main concern was to comply with our instructions and provide correct solutions. This type of activity is indicated by the use of less sophisticated strategies. Level 3, exploring relations and structure, includes students who also exceeded the minimum amount of activity and used sophisticated strategies. The percentage of students who performed free playing was the highest in grade 4, while the majority of students of grade 5 were just looking for answers (Table 5). The percentage of students who explored relations and structure slightly increased from grade 4 to grades 5 and 6.

Table 5
Percentages of students who applied a particular level of online working per grade

	Grade 4	Grade 5	Grade 6	Total
Free playing (Level 1)	60.8	29.5	39.6	43.1
(Just) looking for answers (Level 2)	12.7	41.0	28.1	27.3
Exploring relations and structure (Level 3)	26.6	29.5	32.3	29.6

To explore the effect of the levels of online working we conducted an analysis of covariance with the posttest as independent variable and the pretest as covariate. We performed this analysis per grade (Table 6). We took level 1 as the baseline level. In grade 4 the factor level had a significant effect on the posttest performance ($F(2, 71) = 3.27$, $p <.05$). Furthermore, there was an increase in performance between level 1 and level 2 and between level 2 and level 3. Posthoc tests showed that the difference between level 1 and 3 was significant ($p < .05$), while the difference between level 1 and level 2 was not significant. For grade 5 the factor level was not significant ($F(2, 69) = .90$, $p = .412$), but there was an increase in performance across the levels. In grade 6 the level of online working had a significant effect on the posttest performance ($F(2, 80) = 6.04$, $p < .01$). However, the increase in performance was larger for level 2 than for level 3. Also, we observed a significant ($p < .05$) increase in performance between level 1 and 2. The difference between level 1 and 3 was not significant.

Table 6
Effect of the levels of online working on posttest performance per grade

	B	*SE*	*t*	*p*
Grade 4				
Intercept	.03	.02	1.59	.117
Pretest	.58	.10	6.05	.000
Level 2 vs. Level 1	.03	.039	.64	.524
Level 3 vs. Level 1	.08	.032	2.56	.013
Grade 5				
Intercept	.05	.04	1.13	.264
Pretest	.83	.09	9.18	.000
Level 2 vs. Level 1	.04	.06	.69	.495
Level 3 vs. Level 1	.08	.06	1.34	.186
Grade 6				
Intercept	.08	.04	1.99	.050
Pretest	.82	.07	11.10	.000
Level 2 vs. Level 1	.19	.06	3.45	.001
Level 3 vs. Level 1	.11	.06	1.90	.061

5.5 Excerpts of student's online working

To illustrate how the online environment supported students' problem solving we describe the online working of two sixth-grade students. The log file of student A (Figure 7) shows that she gave a solution to *Problem 3* by trial-and-error. In *Problem 4* she arrived at the general rule for obtaining the result "16 hits–16 misses–16 points" and later on applied this rule in solving *Problem 5* by shooting 10 hits and 10 misses (i.e., the Analogous strategy). Table 7 presents the online activity of student A in a compact format including the total score per event.

```
session: 1   date: 2008/11/17 04:13:58   duration: 00:05:59   total events: 2
event: 1
who shoots: computer   hits: 15   misses: 15   at-random: 0   removed: 3 (hits: 2 misses: 1)
game rule: student   hit: 4 added   miss: 2 less
event: 2
who shoots: computer   hits: 15   misses: 15   at-random: 0   removed: 0
game rule: student   hit: 1 added   miss: 1 added
session: 2   date: 2008/11/17 04:20:43   duration: 00:02:52   total events: 1
event: 1
who shoots: computer   hits: 15   misses: 15   at-random: 0   removed: 0
game rule: student   hit: 5 added   miss: 4 less
                                                        } Problem 3
session: 3   data: 2008/11/25 04:02:24   duration: 00:36:03   total events: 6
event: 1
who shoots: computer   hits: 16   misses: 16   at-random: 0   removed: 0
game rule: student   hit: 2 added   miss: 1 less
event: 2
who shoots: computer   hits: 16   misses: 16   at-random: 0   removed: 0
game rule: student   hit: 3 added   miss: 2 less
event: 3
who shoots: computer   hits: 16   misses: 16   at-random: 0   removed: 32
(hits: 16 misses: 16)
game rule: student   hit: 100 added   miss: 99 less
                                                        } Problem 4
event: 4
who shoots: computer   hits: 10   misses: 10   at-random: 0   removed: 0
game rule: student   hit: 2 added   miss: 1 less
event: 5
who shoots: computer   hits: 10   misses: 10   at-random: 0   removed: 0
game rule: student   hit: 1000 added   miss: 999 less
event: 6
who shoots: computer   hits: 10   misses: 10   at-random: 0   removed: 0
game rule: student   hit: 50 added   miss: 49 less
                                                        } Problem 5
```

Fig. 7 Log file of Student A

Table 7
Focused events of Student A

Problem	Session	Event	Hits	Misses	Game rule		Score
					Points per hit	Points per miss	
15 hits-15 misses-15 points	1	1	15	15	+4	−2	30
		2	15	15	+1	−1	0
	2	1	15	15	+5	−4	15
16 hits-16 misses-16 points	3	1	16	16	+2	−1	16
		2	16	16	+3	−2	16
		3	16	16	+100	−99	16
100 hits-100 misses-100 points		4	10	10	+2	−1	10
		5	10	10	+1000	−999	10
		6	10	10	+50	−49	10

Table 8
Focused events of Student B

Problem	Session	Event	Hits	Misses	Game rule		Score
					Points per hit	Points per miss	
15 hits-15 misses-15 points	1	1	15	15	+2	−1	15
		2	15	15	+4	−2	30
		3	15	15	+4	−3	15
16 hits-16 misses-16 points	2	1	1	0	+16	−16	16
		2	4	3	+16	−16	16
		3	2	1	+16	−16	16
100 hits-100 misses-100 points		4	1	0	+100	−100	100
		5	4	3	+100	−100	100

Student B (Table 8) also applied trial-and-error for solving *Problem 3*. From her log file we derived that she used the results from previous attempts to inform subsequent trials. Based on her initial answer (+2 −1) in the following event she applied an erroneous rule (+4 −2), but subsequently adjusted the values of the game rule, which resulted in a correct answer (+4 −3). To solve problems 4 and 5 she applied the Transpose strategy and provided several pairs of values for the hits and misses that led to the intended score. It seems that the student was able to provide a series of correct solutions due to her discovery of the underlying relation between the variables.

6 Discussion

This study looked at how an online learning environment with a game and a series of problems can contribute to fourth to sixth graders' ability to solve early algebra problems. In particular, we investigated the gain in performance in a paper-and-pencil test, the strategies that students applied while working online and the effects of this online working on students' performance in the paper-and-pencil test.

The analysis revealed that the gain in performance between the pre and posttest was the highest in grade 6. To explain the increase in performance, we examined the students' online working. It was found that the learning environment brought about the application of various types of strategies. As expected, the dynamic character of the game offered students the opportunity to apply sophisticated strategies with which they explored relations and structure. Furthermore, trial-and-error played a significant role. It was the most often applied strategy and the dominant strategy in one of the strategy profiles that had a significant effect on the gain in performance. This major role of trial-and-error is rather intriguing. One the one hand, it is considered a poor solution method. It is mostly related to finding a local solution to a problem. Moreover, as Lannin (2005) suggested, it might be the case that students do not reflect on the process when applying this strategy and therefore do not understand why a particular generalization is valid. In fact, the game allowed students to find an answer without reflection. On the other hand, trial-and-error can be also regarded as a very powerful strategy. Through performing a sequence of purposeful trials students might grasp the relation between the input and output values so that they use this knowledge to determine solutions to algebra problems (Levin, 2008).

Despite the short duration of the intervention, the online working was positively related with success in the posttest. Controlling for the pretest performance, the level of online working had a significant effect on posttest performance in grades 4 and 6. In particular, in grade 4 the students who belonged to the level of exploring relations and structure outperformed their peers. In grade 6 the students who belonged to the level of looking for answers obtained the highest posttest scores. In both grades the students who belonged to the level related to the highest posttest scores performed significantly better than their peers who exhibited free playing. Apparently, free playing did not contribute to success rate. This is in line with previous research that has shown that a simulation by itself is not sufficient to enhance learning; rather, students should be steered by appropriate assignments that stimulate the generation and testing of hypotheses (De Jong, 2005). In our study it means that in order to bring about learning effects, game playing should be accompanied by working on problems that elicit the discovery of the relations between the variables in the game. In addition to working on the problems, the sequence of the problems turned out to be important as well. The log files revealed that the variation in the problems was a crucial factor for discovering these relations. By working on the series of problems students could experience how the values covariate, which prompted the discovery of the general relation between these values.

However, some limitations should be kept in mind when interpreting the results of the study. A first point of attention is the size of our sample and the representativity of it. Because there are implausible grade-wise patterns it could be doubted if students of three grades are typical. It could be the case that due to cluster sampling of students, there is a higher probability of getting non-representative samples leading to counter-intuitive results. Another issue related to the sample size is the lack of significant differences in the effects of the levels of online working on the posttest performance. To examine whether a larger sample size would give a different result we carried out a posthoc power analysis. Given the observed effect size of the explained proportion of variance of .08 in the analysis of covariance for grade 4, a sample of about 110 students is necessary to get a power of .80 to reject the null hypothesis of no effect (Murphy, Myors, & Wolach, 2009). Therefore, in replications studies of this experiment a somewhat increased sample size would be helpful to detect significant differences.

Other shortcomings have to do with the nature of the online environment. Although we assume that the student-driven animation supported the understanding of covariation, the opposite might also be the case at least for a part of the students. Because of their association with entertainment, dynamic visualizations might create an illusion of understanding (Bétrancourt, 2005) or result in a superficial engagement with the content (Lowe, 2004). Since students can check the correctness of an answer by shooting arrows, they might not need to engage in the more tedious work of exploring the relations between the variables. Moreover, our conclusions are based on the log files of the students' online activity, which might not entirely capture the students' cognitive processes, such as the mental calculations that they performed. Nevertheless, the collected data provided a step-by-step account of students' interaction with the computer environment.

Furthermore, we could not control whether the students worked on the online problems on their own or with the help of a peer or family member. However, it is less likely that a voluntary activity that was not related to formal assessment induced the need for assistance. In addition, the diversity in the students' engagement in the online computer activity is an indication that the students did work on their own; had parents been involved, it would have been more likely for the students to try to solve all the presented problems.

However, in general, our results suggest that a dynamic computer environment can stimulate algebraic reasoning in the primary grades. Home computing may create an effective learning environment supporting and extending school learning. Yet, further investigations are needed; in particular, it would be interesting to examine the influence of the game when introduced as a classroom activity supervised and supported by the teacher.

Notes

1. The game *Hit the target* is developed by the second author and programmed by our colleague Huub Nilwik.
2. The Digital Mathematics Environment (DME) is developed by our colleague Peter Boon.

References

Al-Murani, T. (2006). Teachers' awareness of dimensions of variation: A mathematics intervention project. In J. Novotná. H. Moraová. M. Krátká & N. Stehlikova (Eds.), *Proceeding of the 30th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 25–32). Prague. Czech Republic: PME.

Bétrancourt, M. (2005). The animation and interactivity principles in multimedia learning. In R.E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 287–296). Cambridge University Press.

Cai, J. (2004). Developing algebraic thinking in the earlier grades from an international perspective. *The Mathematics Educator, 8*, 1–5.

Carraher, D.W., & Schliemann, A.D. (2007). Early Algebra and Algebraic Reasoning. In F. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (Vol. 2, pp. 669–705). Charlotte. NC: Information Age Publishing.

Goldenberg, E.P., Shteingold, N., & Feurzeig, N. (2003). Mathematical habits of mind for young children. In F. K. Lester & R. I. Charles (Eds.), *Teaching mathematics through problem solving: Prekindergarten-Grade 6* (pp. 15-29). Reston. VA: NCTM.

Crown, W.D. (2003). Using technology to enhance a problem-based approach to teaching: What will and what will not work. In F. K. Lester & R. I. Charles (Eds.), *Teaching Mathematics through Problem Solving. Prekindergarten-Grade 6* (pp. 217–228). Reston. VA: NCTM.

De Jong, T. (2005). The guided discovery principle in multimedia learning. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 215–228). New York: Cambridge University Press.

Hoyles, C., & Sutherland, R. (1989). *Logo Mathematics in the Classroom*. London: Routledge.

Harel, G. (2008). What is mathematics? A pedagogical answer to a philosophical question. In B. Gold & R. A. Simons (Eds.), *Proof and Other Dilemmas: Mathematics and Philosophy* (pp. 265–290). Washington. DC: Mathematical American Association.

Humberstone, J. & Reeve, R.A. (2008). Profiles of algebraic competence. *Learning and Instruction 18*, 354–367.

Johanning, D.I. (2004). Supporting the development of algebraic thinking in middle school: A closer look at students' informal strategies. *Journal of Mathematical Behavior, 23*, 371–388.

Kaput, J. (2007). What is algebra? What is algebraic reasoning? In J. Kaput, D. Carraher, & M. Blanton (Eds.), *Algebra in the Early Grades* (pp. 5–18). New York: Lawrence Erlbaum.

Kaput, J., Carraher. D.W., & Blanton. M. (Eds.) (2007). *Algebra in the early grades*. New York: Lawrence Erlbaum.

Kebritchi, M., Hirumi, A., & Bai, H. (2010). The effects of modern mathematics computer games on mathematics achievement and class motivation. *Computers & Education, 55*, 427–443.

Kieran, C. (2004). Algebraic thinking in the early grades: What is it? *The Mathematics Educator, 8*, 139–151.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.) (2001). *Adding it up: Helping children learn mathematics*. Washington, D.C.: National Academy Press.

Kolovou, A., Van den Heuvel-Panhuizen, M., Bakker, A. and Elia, I. (2008). An ICT environment to assess and support students' mathematical problem-solving performance in non-routine puzzle-like word problems In A. Gagatsis (Ed.), Research in Mathematics Education (pp. 175 – 190). Nicosia, Cyprus: University of Cyprus.

Lannin, J. (2005). Generalization and justification: The challenge of introducing algebraic reasoning through patterning activities. *Mathematical Thinking and Learning, 7*, 231–258.

Levin, M. (2008). The potential for developing algebraic thinking through purposeful guessing and checking. In G. Kanselaar, J. van Merriënboer, P. Kirschner, & T. de Jong (Eds.), *Proceedings of the 8th the International Conference of the Learning Sciences* (Vol. 1, pp. 469–476). Utrecht, The Netherlands: International Society of the Learning Sciences.

Li, Q., & Ma, X. (2010). A Meta-analysis of the Effects of Computer Technology on School Students' Mathematics Learning. *Educational Psychology Review*, *22*, 215–243.

Lou, Y., Abrami, P.C., & d'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research*, *71*, 449–521.

Lowe, R.K. (2004). Animation and learning: Value for money? In R. Atkinson, C. McBeath, D. Jonas-Dwyer & R. Phillips (Eds), *Beyond the comfort zone: Proceedings of the 21st ASCILITE Conference* (pp. 558-561). Perth, Australia: ASCILITE. Retrieved 16 February, 2011
http://www.ascilite.org.au/conferences/perth04/procs/lower.html

Marton, F., & Tsui, A.B.M. (2004). Classroom Discourse and the Space of Learning. Mahwah, N.J: Lawrence Erlbaum.

Moss, J., & Beatty, R. (2006). Knowledge building in mathematics: Supporting collaborative learning in pattern problems. *International Journal of Computer-Supported Collaborative Learning, 1*, 441–466.

Murphy, K.R., Myors, B., & Wolach, A. (2009). *Statistical power analysis*. New York: Routledge.

Nathan, M.J. (1998). Knowledge and situational feedback in a learning environment for algebra story problem solving. *Interactive Learning Environments, 5*, 135–159.

Nathan, M.J. & Koedinger, K.R. (2000). Teachers' and researchers' beliefs about the development of algebraic reasoning. *Journal for Research in Mathematics Education*, *31*, 168–190.

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Reston. VA: NCTM.

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.

Runesson, U. (2006). What is possible to learn? On variation as necessary conditions for learning. *Scandinavian Journal of Educational Research, 50*(4), 397–410.

Schliemann, A. D., Carraher, D. W., & Brizuela, B. M. (2001). When tables become function tables. In M. Van den Heuvel-Panhuizen (Ed.), *Proceedings of the 25th conference of the international group for the psychology of mathematics education* (Vol. 4, pp. 145–152). Utrecht: Freudenthal Institute.

Van Amerom. B, (2003). Focusing on informal strategies when linking arithmetic to early algebra. *Educational studies in Mathematics, 54*, 63–75.

Van Dooren, W., Verschaffel, L., & Onghena, P. (2002). The impact of preservice teachers' content knowledge on their evaluation of students' strategies for solving arithmetic and algebra word problems. *Journal for Research in Mathematics Education, 33*, 319–351.

Warren, E.A., Cooper, T.J., & Lamb, J.T. (2006). Investigating functional thinking in the elementary classroom: Foundations of early algebraic reasoning. *Journal of Mathematical Behavior, 25*, 208–223.

Watson, A., & Mason, J. (2006). Seeing an exercise as a single mathematical object: Using variation to structure sense-making. *Mathematical Thinking and Learning*, *8*, 91–111.

Zbiek, R., Heid, M., Blume, G., & Dick, T. (2007). Research on technology in mathematics education: The perspective of constructs. In F. Lester (Ed.), *Handbook of research on mathematics teaching and learning* (Vol. 2, pp. 1169–1207). Charlotte, NC: Information Age Publishing.

Chapter 8

Conclusion

Conclusion

1 Gains from the study on problem solving

The main purpose of this PhD study was to investigate and support the problem solving competence of primary school students. The study was carried out in the Netherlands and the focus was on solving context problems with covariating values. These problems require higher-order thinking and modeling, and cannot be solved in a straightforward routine way. Besides that these problems are important as such, because they offer students' opportunities to develop problem solving skills and mathematical reasoning, they also provide entry points to the development of algebraic thinking and prepare students for the study of formal algebra in the secondary school. Primary school students can apply informal context-connected strategies to deal with these problems. However, the question is whether they succeed in solving them, how they solve them and what kind of support contributes to their problem solving ability. These are the issues that this study addressed.

1.1 Students' problem solving performance and strategies

The first research question was related to the performance and strategies of Dutch primary school students in non-routine problem solving. To answer this question we analyzed students' work collected during the first phase of the POPO project and in particular, the written down solution steps of 152 high achievers in grade 4. We investigated the nature of the students' strategies as well as the flexibility in strategy use. The results of this study confirmed what the preliminary results of the POPO project (Van den Heuvel-Panhuizen & Bodin, 2004) had already shown. The problem solving performance of high achievers fell short of our expectations and the students often lacked essential problem solving skills. The percentages of (observed) strategy use varied from problem to problem and in general were relatively low. Students were not prone to writing down the solution steps. This tendency might be related to the students' belief that solving a problem mentally indicates a higher level of reasoning and that using scrap paper to make notes, although it can be very helpful (students can try out different solutions or check their results), means a lower level of understanding. Moreover, in discussions with students and teachers and other professionals at schools, it came to the fore that according to students a good solution is a neat solution that leads to the correct answer. It goes without saying that such a belief is in sharp contrast with the nature of problem solving as a complex and non-linear activity and seriously hampers productive solution processes. Research has shown that writing down notes can be an effective tool in supporting metacognitive behaviors in problem solving (Pugalee, 2004) and it is positively related to problem solving success (Williams, 2003; Van den Heuvel-Panhuizen & Bodin, 2004; Kolovou, Van den Heuvel-Panhuizen, & Elia, 2009).

Among the strategies that the students applied, trial-and-error had the highest potential to lead to a correct solution. Although considered a less sophisticated strategy, trial-and-error can support students in handling relations between quantities (Johanning, 2004), which is crucial in solving the types of problems that we presented to the students. Also, Stacey and McGregor (2000) showed that purposeful trials, in which information from previous guesses is used to make a subsequent guess, led to the correct solution of algebra problems. Our study showed that with respect to strategy flexibility, students mainly used a single strategy and failed to consider any alternative or complementary strategy. Furthermore, we found that the construction of a complete model of a problem is the foundation for a successful solution. In other words, the ability to (flexibly) apply problem solving strategies does not lead to the correct solution when students fail to use all problem information.

1.2 Opportunities to learn problem solving

We conducted a textbook analysis to answer the second research question about the opportunities that textbooks offer students to learn problem solving. Since our first results were drawn from students in grade 4, we analyzed the cognitive demand of the tasks in the textbook series of this grade. To do this we developed a framework encompassing three main categories of tasks: Straightforward tasks, non-routine puzzle-like tasks (mainly tasks that in later studies we called early algebra problems) and 'gray-area' tasks. This latter category includes tasks that have a lower cognitive demand, but are still not straightforward. The analysis showed that challenging non-routine puzzle-like tasks are scarce and are also mostly to be found outside the main book of the textbook series. Thus, students rarely have the opportunity to encounter tasks that could support them in developing problem solving skills. Although the written curriculum and the enacted curriculum (how the tasks in the textbooks are actually dealt with in the classroom) are two different entities, the absence of such tasks in the textbooks strongly suggests that they are absent in mathematics classes as well. In fact, Dutch primary school teachers, who teach all subjects, use textbooks as their main guide for planning their teaching.

These findings about the marginal place problem solving has in Dutch primary school and, consequently, the lack of models of how to teach it, set us on the track of searching for ways to support students' problem solving competence.

1.3 Helpful support in problem solving

The study described in Chapter 2 revealed that students' difficulties in the presented problems did not lie in performing the necessary calculations, but in grasping the relations between the numbers or quantities involved in the problems. For example, in the *Quiz* problem (correct answer 2 points added; incorrect answer 1 point subtracted; 10 questions; 8 points; how many correct answers), several students failed to take into account that the score is not only determined by the number of the correct answers, but also by the number of the wrong answers. As a result, they claimed that the number of correct answers is $4 \times 2 = 8$ and did not consider the number of wrong answers or the total number of questions. Therefore, we assumed that an environment where relations between numbers or quantities become more explicit can offer students opportunities to experience the interdependency of numbers or quantities. The environment that was developed for this purpose was the computer game *Hit the target,* which is a simulation of an archery game. In this environment the number of hits, misses, the rule of the game and the score are dynamically linked; by manipulating the values of the hits and the misses or the game rule, the student can observe the resulting changes on the score.

The game was first tried out by three students from grade 4. Based on these observations some changes in the game interface were implemented. Next, we piloted the game in a small-scale experiment with 24 high-achievers from grade 4. The students of the experimental group worked in pairs following a pre-defined scenario including three problem solving tasks and their work was registered by screen videos. The analysis of the audiovisual data revealed that all pairs of students solved the tasks successfully by generating and testing hypotheses and sharing ideas. Moreover, the students applied various strategies and were able to generalize their solutions. However, the students demonstrated different levels of problem solving activity. Some students were deliberately checking their hypotheses with the game, while others were just trying out numbers at random. The results showed that the experimental group gained slightly from the treatment. Because the data were too few, it did not make sense to look for a possible effect of working in the ICT environment on students' learning.

Therefore, we replicated the study with a large sample of 785 students from grades 4, 5, and 6 from ten primary schools in Utrecht. Because of the number of participants, we opted for monitoring students' work by log files. Compared to screen videos, the log files do not require real-time analysis of the students' activity. However, they are less precise than screen videos, because they display students' work in a more condensed way (Van den Heuvel-Panhuizen, Kolovou, & Peltenburg, 2011). The students of the five experimental schools were invited to log in the online environment at home and carry out a number of problems. These problems varied from finding the pair of hits and misses that produces a particular score to generating a general rule by systematizing all solution pairs. In particular, we wanted to investigate whether the online environment could support students' problem solving performance.

The students of the experimental group received a unique login account and a series of eight problems split in three sets and were asked to solve the problems at home. The intervention period lasted three weeks. At the end of each week the students presented their answers in short follow-up discussions in class. The purpose of these discussions was not to teach students how to solve the problems, but mainly to provide feedback on their answers and sustain their participation. Students' engagement in the game-playing was quite divergent. Some students did not log in at all, while others solved a part of the given problems. This was an expected outcome, since we initiated the game-playing as a voluntary homework activity.

The analysis of the data for grade 6 showed that the students of the experimental group outperformed their peers in the control group. When we compared the group of logged-in students with the control group, we found similar results. However, when we compared the students who logged in with those who did not, we found no significant differences. A reason for this could be that the number of students who did not log in was quite small, which might mean that the statistical analysis lacked the power to yield significant results. Another reason is more substantial. Because the students who did not log in were part of the experimental group they were also exposed to the dialogues in class, which may have contributed to an improved performance in the posttest.

The effort put into the online activity, indicating frequency, duration, and focus of the online working, did not have a significant effect on the gain in performance. Although this result is counterintuitive, it might be that the improvement in the problem solving performance cannot be solely attributed to the amount of effort. For example, although girls outperformed boys in some effort indicators, their achievement gains were the same as for boys. It seems that the role of effort on performance is a complex one and is worthy of further investigations.

To conclude, online computer tools, accessible at any time or place, might offer students learning opportunities to develop sophisticated thinking skills. Furthermore, our results showed that a homework assignment in which the students take the responsibility for accomplishing it, can anyway contribute to positive effects on learning. Teachers should be aware of this.

In a closer analysis of students' work in the online environment we found that students applied various strategies ranging from trial-and-error to strategies that imply exploring the relations between the variables. As in the study on students' strategies in written problems, described in Chapter 2, trial-and-error also played a significant role in solving the problems in the online environment. Furthermore, in line with the results of the small-scale study described in Chapter 4, some students just tried out numbers without insight into the relations between the variables, while others applied solutions that captured the general relations. Nonetheless, trial-and-error is not a poor strategy per se. On the contrary, it is a strategy with high potential, because it can support students in grasping the relations between the variables and in general provide access into the solution of a problem.

Taking into account the exerted effort and the applied strategies we furthermore defined three levels of online working: free playing, just looking for answers and exploring relations and structure. We found that the level of online working had a significant effect on the gain in performance of the students in grades 4 and 6. In grade 4, the highest performance was found for the students who were exploring relations and structure, while in grade 6 this was the case for the students who were (just) looking for answers. In general, students who belonged to the level of free playing had the lowest posttest performance in all grades. Although the students' activity did not follow an overall trend in which free playing decreases and exploring relations and structure increases from grade 4, to grades 5 and 6, there was a small increase in the more advanced level of online working from grade 4 to 6. A problem with investigating trends in the levels of online working is that the observations did not come from a cohort of students, but was a cross-sectional sample from three grades.

In any case, the analysis showed that free playing did not support students in achieving a better performance. Rather, it was the students' working on a series of problems that might have led to positive outcomes. A crucial aspect of the series of problems is the variation in the numbers involved as well as in the given and unknown quantities. This variation elicits the use of a solution previously found to formulate a rule that describes the relation between the quantities. A prerequisite for generating a general solution though, is that students understand that the involved quantities vary together. The students in our study might have developed this understanding through the interactive dynamic character of the computer environment. While playing the game, students may comprehend that the arrows, the score and the game rule are related to each other so that a modification in the value of one of these variables has a direct effect on the other variables. In this environment students are continuously confronted, on a very natural way, with covarying variables.

To conclude, the interactive dynamic character of the computer environment and the variation in the series of problems are significant elements of the support students need in order to develop problem solving skills and mathematical reasoning necessary for dealing with early algebra problems.

2 Suggestions for further research

In our experiments we used only one type of early algebra tasks, namely contextual problems with interrelated values. Therefore, the operationalization of early algebra in this study was quite narrow. The implementation of other early algebra activities such as working with patterns, exploring functional relations and using graphs to represent change might bring new insights into primary school students' ability to reason algebraically. Possible tasks might be exploring the relation between the length and perimeter of a square or other geometrical figures and studying the relation between distance and time.

Besides broadening the range of early algebra tasks, the development of early algebraic reasoning could be also investigated in lower grades. Exploring patterns, studying part-whole relations, and analyzing relations between quantities might be appropriate activities for younger students. In addition, it would be interesting to investigate beginning secondary school students' solutions in early algebra problems. The examination of their approaches and reasoning would reveal whether these students prefer algebraic methods over arithmetic ones or vice versa, and how students' reasoning develops as they are introduced to formal algebra.

Another issue that needs further deliberation is the role of writing down the solution steps in problem solving. In particular, how does writing support students in solving problems and what kind of notations assist students in understanding a problem and guiding their solution process? With respect to the development of algebraic reasoning an issue that could be investigated is how students' notations could be further elaborated and developed into more formal ones.

The PhD study focused mainly on individual students' work on the computer outside the classroom. It is necessary, however, to address the complexities and challenges that arise when a technological tool is used in classroom. In particular, the role that the teacher can play to make the game activity effective and the interaction that the game stimulates between students should be further investigated.

In addition to studying the role of instruction given by the teacher, research on teacher beliefs and attitudes towards problem solving and early algebra is also necessary. Acquiring such knowledge not only can clarify aspects of students' performance, but is also a first step towards implementing problem solving and early algebra in the primary school curriculum.

3 How can arithmetic education in the Netherlands become more mathematical: A personal reflection to conclude

During these four years of intensive study of problem solving, some observations have made a great impression on me. These observations are related to students' written work, their dialogues and log files while solving problems on the computer, and conversations with teachers and other staff members at schools. In addition, an analysis of the textbook series for grade 4 gave me a good picture of the content and competencies that different textbooks endorse.

When students were confronted with non-routine problems, quite different reactions came to the fore. While some students were trying hard to get through to a solution and were dealing with the complex problem situations with relative confidence, others displayed a remarkable aversion to them. They refused to solve the problems and complained that they were very difficult and unfamiliar to them. Exclamations of satisfaction as soon as a student has cracked the problem he or she was working on contradicted with expression of discomfort and unwillingness to make or sustain an attempt to solve the problems.

Far too often students' reactions and engagement seemed to be in accord with their teachers' stance towards the problems. In the case were the teacher recognized problem solving as an important skill and conveyed the notion that by trying out several ideas the students might be able to find a solution, students were keener to accept the challenge of solving an unfamiliar problem. By talking with teachers I also realized that the more a teacher was feeling confident about his or her mathematical and didactical knowledge the more he or she was willing to accept the importance of solving non-routine problems. More often than not, teachers believed that these problems could only be solved by the more able students. According to them the problems were much too difficult for the less able students and could only cause them frustration.

While students' mathematical competence did play a significant role on their ability to solve the problems, we should not forget that most of the students do not encounter a single non-routine problem. The textbook analysis revealed that in most textbook series these problems have a marginal place. In general, the percentage of non-routine tasks in the textbooks is too low and these tasks are often to be found outside the main textbook in supplementary material meant for the more able students. In fact, the most widely used textbook series does not contain any non-routine task at all. No wonder then, that these problems are unfamiliar to students and are a source of serious difficulties. The same situation is true for the so-called gray-area tasks that provide students with opportunities to develop crucial problem solving skills and algebraic reasoning. Unfortunately, students rarely get this opportunity. Most of the time, they are dealing with tasks that simply require the use of algorithms.

My conclusion is that hardly any attention is given to problem solving and early algebraic reasoning in primary school mathematics education in the Netherlands. Van den Heuvel-Panhuizen (2008; 2009) has sketched a similar situation and argued that primary school arithmetic education should become more mathematical. This leads to the following question:

> *What should be done in order to bring about more mathematical reasoning in primary school mathematics education in the Netherlands?*

An answer to such a question is by no means straightforward. In my opinion, a change in primary school mathematics education in favor of problem solving and algebraic reasoning is a decision that should be taken at different educational levels. Nevertheless, I will try to give some suggestions on the basis of the observations that stemmed from my PhD study.

My first suggestion towards teachers is that they should be more bold in presenting their students with non-routine tasks. By going through the solution process together with their students, the students will come to see what it takes to solve a non-routine problem – among others, a good deliberation of the situation and many steps back and forth until a solution is found. More importantly, students will realize that a genuine problem cannot be solved within a few minutes and will build up confidence in their ability to solve complex problems. The positive impact of maintaining high expectations for students is enhanced as teachers encourage students to deal with cumbersome tasks. Of course, the teachers are the ones to decide when and how often non-routine tasks should be posed in their classrooms. They could first implement more accessible tasks (i.e., gray-area tasks) in order to lay a basis for handling more advanced problems and see how far they can go with their classroom. It goes without saying that such claims place also the responsibility on textbook authors to include more challenging tasks in the textbooks and on teacher educators to support student teachers and in-service teachers in implementing such tasks in their classroom.

A second suggestion to teachers is to encourage their students to write down their solution steps. Writing helps students monitor their solution process and work systematically towards a solution. However, students do not necessarily know why and what is appropriate to write down. For example, when I asked students to write down how they reached a solution, some of them provided lengthy descriptions of their though processes, while others just justified their solutions by checking the correctness of the result. Moreover, students who found a correct answer very quickly often replied that they *just* knew it was correct. These students did not just guess the answer; they indeed found a correct answer, but it was difficult for them to capture or retrieve it. Apparently, the articulation of one's thinking in writing is not a simple task, especially when one is engaged in a complex cognitive process like problem solving. This means that teachers should also be provided with guidelines about how to support students in developing this skill.

Another suggestion to teachers is based on the intriguing observation that trial-and-error was a strategy with a high potential to lead to a correct solution. However, in order for trial-and-error to be an efficient approach, students have to monitor their trials and use prior trials in order to inform subsequent ones. Teachers should, therefore, provide support in this direction.

Besides trial-and-error students were also able to apply strategies that imply a focus on structure, provided they had the opportunity to do so. This behavior brings me to another suggestion that concerns the utilization of technological tools in developing algebraic reasoning. My observations showed that a dynamic and interactive computer environment can support students in making generalizations. Such tools are also accessible at any time and any place allowing students to take control over their learning processed. My experience is that an online homework assignment with minimal guidance can influence

learning. Thus, designing and making available such tools to students and teachers should also be high on the agenda of software developers, task designers, and publishers.
Another suggestion to the designers of mathematics tasks is to develop tasks with systematic variation in the mathematical aspect that is to be emphasized. Based on my observations, I argue that variation can play a significant role in supporting students grasping the relations between variables in problems.

Besides what kind of tasks to include in the written curriculum, the assessed curriculum should also reflect the importance of mathematical reasoning and problem solving. Teaching, learning, and assessment are highly intertwined. Not only does assessment reflect what is considered important to learn in a particular educational system, but it often defines the content of learning as well. I recommend that problem solving becomes an explicit objective in the curriculum and assessment documents in primary school. Moreover, attention should be paid not only to the content but to the format of assessment as well. For example, ICT enables the presentation of more complex tasks, but the assessment format influences heavily the assessed competencies and the outcomes. This is made clear by comparing students' solutions in the online environment with their paper-and-pencil solutions. Therefore, the role and consequences of ICT-based assessment in formative and summative assessment are issues worth pursuing by the educational community.

The above mentioned suggestions involve all educational stakeholders. Teachers, parents, textbook authors, publishers, developers, mathematics educators, and policy makers should share the same vision: Do justice to the fundamental right of students to develop their mathematical reasoning. It is the task of mathematics education researchers to (co)develop appropriate tools for realizing this aim.

References

Johanning, D.I. (2004). Supporting the development of algebraic thinking in middle school: A closer look at students' informal strategies. *Journal of Mathematical Behavior, 23*, 371–388.

Kolovou, A., van den Heuvel-Panhuizen, & Elia, I (2009). An online ICT environment to support and assess primary school students' solving of non-routine puzzle-like mathematical word problems. In M. Tzekaki, M. Kaldrimidou, & H. Sakonidis (Eds.), *Proceedings of the 33rd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 386–392). Thessaloniki, Greece: PME.

Pugalee, D.K. (2004). A Comparison of verbal and written descriptions of students' problem solving processes. *Educational Studies in Mathematics, 55*, 27–47.

Stacey, K. & MacGregor, M. (2000). Learning the algebraic method of solving problems. *Journal of Mathematical Behavior, 18*, 149–167.

Van den Heuvel Panhuizen, M. (2008, January). *Hoe wiskundig is ons rekenonderwijs op de basisschool?* [How mathematical is our arithmetic education in primary school?]. Plenary lecture at the 26th Panama conference, Noordwijkerhout, the Netherlands.

Van den Heuvel Panhuizen, M. (2009). *Hoe rekent Nederland?* [How does the Netherlands calculate?]. Inaugural lecture. Utrecht: Universiteit Utrecht, Fisme.

Van den Heuvel-Panhuizen, M., & Bodin-Baarends, C. (2004). All or nothing: Problem solving by high achievers in mathematics. *Journal of the Korea Society of Mathematical Education, 8*(3), 115–121.

Van den Heuvel-Panhuizen, M., Kolovou, A., Peltenburg, M. (2011). Using ICT to improve assessment. In B. Kaur, & W.K. Yoong (Eds.), *Assessment in the mathematics classroom: Yearbook 2011, Association of Mathematics Educators* (pp. 165–185). Singapore: World Scientific and AME.

Williams, K. (2003). Writing about the problem-solving process to improve problem-solving performance. *The Mathematics Teacher, 96*(3), 185–187.

Summary

Samenvatting

Acknowledgements

Curriculum vitae

Overview of publications

Overview of presentations

FIsme Scientific Library

ICO Dissertation Series

Summary

The aim of this thesis was to provide insight into Dutch primary school students' competence in solving non-routine number problems. In contrast to routine problems, such problems cannot be solved by just applying a standard operation. An example of such a puzzle-like number problem is the following. *In a quiz you get two points for each correct answer. If a question is not answered or the answer is wrong, one point is subtracted from your score. The quiz contains ten questions. Tina received eight points in total. How many questions did Tina answer correctly?* The complexity of this problem is not due to the size of the numbers and the calculations that have to be carried out. What makes this problem perplexing for primary school students is that they have to understand the interrelatedness of the numbers in the problem. Introducing such problems in primary school is important for two reasons. These problems offer students opportunities to develop problem solving and reasoning skills with numbers and they provide them with entry points for learning algebra. Therefore, these problems can be called early algebra problems. In our study these terms are used interchangeably.

The study addressed three research questions:

1. *How able are Dutch primary school students in solving non-routine puzzle-like problems?*
2. *To what degree do textbooks offer students opportunities to learn solving non-routine puzzle-like problems?*
3. *How can we support students in solving non-routine puzzle-like problems?*

To answer these research questions several studies were carried out, which are reported in chapters 2 to 7 of this thesis. First, we investigated the ability of high-achieving students to solve early algebra problems as well as the strategies that students applied. Next, we examined the opportunities that students get to learn to solve these problems within the current primary school curriculum. Finally, we developed an ICT environment for solving these problems. Students' activity in this environment was analyzed and related to their problem solving performance. Both qualitative and quantitative analyses were carried out in these studies.

Chapter 2 reports on the study that investigated strategy use and strategy flexibility and their relations with performance in early algebra problems. The study was based on three paper-and-pencil problems and examined the solutions of 152 high-achieving students in grade 4. Two types of strategy flexibility were investigated, namely inter-task flexibility (i.e., changing strategies across problems) and intra-task flexibility (i.e., changing strategies within problems). The findings showed that students did not always write down their strategy. In the cases in which we could identify the use of a particular strategy, we found that trial-and-error had the highest potential to lead to a correct solution. The two types of

flexibility were seldom applied, yet students who showed inter-task strategy flexibility performed better than those who persevered with the same strategy across problems. Moreover, contrary to our expectations, intra-task strategy flexibility did not support students in providing the correct answer. However, we found that this result was due rather to the construction of an incomplete model of the problem situation than to the use of intra-task strategy flexibility.

To shed light on the students' low performance we investigated the opportunities that students get to practice solving non-routine puzzle-like problems in classroom. Although several factors influence instruction, there is much evidence that the curriculum and the textbooks determine to a large extent what teachers teach and, consequently, what students learn. *Chapter* 3 reports on a study in which the nature of tasks in six mathematics textbook series for grade 4 was analyzed. The analytical framework that we used was designed through an iterative process of classifying tasks according to their complexity. The final version of the framework contained three main categories of tasks: straightforward tasks, gray-area tasks, and non-routine puzzle-like tasks. Straightforward tasks require accurately carrying out arithmetical calculations, while in non-routine puzzle-like tasks students have to focus on the relations between the numbers or quantities included in the task. Gray-area tasks may prompt students to do investigations and develop problem solving strategies. For example, finding all possible combinations of given numbers and searching for a pattern in a series of numbers. The analysis of the tasks in the six textbook series revealed that challenging non-routine tasks are rare and are mostly found in the additional materials that accompany the textbook series. In some textbook series these non-routine tasks are completely absent. The results of our study raised questions and concerns about the mathematical nature of arithmetic education in Dutch primary schools.

Subsequently, we investigated ways to support students in dealing with early algebra problems. For this aim we provided them with an environment for experiencing interdependency of values in a problem situation. A dynamic interactive computer game called Hit the target was developed for this purpose. In this environment the students can observe how the game score varies as they manipulate the values of the hits, the misses, or the game rule. *Chapter* 4 reports on a pilot study that employed the ICT environment both as a tool to support students' learning by offering them opportunities to experiment, produce and reflect on their solutions, and as a tool to monitor and assess students' problem solving processes. In total, 24 high achieving students from grade 4 took part in this study. The students of the experimental group carried out a number of assignments in the ICT environment in pairs. The analysis of their dialogues and actions provided us with a detailed account of their problem solving processes. Students demonstrated different levels of problem solving activity and some interesting processes were revealed, for example the so-called bouncing effect. This means that the students first come up with a correct general solution that captures the relations between the variables, but nevertheless later revert to an incorrect solution. The data of a written test on problem solving collected before and after this short treatment did not form a sufficient basis for drawing conclusions about the potential of the dynamic interactive computer game to improve students' problem solving performance.

Therefore we conducted a large-scale study with a pretest-posttest control group design with 785 students from grades 4, 5, and 6. The students of the experimental group were asked to play the game at home online and to use it for solving a series of problems. The students' online computer activity was registered with special software. Before and after the intervention all students' problem solving performances were measured with a paper-and-pencil test.

Chapter 5 describes the study in which we explored the role of feedback that was generated by the game when solving the series of problems. This feedback neither diagnoses the students' solution processes, nor indicates whether their answers are correct or incorrect. Instead, it bears information that students have to interpret in order to evaluate their answers themselves. The analysis of the data for grade 6 showed that the game-generated feedback supported students in detecting and correcting their errors. The students performed better in the problems they had to solve in the online environment than in the paper-and-pencil test where no feedback was available. Furthermore, in the posttest their performance and the frequency of answer verification were higher than in the pretest. The last finding suggests that game-generated feedback stimulated student-generated feedback.

Chapter 6 describes the study in which we investigated the differences in performance on early algebra problems between the experimental group and the control group. Statistical analyses on the data of 236 sixth graders showed a significant positive effect of the intervention on the posttest performance. Furthermore, we neither detected a significant effect of the amount of effort put into the online work, nor of the sex of the students on performance. However, effort turned out to play a complex role in this study; although girls put more effort into the online work than boys did, their gains in achievement were the same as for boys. This finding might imply that boys are more efficient users of computer tools.

Chapter 7 focuses on the strategies that the students in grades 4, 5, and 6 applied when working in the online environment. The analysis revealed that the learning environment stimulated the application of various strategies and triggered the exploration of the relations between the variables. Moreover, the online work was positively related with success in the paper-and-pencil posttest and the gain between the pre and posttest was the highest in grade 6. The study indicated that the interactive dynamic character of the computer environment and the systematic variation of the values in the series of problems had a significant role in supporting students' reasoning when dealing with early algebra problems.

Chapter 8 gives an overview of the findings and provides suggestions for further research and educational practice. Conclusions that can be drawn from this thesis are that Dutch primary students have difficulties in solving non-routine puzzle-like problems and that the tasks included in Dutch textbook series are an indication that hardly any attention is given to problems that introduce students to algebraic reasoning. This situation calls for a concerted effort from all educational stakeholders in order to include more mathematics in primary school arithmetic education in the Netherlands. This thesis has provided ideas for changing this situation. It has shown that computer tools can enable students to explore concepts related to dynamic processes like covariation and thus can open up new possibilities for familiarizing primary school students with early algebra problems.

Samenvatting

Het onderzoek dat in het kader van dit proefschrift is uitgevoerd, had als doel inzicht te verwerven in de vaardigheid van Nederlandse basisschoolleerlingen in het oplossen van niet-routinematige, puzzelachtige rekenopgaven. Kenmerkend voor deze opgaven is dat de leerlingen voor het oplossen ervan geen standaardprocedure beschikbaar hebben. De volgende opgave kan als een voorbeeld van dergelijke opgaven worden gezien. *In een quiz krijg je twee punten voor een goed antwoord. Geef je geen antwoord of is het antwoord fout, dan wordt er een punt van je score afgehaald. Tina heeft na tien vragen acht punten. Hoeveel vragen heeft Tina goed beantwoord?* De complexiteit van deze opgave ligt niet in de grootte van de getallen of in de berekeningen die uitgevoerd moeten worden, maar de opgave is ingewikkeld omdat de leerlingen rekening moeten houden met de onderlinge relaties tussen de getallen. De invoering van dit soort opgaven in de basisschool is om twee redenen belangrijk. Ten eerste worden op die manier de ontwikkeling van de probleemoplossingsvaardigheden en het kunnen redeneren met getallen bevorderd. Ten tweede worden leerlingen door niet-routinematige, puzzelachtige rekenopgaven voorbereid op het leren van algebra. Daarom worden deze niet-routinematige, puzzelachtige rekenopgaven ook wel *early algebra* opgaven genoemd.

In het dissertatieonderzoek komen drie onderzoeksvragen aan de orde:

1. *Hoe goed zijn Nederlandse basisschoolleerlingen in probleem oplosssen, d.w.z. in het oplossen van niet-routinematige, puzzelacthige rekenopgaven?*
2. *In hoeverre bieden reken-wiskundemethodes aan leerlingen de mogelijkheid om niet-routinematige, puzzelachtige rekenopgaven te leren maken?*
3. *Wat is een goede manier om leerlingen niet-routinematige, puzzelachtige rekenopgaven te leren maken?*

Deze onderzoeksvragen zijn beantwoord in een aantal deelstudies die achtereenvolgens beschreven zijn in de hoofdstukken 2 tot 7 van dit proefschrift. Het eerste onderzoek richtte zich op de vaardigheid van goede rekenaars in het oplossen van early algera opgaven en de strategieën die ze hierbij hebben gebruikt. Vervolgens hebben wij gekeken naar de mate waarin deze niet-routinematige, puzzelachtige rekenopgaven voorkomen in de reken-wiskundemethodes. Ten slotte is binnen het project een ICT-omgeving ontwikkeld voor het oplossen van dergelijke opgaven. Het werken van de leerlingen in deze omgeving is geanalyseerd en gerelateerd aan hun vaardigheid om niet-routinematige, puzzelachtige rekenopgaven op te lossen. Hierbij zijn zowel kwalitatieve als kwantitatieve analyses toegepast.

Hoofdstuk 2 vormt het verslag van het onderzoek naar het verband tussen strategiegebruik, strategieflexibiliteit en de prestaties bij de *early algebra* opgaven. Het onderzoek is gebaseerd op drie opgaven in een schriftelijk toets. In totaal zijn de oplossingsstrategieën van 152 goede rekenaars uit groep 6 geanalyseerd. Hierbij is onder andere gekeken naar twee soorten strategie-flexibiliteit: inter-flexibiliteit (d.w.z., bij de ene opgave een andere strategie toepassen dan bij de andere) en intra-flexibiliteit (d.w.z., het wisselen van strategie binnen een opgave). De resultaten lieten op de eerste plaats zien dat de leerlingen niet altijd hun strategieën noteerden. Onder de strategieën die wij wel konden identificeren, leidde de trial-and-error strategie het vaakst tot een correcte oplossing. De twee soorten flexibiliteit werden niet vaak toegepast, maar de leerlingen die bij de ene opgave een andere strategie toepasten dan bij de andere, presteerden hoger dan degenen die aan eenzelfde strategie vasthielden. Tegen onze verwachtingen in, hielp intra-flexibiliteit niet bij het vinden van een juist antwoord. Dit had echter eerder te maken met het feit dat de leerlingen zich geen volledig beeld van de probleemsituatie hadden gevormd dan met het wisselen van strategie binnen een opgave.

Om de lage prestaties van Nederlandse basisschoolleerlingen beter te begrijpen, hebben we daarna onderzocht in hoeverre de leerlingen ook werkelijk de kans krijgen om op school ervaring op te doen met niet-routinematige, puzzelachtige rekenopgaven. Hoewel meerdere factoren van invloed zijn op het gegeven onderwijs, zijn er voldoende aanwijzingen dat het curriculum en de gebruikte reken-wiskundemethode in hoge mate bepalen wat er wordt onderwezen en wat de kinderen vervolgens leren. Daarom hebben we een methodeanalyse uitgevoerd. In *hoofdstuk* 3 wordt hierover gerapporteerd. In deze methodeanalyse is de aard onderzocht van alle opgaven die te vinden zijn in zes reken-wiskundemethodes voor groep 6. Het raamwerk dat hiervoor is gebruikt, is in een iteratief proces ontwikkeld. De opgaven zijn hierbij geclassificeerd op basis van hun complexiteit. De eindversie van het raamwerk bevatte drie hoofdcategorieën: recht-toe-recht-aan opgaven, grijsgebied opgaven, en niet-routinematige, puzzelachtige opgaven. Recht-toe-recht-aan opgaven vereisen het nauwkeurig uitvoeren van berekeningen, terwijl in niet-routinematige, puzzelachtige opgaven de leerlingen eerst zicht moeten krijgen op de relaties tussen de getallen. Grijsgebied opgaven liggen tussen de andere twee typen opgaven in. Ze kunnen de leerlingen ertoe aanzetten verschillende oplossingen te proberen en oplossingsstrategieën te ontwikkelen. Voorbeelden daarvan zijn opgaven waarbij de leerlingen alle mogelijke combinaties van bepaalde gegeven getallen moeten vinden of patronen moeten zoeken in reeksen getallen. De resultaten van onze methodeanalyse lieten zien dat het aantal niet-routinematige, puzzelachtige opgaven in de methodes zeer klein is en dat zij bovendien vaker in het aanvullend materiaal van een methode te vinden zijn. In sommige methodes waren deze opgaven totaal afwezig. De resultaten van deze analyse roepen vragen en ook zorgen op over het wiskundige karakter van het Nederlandse rekenonderwijs.

Als volgende stap in het onderzoek zijn we nagegaan op welke manier leerlingen ondersteund kunnen worden bij het leren oplossen van *early algebra* opgaven. Voor dit doel hebben wij de leerlingen een omgeving verschaft, waarin zij ervaring kunnen opdoen met probleemsituaties waarin getalwaardes met elkaar samenhangen en samen variëren.

Deze omgeving bestond uit een dynamisch, interactief pijl-en-boogspelletje op de computer, getiteld Hit the target. In deze computeromgeving kunnen de leerlingen ondervinden hoe de behaalde spelscore kan variëren door het aantal rake en gemiste pijlen en de spelregel te variëren.

Hoofdstuk 4 rapporteert over de pilotstudie waarin een ICT-omgeving op twee manieren is gebruikt; enerzijds als hulpmiddel bij het leren van probleem oplossen – leerlingen kunnen het spel spelen en zo experimenteren en zelf oplossingen bedenken en over hun oplossingen reflecteren – en anderzijds als hulpmiddel om het oplossingsproces van de leerlingen te volgen en op deze manier te meten. In totaal namen 24 goede rekenaars uit groep 6 aan dit onderzoek deel. De leerlingen in de experimentele groep hebben in tweetallen in de ICT-omgeving het spelletje gespeeld en als hulp gebruikt bij het oplossen van een paar opgaven. De analyse van de dialogen en de handelingen van de leerlingen gaf een gedetailleerd beeld van hun oplossingsprocessen. Er werd op verschillende niveaus gewerkt en enkele interessante kenmerken van hun manier van werken zijn zichtbaar gemaakt. Een voorbeeld hiervan is het zogenaamde *bouncing effect*. Dit houdt in dat de leerlingen eerst met een correcte oplossing komen gebaseerd op een algemene regel en bij een volgende opgave toch weer een onjuiste oplossing laten zien. Door het kleine aantal leerlingen dat heeft meegedaan aan deze deelstudie konden we echter geen conclusies trekken over de bijdrage van het computerspelletje aan de prestaties van de leerlingen.

Om meer robuuste resultaten te verkrijgen, is vervolgens een grootschalig onderzoek uitgevoerd bij 785 leerlingen uit groepen 6, 7, en 8. Dit onderzoek is opgezet volgens een voortoets-natoets-controle-groep design. Aan de leerlingen van de experimentele groep is gevraagd het spelletje online thuis te spelen en met behulp van het spelletje een reeks opgaven op te lossen. Speciale software zorgde ervoor dat de online-activiteiten van de leerlingen tijdens het spelen van het spelletje werden geregistreerd. Verder werden de probleemoplossingsvaardigheden van alle leerlingen met een schriftelijke voor- en natoets gemeten.

Hoofdstuk 5 beschrijft de deelstudie waarin de rol van feedback is onderzocht. Dit betreft de feedback die het computerspelletje genereert. Als de leerlingen bepaalde oplossingen invoeren in het spelletje geeft het informatie die de leerlingen zelf moeten interpreteren om te kunnen beslissen of hun antwoorden goed zijn. Deze feedback levert de leerlingen dus geen diagnose over hun wijze van oplossen en geeft ook niet aan of hun antwoorden juist of onjuist zijn, maar laat de leerlingen alleen de consequenties zien van de waarden die ze hebben ingevoerd. De analyse van de gegevens uit groep 6 bracht aan het licht dat de door het spelletje gegenereerde feedback de leerlingen inderdaad heeft ondersteund bij het ontdekken en corrigeren van hun fouten. Daarnaast hebben we ook gevonden dat de leerlingen beter presteerden op de opgaven tijdens het online werken dan bij de schriftelijke toets waar geen feedback beschikbaar was. Bovendien waren er ook verschillen tussen de scores op de voortoets en de natoets. Bij de natoets scoorden de leerlingen hoger en controleerden zij hun antwoorden vaker. Dit laatste resultaat suggereert dat de door het spelletje gegenereerde feedback ertoe kan bijdragen dat de leerlingen ook zelf feedback gaan genereren, bijvoorbeeld door hun antwoorden te controleren.

Hoofdstuk 6 behandelt de deelstudie waarin de verschillen in prestaties tussen de experimentele en de controlegroep zijn onderzocht. Statistische analyses van de data van 236 leerlingen uit groep 6 lieten een significant positief effect zien van de interventie op de scores op de natoets. Er is echter geen significant effect gevonden van de manier waarop de leerlingen in de online-omgeving gewerkt hebben. Meer inspanning (langer met het spelletje bezig zijn, meer handelingen verrichten) leidde niet tot een hogere score. Ook het geslacht van de leerlingen bleek geen direct effect te hebben op de prestaties. Maar we vonden wel, dat de geleverde inspanning een complexe rol in dit geheel speelde. Ofschoon de jongens minder inspanning leverden bij het online werken dan de meisjes, bleken de jongens toch evenveel profijt van het spelletje te hebben gehad als de meisjes. Dit zou kunnen impliceren dat jongens efficiëntere gebruikers van computerprogramma's zijn.

Hoofdstuk 7 richt zich op de oplossingsstrategieën die leerlingen uit groepen 6, 7 en 8 in de computeromgeving hebben toegepast. Uit de analyse bleek dat de online-omgeving de toepassing van diverse strategieën en het verkennen van relaties tussen de variabelen bevorderde. Bovendien hebben we een significante relatie vastgesteld tussen het online werken en de mate van succes op de schriftelijke toets. De grootste toename in de toetsscores vonden we in groep 6. Deze bevindingen suggereren dat een dynamisch interactief computerspelletje in combinatie met een serie opgaven waarin de waardes systematisch zijn gevarieerd, de leerlingen tot steun kunnen zijn bij het leren oplossen van *early algebra* opgaven.

Hoofdstuk 8 geeft een overzicht van onze bevindingen en komt met suggesties voor vervolgonderzoek en de onderwijspraktijk. Dit dissertatieonderzoek heeft laten zien dat Nederlandse basisschoolleerlingen moeite hebben met het oplossen van niet-routinematige, puzzelachtige opgaven en dat er tegelijkertijd in het Nederlandse reken-wiskundeonderwijs waarschijnlijk weinig aandacht is voor opgaven die leerlingen voorbereiden op algebraïsch denken. Dit is een situatie die zorgen baart en die vraagt om een gezamenlijke actie van alle bij het onderwijs betrokken partijen om hier iets aan te doen door meer wiskunde in het Nederlandse rekenonderwijs op de basisschool te brengen. Het onderzoek dat in het kader van dit proefschrift is uitgevoerd, heeft hiervoor ideeën aangeleverd. Het heeft laten zien dat leerlingen met behulp van ICT-tools ervaring kunnen opdoen in het exploreren van dynamische concepten zoals covariatie en dat op deze manier de weg geopend kan worden voor *early algebra* op de basisschool.

Acknowledgements

The completion of this thesis signifies the end of a long journey, literally and metaphorically; at least longer than I expected when I first came to the Netherlands seven years ago to teach Greek to students of Greek parentage. Actually, one of my motives to come to the Netherlands was to get involved in the work of the Freudenthal Institute of which I had heard so much while studying at the University of Athens.

During my first visit at the Institute I was introduced to my future supervisor Marja van den Heuvel-Panhuizen. Her openness was overwhelming. In no time I got the opportunity to contribute to a project on problem solving and after a year I got a paid educational leave from the Greek ministry of Education and I was able to start a full-time PhD trajectory. A dream was actually coming true.

Throughout these years there had been moments of joy and moments of disappointment: *It's all in the game*. A very important thing that I have learned though is not to give up and keep on trying improving myself. Also, that hard work can get you far. Nothing can be closer to truth when you work with Marja. I thank her especially for the expertise and energy she has put in this thesis and for the precious moments of collaboration and inspiration. During my PhD studies I had also the luck to collaborate with Iliada Elia from Cyprus and Olaf Köller, Hendrik Winkelmann, and Alexander Robitzsch from Germany. Their input was very important for the completion of this thesis. I would also like to thank my colleague Marjolijn Peltenburg for her support and warmth; we have been together in plenty of seminars and conferences and we have collaborated and exchanged views extensively. Also, my colleague and roommate Adri Dierdorp and my colleagues at the Institute who showed interest in my research and encouraged me to go on. Especially, I am grateful to Nathalie Kuijpers and Betty Heijman for helping me out with turning my manuscript into a book. Last but not least, I am grateful for the opportunities I got while working at the Institute to participate in international conferences and several master classes. The teachers and students at St. Dominicusschool, De Baanbreker, De Beiaard, W.G. van de Hulstschool, Julianaschool, De Koekoek, Marcusschool, De Spits, Parkwijck, Paulusschool, De Schakel and Waterrijk, also deserve my gratitude. Without them this research would not have been possible.

Finally, I would like to thank the ones who have supported me in many significant ways. Some of them were not physically present, but were always in my mind and in my heart: My parents who value education very highly and showed trust; my sister and my nephew for their unconditional love; my dearest friends Ioanna and Apostolos who have been always there for me; and Nondas who would have been very proud of me if he was here. I also want to thank my Dutch friends Marco, Annelies, and Anton. They have helped me a

lot to cope with living in a foreign country. Also, my Greek (and non-Greek) friends in the Netherlands who made things easier when times were rough. En Petra, dank je wel!

I am sure by now: Ithaca has not deceived me; it has given me a beautiful journey. I am off for the next one!

Curriculum vitae

Angeliki Kolovou was born in May 30, 1974 in Athens, Greece. She got her high school diploma in 1992 from the public Lyseum of Kareas in Athens. In 1993, after state examinations, she was accepted at the Department of Primary Education of the National Kapodistrian University of Athens. Four years later she got her bachelor degree and continued with the Master's program *Mathematics in Primary Education* at the same university department under the supervision of prof. Ada Boufi. In 2001 she got her Master's degree and participated in the state examinations for a post as a primary school teacher. From 2001 to 2004 she worked in primary schools in Athens. In 2004 she was seconded to the department of Greek language in Utrecht, the Netherlands where she taught Greek as second language to primary school students until 2007. In parallel, from 2006 to 2007 she participated in the Problem Solving in Primary School (POPO) project at the Freudenthal Institute for Science and Mathematics Education of Utrecht University. After getting a paid educational leave from the Greek Ministry of Education she became a full-time PhD student in the POPO project. Her supervisor within this project was prof. dr. Marja van den Heuvel-Panhuizen. Her PhD study was accepted by the prestigious Interuniversity Centre for Educational Research (ICO) in the Netherlands for which she fulfilled all requirements to be an ICO PhD member.

Overview of publications

Kolovou, A., Van den Heuvel-Panhuizen, M., &. Robitzsch A. (submitted). An online game as a learning environment for early algebraic problem solving by students in upper primary school.

Kolovou A., Van den Heuvel-Panhuizen M., & Köller O. (submitted). Can an intervention including an online game to be played at home contribute to students' early algebra performance?

Kolovou, A., Van den Heuvel-Panhuizen, M. (accepted). An online game as a learning environment for early algebraic problem solving by upper primary school students. *Proceedings of the CERME 7 conference*.

Van den Heuvel-Panhuizen, M., Kolovou, A., Peltenburg, M. (2011). Using ICT to improve assessment. In B. Kaur, & W.K. Yoong (Eds.), *Assessment in the mathematics classroom: Yearbook 2011, Association of Mathematics Educators* (pp. 165–185). Singapore: World Scientific and AME.

Kolovou, A. & Van den Heuvel-Panhuizen M. (2011). Ένα διαδικτυακό παιχνίδι ως περιβάλλον επίλυσης αλγεβρικών προβλημάτων από μαθητές των τριών τελευταίων τάξεων του Δημοτικού [An online game as a learning environment for early algebraic problem solving by upper primary school students]. In Ch. Th. Panagiotakopoulos (Ed.), Proceedings of the 2nd Panhellenic conference on the integration and use of ICT in the educational process (pp. 363–372). Patras, Greece: CETL–HICT.

Kolovou, A., & Van den Heuvel-Panhuizen, M. (2010). Online game-generated feedback as a way to support early algebraic reasoning. *International J. Continuing Engineering Education and Life-Long Learning, 20*(2), 224–238.

Kolovou, A., Van den Heuvel-Panhuizen, M., & Elia, I. (2009). An online ICT environment to support and assess primary school students' solving of non-routine puzzle-like mathematical word problems. In Tzekaki, M., Kaldrimidou, M. & Sakonidis, H. (Eds.), Proceedings of the 33rd Conference of the International Group for the Psychology of Mathematics Education, (Vol. 3, pp 385–392). Thessaloniki, Greece: PME.

Kolovou, A., Van den Heuvel-Panhuizen, M., & Bakker, A. (2009). Non-routine problem solving tasks in primary school mathematics textbooks-A needle in a haystack. *Mediterranean Journal for Research Mathematics Education 8*(2), 29–66.

Kolovou, A. & Van den Heuvel-Panhuizen, M. (2009). Hoeveel probleemoplossingsopgaven zitten er in onze reken-wiskunde methoden? In M. van Zanten (Ed.) *Doorgaande ontwikkelingen rekenen-wiskunde, verslag van de 26ste Panama-conferentie* (pp. 101–106). Utrecht: Fisme, Utrecht University.

Van den Heuvel-Panhuizen, M., & Kolovou, A. (2009). Een ICT-omgeving om het oplossen van puzzelachtige verhaaltjessommen te toetsen en te leren. *Volgens Bartjens 28*(3), 22–26.

Elia, I., Van den Heuvel-Panhuizen, M., & Kolovou, A. (2009). Exploring strategy use and strategy flexibility in non-routine problem solving by primary school high achievers in mathematics. *ZDM The International Journal on Mathematics Education*, *41*, 605–618.

Kolovou, A., van den Heuvel-Panhuizen, M., Bakker, A, & Elia, I. (2008). *An ICT environment to assess and support students' mathematical problem-solving performance in non-routine puzzle-like word problems*. In A. Gagatsis (Ed.), Research in Mathematics Education (pp. 175–190). Nicosia, Cyprus: University of Cyprus.

Overview of presentations

Kolovou, A. & Van den Heuvel-Panhuizen M. (2011). Ένα διαδικτυακό παιχνίδι ως περιβάλλον επίλυσης αλγεβρικών προβλημάτων από μαθητές των τριών τελευταίων τάξεων του Δημοτικού [An online game as a learning environment for early algebraic problem solving by upper primary school students]. Paper presentation at the 2nd Panhellenic conference on the integration and use of ICT in the educational process (Patras, Greece, April 28 – 30, 2011).

Kolovou, A. & Van den Heuvel-Panhuizen M. (2011). An online game as a learning environment for early algebraic problem solving by upper primary school students. Paper presentation at the CERME 7 conference (Rzeszów, Poland, February 9 – 13, 2011).

Kolovou, A. & Van den Heuvel-Panhuizen M. (2010). Can an intervention including an online game to be played at home contribute to students' early algebra performance? Paper presented at the ICO Toogdag research meeting (Amstelveen, the Netherlands, November 4, 2010).

Kolovou, A. & Van den Heuvel-Panhuizen M. (2010). Η ανατροφοδότηση από ένα διαδικτυακό παιχνίδι ως μέσο στήριξης της πρώιμης αλγεβρικής σκέψης [Online game-inherent feedback as a way to support early algebraic reasoning]. Paper presentation at the 2nd Panhellenic Conference of Educational Sciences (Athens, Greece, May 27 – 30, 2010).

Kolovou, A., Van den Heuvel-Panhuizen M., & Elia I. (2009). An online ICT environment to support and assess primary school students' solving of non-routine puzzle-like mathematical word problems. Round table presentation at the EARLI 2009 conference (Amsterdam, the Netherlands, August 25 – 29, 2009).

Kolovou, A., Van den Heuvel-Panhuizen M., & Elia I. (2009). An ICT environment to assess and support students' mathematical problem-solving performance in non-routine puzzle-like word problems. Paper presentation at the JURE 2009 conference (Amsterdam, the Netherlands, August 24 – 25, 2009).

Kolovou, A., Van den Heuvel-Panhuizen M., & Elia I. (2009). An online ICT environment to support and assess primary school students' solving of non-routine puzzle-like mathematical word problems. Research report presentation at the PME 33 conference (Thessaloniki, Greece, July 19 – 24, 2009).

Kolovou, A., Van den Heuvel-Panhuizen M., & Elia I. (2009). An ICT environment to support and access students' problem-solving processes and performance in non-routine number puzzle-like word problems. Poster presentation at the CERME 6 conference (Lyon, France, January 28 – February 1, 2009).

Kolovou, A., Van den Heuvel-Panhuizen M., & Elia I. (2009). An ICT environment to support and access students' problem-solving processes and performance in non-routine number puzzle-like word problems. Poster presentation at the Expertmeeting Rekenen (Utrecht, the Netherlands, November 6, 2009).

Kolovou, A., & Van den Heuvel-Panhuizen M. (2008). Hoeveel probleemoplossen zit er in onze reken -wiskundemethoden? Workshop at the 26ste Panama conference (Noordwijkerhout, the Netherlands, January 23 – 25, 2008).

Kolovou, A., Van den Heuvel-Panhuizen M., & Elia I. (2009). An ICT environment to assess and support students' mathematical problem-solving performance in non-routine puzzle-like word problems. Paper presentation at ISDDE 2008 conference (Egmond aan Zee, the Netherlands, June 29 – July 2, 2008).

Kolovou, A., Van den Heuvel-Panhuizen M., Bakker, A., & Elia I. (2009). An ICT environment to assess and support students' mathematical problem-solving performance in non-routine puzzle-like word problems. Paper presentation at ICME11 conference (Monterey, Mexico, July 6 – 13, 2008).

FIsme Scientific Library (formerly published as CD-β Scientific Library)

66. Kolovou, A. (2011). *Mathematical problem solving in primary school*.
65. Meijer, M. R. (2011). *Macro-meso-micro thinking with structure-property relations for chemistry. An explorative design-based study*.
64. Kortland, J. & Klaassen, C. J. W. M. (2010). *Designing theory-based teaching-learning sequences for science. Proceedings of the symposium in honour of Piet Lijnse at the time of his retirement as professor of Physics Didactics at Utrecht University*.
63. Prins, G. T. (2010).*Teaching and learning of modelling in chemistry education. Authentic practices as contexts for learning*.
62. Boerwinkel, D. J. & Waarlo, A. J. (2010). *Rethinking science curricula in the genomics era. Proceedings of an invitational workshop*.
61. Ormel, B. J. B. (2010). *Het natuurwetenschappelijk modelleren van dynamische systemen. Naar een didactiek voor het voortgezet onderwijs*.
60. Hammann, M., Waarlo, A. J., & Boersma, K. Th. (Eds.) (2010). *The nature of research in biological education: Old and new perspectives on theoretical and methodological issues – A selection of papers presented at the VIIth Conference of European Researchers in Didactics of Biology*.
59. Van Nes, F. (2009). *Young children's spatial structuring ability and emerging number sense*.
58. Engelbarts, M. (2009*). Op weg naar een didactiek voor natuurkunde-experimenten op afstand. Ontwerp en evaluatie van een via internet uitvoerbaar experiment voor leerlingen uit het voortgezet onderwijs*.
57. Buijs, K. (2008). *Leren vermenigvuldigen met meercijferige getallen*.
56. Westra, R. H. V. (2008). *Learning and teaching ecosystem behaviour in secondary education: Systems thinking and modelling in authentic practices*.
55. Hovinga, D. (2007). *Ont-dekken en toe-dekken: Leren over de veelvormige relatie van mensen met natuur in NME-leertrajecten duurzame ontwikkeling*.
54. Westra, A. S. (2006). *A new approach to teaching and learning mechanics*.
53. Van Berkel, B. (2005). *The structure of school chemistry: A quest for conditions for escape*.
52. Westbroek, H. B. (2005). *Characteristics of meaningful chemistry education: The case of water quality*.
51. Doorman, L. M. (2005). *Modelling motion: from trace graphs to instantaneous change*.
50. Bakker, A. (2004). *Design research in statistics education: on symbolizing and computer tools*.
49. Verhoeff, R. P. (2003). *Towards systems thinking in cell biology education*.

48. Drijvers, P. (2003). *Learning algebra in a computer algebra environment. Design research on the understanding of the concept of parameter*.
47. Van den Boer, C. (2003). *Een zoektocht naar verklaringen voor achterblijvende prestaties van allochtone leerlingen in het wiskundeonderwijs*.
46. Boerwinkel, D.J. (2003). *Het vormfunctieperspectief als leerdoel van natuuronderwijs. Leren kijken door de ontwerpersbril.*
45. Keijzer, R. (2003). *Teaching formal mathematics in primary education. Fraction learning as mathematising process*.
44. Smits, Th. J. M. (2003). *Werken aan kwaliteitsverbetering van leerlingonderzoek: Een studie naar de ontwikkeling en het resultaat van een scholing voor docenten*.
43. Knippels, M. C. P. J. (2002). *Coping with the abstract and complex nature of genetics in biology education – The yo-yo learning and teaching strategy*.
42. Dressler, M. (2002). *Education in Israel on collaborative management of shared water resources.*
41. Van Amerom, B.A. (2002). *Reinvention of early algebra: Developmental research on the transition from arithmetic to algebra*.
40. Van Groenestijn, M. (2002). *A gateway to numeracy. A study of numeracy in adult basic education.*
39. Menne, J. J. M. (2001). *Met sprongen vooruit: een productief oefenprogramma voor zwakke rekenaars in het getallengebied tot 100 – een onderwijsexperiment*.
38. De Jong, O., Savelsbergh, E.R. & Alblas, A. (2001). *Teaching for scientific literacy: context, competency, and curriculum.*
37. Kortland, J. (2001). *A problem-posing approach to teaching decision making about the waste issue.*
36. Lijmbach, S., Broens, M., & Hovinga, D. (2000). *Duurzaamheid als leergebied; conceptuele analyse en educatieve uitwerking*.
35. Margadant-van Arcken, M. & Van den Berg, C. (2000). *Natuur in pluralistisch perspectief – Theoretisch kader en voorbeeldlesmateriaal voor het omgaan met een veelheid aan natuurbeelden*.
34. Janssen, F. J. J. M. (1999). *Ontwerpend leren in het biologieonderwijs. Uitgewerkt en beproefd voor immunologie in het voortgezet onderwijs*.
33. De Moor, E. W. A. (1999). *Van vormleer naar realistische meetkunde – Een historisch-didactisch onderzoek van het meetkundeonderwijs aan kinderen van vier tot veertien jaar in Nederland gedurende de negentiende en twintigste eeuw*.
32. Van den Heuvel-Panhuizen, M. & Vermeer, H. J. (1999). *Verschillen tussen meisjes en jongens bij het vak rekenen-wiskunde op de basisschool – Eindrapport MOOJ-onderzoek*.
31. Beeftink, C. (2000). *Met het oog op integratie – Een studie over integratie van leerstof uit de natuurwetenschappelijke vakken in de tweede fase van het voortgezet onderwijs*.

30. Vollebregt, M. J. (1998). *A problem posing approach to teaching an initial particle model*.
29. Klein, A. S. (1998). *Flexibilization of mental arithmeticsstrategies on a different knowledge base – The empty number line in a realistic versus gradual program design*.
28. Genseberger, R. (1997). *Interessegeoriënteerd natuur- en scheikundeonderwijs – Een studie naar onderwijsontwikkeling op de Open Schoolgemeenschap Bijlmer*.
27. Kaper, W. H. (1997). *Thermodynamica leren onderwijzen*.
26. Gravemeijer, K. (1997). *The role of context and models in the development of mathematical strategies and procedures*.
25. Acampo, J. J. C. (1997). *Teaching electrochemical cells – A study on teachers' conceptions and teaching problems in secondary education*.
24. Reygel, P. C. F. (1997). *Het thema 'reproductie' in het schoolvak biologie*.
23. Roebertsen, H. (1996). *Integratie en toepassing van biologische kennis – Ontwikkeling en onderzoek van een curriculum rond het thema 'Lichaamsprocessen en Vergift'*.
22. Lijnse, P. L. & Wubbels, T. (1996). *Over natuurkundedidactiek, curriculumontwikkeling en lerarenopleiding*.
21. Buddingh', J. (1997). *Regulatie en homeostase als onderwijsthema: een biologie-didactisch onderzoek*.
20. Van Hoeve-Brouwer G. M. (1996). *Teaching structures in chemistry – An educational structure for chemical bonding*.
19. Van den Heuvel-Panhuizen, M. (1996). *Assessment and realistic mathematics education*.
18. Klaassen, C. W. J. M. (1995). *A problem-posing approach to teaching the topic of radioactivity*.
17. De Jong, O., Van Roon, P. H. & De Vos, W. (1995). *Perspectives on research in chemical education.*
16. Van Keulen, H. (1995). *Making sense – Simulation-of-research in organic chemistry education*.
15. Doorman, L. M., Drijvers, P. & Kindt, M. (1994). *De grafische rekenmachine in het wiskundeonderwijs*.
14. Gravemeijer, K. (1994). *Realistic mathematics education*.
13. Lijnse, P. L. (Ed.) (1993). *European research in science education*.
12. Zuidema, J. & Van der Gaag, L. (1993). *De volgende opgave van de computer*.
11. Gravemeijer, K, Van den Heuvel Panhuizen, M., Van Donselaar, G., Ruesink, N., Streefland, L., Vermeulen, W., Te Woerd, E., & Van der Ploeg, D. (1993). *Methoden in het reken-wiskundeonderwijs, een rijke context voor vergelijkend onderzoek*.
10. Van der Valk, A. E. (1992). *Ontwikkeling in Energieonderwijs*.
9. Streefland, L. (Ed.) (1991). *Realistic mathematics education in primary schools*.

8. Van Galen, F., Dolk, M., Feijs, E., & Jonker, V. (1991). *Interactieve video in de nascholing reken-wiskunde.*
7. Elzenga, H. E. (1991). *Kwaliteit van kwantiteit*.
6. Lijnse, P. L., Licht, P., De Vos, W. & Waarlo, A. J. (Eds.) (1990). *Relating macroscopic phenomena to microscopic particles: a central problem in secondary science education*.
5. Van Driel, J. H. (1990). *Betrokken bij evenwicht*.
4. Vogelezang, M. J. (1990). *Een onverdeelbare eenheid*.
3. Wierstra, R. F. A. (1990). *Natuurkunde-onderwijs tussen leefwereld en vakstructuur*.
2. Eijkelhof, H. M. C. (1990). *Radiation and risk in physics education*.
1. Lijnse, P. L. & De Vos, W. (Eds.) (1990). *Didactiek in perspectief.*

Recent publications in the ICO Dissertation Series

In the ICO Dissertation Series dissertations are published of graduate students from faculties and institutes on educational research within the following universities: Eindhoven University of Technology, Leiden University, Maastricht University, Open University of the Netherlands, University of Amsterdam, University of Groningen, University of Twente, Utrecht University, VU University Amsterdam, and Wageningen University (and formerly Radboud University Nijmegen and Tilburg University).

220. Bruin-Muurling, G. (21-12-2010). *The development of proficiency in the fraction domain: Affordances and constraints in the curriculum*. Eindhoven: Eindhoven University of Technology.
219. Kostons, D.D.N.M. (05-11-2010). *On the role of self-assessment and task-selection skills in self-regulated learning*. Heerlen: Open University of the Netherlands.
218. Vos, M.A.J. (30-09-2010). *Interaction between teachers and teaching materials: On the implementation of context-based chemistry education*. Eindhoven: Eindhoven University of Technology.
217. Bonestroo, W.J. (24-09-2010). *Planning with graphical overview: Effects of support tools on self-regulated learning*. Enschede: University of Twente.
216. Groenier, M. (10-09-2010). *The decisive moment: Making diagnostic decisions and designing treatments*. Enschede: University of Twente.
215. De Bakker, G.M. (08-09-2010). *Allocated online reciprocal peer support as a candidate for decreasing the tutoring load of teachers*. Eindhoven: Eindhoven University of Technology.
214. Endedijk, M.D. (02-07-2010). *Student teachers' self-regulated learning*. Utrecht: Utrecht University.
213. Kessels, C.C. (30-06-2010). *The influence of induction programs on beginning teachers' well-being and professional development*. Leiden: Leiden University.
212. Duijnhouwer, H. (04-06-2010). *Feedback effects on students' writing motivation, process, and performance*. Utrecht: Utrecht University.
211. Moolenaar, N.M. (01-06-2010). *Ties with potential: Nature, antecedents, and consequences of social networks in school teams*. Amsterdam: University of Amsterdam.
210. Mittendorff, K. M. (12-03-2010). *Career conversations in senior secondary vocational education*. Eindhoven: Eindhoven University of Technology.
209. Platteel, T. (11-02-2010). *Knowledge development of secondary school L1 teachers on concept-context rich education in an action-research setting*. Leiden: Leiden University.
208. Koopman, M. (11-02-2010). *Students' goal orientations, information processing strategies and knowledge development in competence-based pre-vocational secondary education*. Eindhoven: Eindhoven University of Technology.

207. Zitter, I.I. (04-02-2010). *Designing for learning: Studying learning environments in higher professional education from a design perspective*. Utrecht: Utrecht University.
206. Wetzels, S.A.J. (18-12-2009). *Individualized strategies for prior knowledge activation*. Heerlen: Open University of the Netherlands.
205. Kirschner, F.C. (11-12-2009). *United brains for complex learning*. Heerlen: Open University of the Netherlands.
204. Rohaan, E.J. (09-12-2009). *Testing teacher knowledge for technology teaching in primary schools*. Eindhoven: Eindhoven University of Technology.
203. Seezink, A. (18-11-2009). *Continuing teacher development for competence-based teaching*. Tilburg: Tilburg University.
202. Van Ewijk, R. (20-10-2009). *Empirical essays on education and health*. Amsterdam: University of Amsterdam.
201. Mainhard, M.T. (25-09-2009). *Time consistency in teacher-class relationships*. Utrecht: Utrecht University.
200. Van der Rijst, R.M. (23-09-2009). *The research-teaching nexus in the sciences: Scientific research dispositions and teaching practice*. Leiden: Leiden University.
199. Visser-Wijnveen, G.J. (23-09-2009). *The research-teaching nexus in the humanities*: Variations among academics. Leiden: Leiden University.
198. Nievelstein, F.E.R.M. (18-09-2009). *Learning law: Expertise differences and the effect of instructional support*. Heerlen: Open University of the Netherlands.
197. Handelzalts, A. (17-09-2009). *Collaborative curriculum development in teacher design teams*. Enschede: University of Twente.
196. Van Borkulo, S.P. (26-06-2009). *The assessment of learning outcomes of computer modelling in secondary science education*. Enschede: University of Twente.
195. Du Chatenier, E. (09-09-2009). *Open innovation competence: Towards a competence profile for inter-organizational collaboration in innovation teams*. Wageningen: Wageningen University.
194. Radstake, H. (14-05-2009).*Teaching in diversity: Teachers and pupils about tense situations in ethnically heterogeneous classes*. Amsterdam: University of Amsterdam.