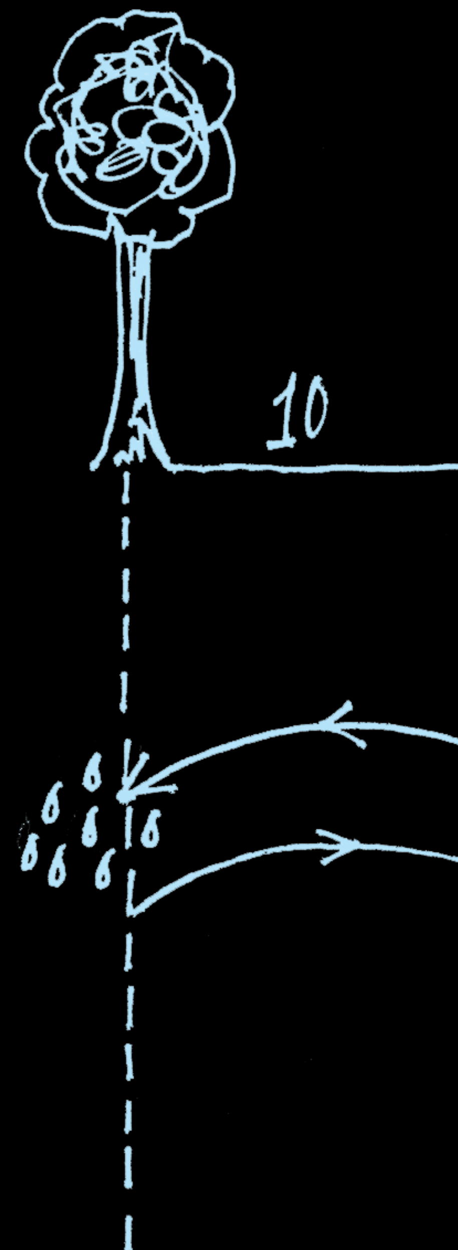
COMPREHEND, VISUALIZE & CALCULATE

Solving mathematical word problems in contemporary math education



Anton J. H. Boonen

COMPREHEND, VISUALIZE & CALCULATE:

Solving mathematical word problems in contemporary math education

BEGRIJPEN, VERBEELDEN & BEREKENEN:

Het oplossen van talige rekenopgaven in het hedendaagse rekenonderwijs

Cover pictures:	Marlies ter Beek-Geertse
Cover design & layout:	Lyanne Tonk, persoonlijk proefschrift.nl
ISBN:	978-94-6259-564-4
Printed by:	Ipskamp Drukkers

VRIJE UNIVERSITEIT

COMPREHEND, VISUALIZE & CALCULATE:
Solving mathematical word problems in contemporary math education

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. F.A. van der Duyn Schouten,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Psychologie en Pedagogiek
op donderdag 12 maart 2015 om 11.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Antonius Josephus Henricus Boonen

geboren te Heeze

promotor: prof.dr. J. Jolles
copromotor: dr. M. van der Schoot

Beoordelingscommissie
prof. dr. L. Verschaffel. University of Leuven
prof. dr. J.E.H. van Luit, Utrecht University
prof. dr. M. Brekelmans, Utrecht University
prof. dr. E.C.D.M. van Lieshout, VU University

Paranimfen
Jeffrey Derks
Paul van den Eijnden

COLOFON

De in dit proefschrift beschreven studies zijn mogelijk gemaakt door financiële steun van:

Vrije Universiteit Amsterdam

Hogeschool Windesheim

De totstandkoming van dit proefschrift werd mede mogelijk gemaakt door financiële steun van:

CPS
Onderwijsontwikkeling en advies

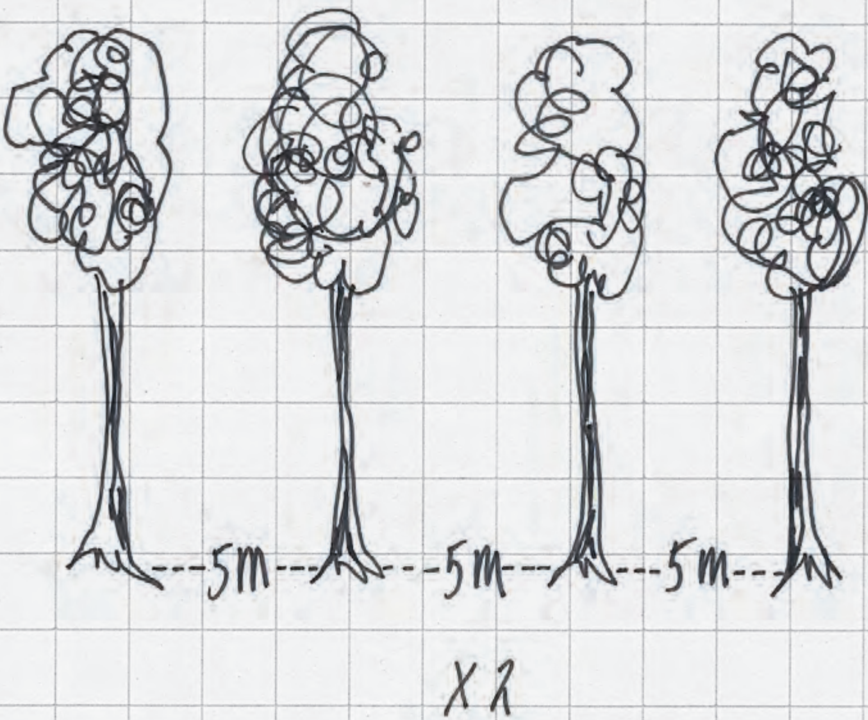
APS Onderwijsadviesbureau te Utrecht

Zwijsen te Tilburg

CONTENTS


UHHH... HOUSTON, WE HAVE A WORD PROBLEM.
IF A TRAIN LEAVES THE STATION AT 9:00AM HEADING NORTH AT 55 MPH, AND ANOTHER TRAIN
STATION AT
SOUTH AT 72
BOTH
©2010 Scott Hilburn/Distributed by Universal Uclick 1/16

1 General Introduction



At each of the two ends of a straight path, a man planted a tree and then every 5 meters along the path he planted another tree. The length of the path is 15 meters. How many trees were planted?

INTRODUCTION

In the last decades, mathematical word problem solving has gained much attention from both researchers and educational practitioners (Campell, 1992; Depaepe, De Corte, & Verschaffel, 2010; Hegarty, Mayer, & Monk, 1995; Hickendorff, 2011; Moreno, Ozogul, & Reisslein, 2011; Swanson, Lussler, & Orosco, 2013). The main focus of the scientific literature on this subject has generally been the word problem solving performances of elementary, middle school and undergraduate students and their use of superficial and/or sophisticated representation strategies (see Cummins, Kintsch, Reusser, & Weimer, 1988; Hegarty & Kozhevinkov, 1999; Pape, 2003; Van der Schoot, Bakker-Arkema, Horsley, & Van Lieshout, 2009; Verschaffel, De Corte, & Pauwels, 1992). However, previous studies provide limited insight into the exact nature of these representation strategies, and investigate the visual-spatial and semantic-linguistic components skills and abilities underlying word problem solving separately from each other. The research reported in the present thesis has therefore been designed to examine students' representation strategies in more detail, and to investigate the underlying components and skills in conjunction with each other.

The effectiveness of word problem solving instructions has been another main area of focus in past research (Jitendra et al., 2013; Jitendra & Star, 2012; Jitendra et al., 2009; Krawec, 2010, 2012). However, until now instructions have been generally executed by researchers in small groups of low ability students in special education. In contrast to previous studies, the research reported in the current thesis examines the role of the teacher while implementing a word problem solving instruction in his/her own regular classroom practice. This is important, given that mainstream schools are becoming more inclusive, and that a greater number of students attending them have mild to severe learning difficulties (Jitendra & Star, 2012; Sharma, Loreman, & Forlin, 2012).

1

BACKGROUND

The findings of the studies that were conducted as part of the research presented in this thesis have implications for the practice of contemporary math education. Drawing clear recommendations based on these findings could contribute to the improvement of mathematics learning and teaching in schools, particularly when it concerns mathematical word problem solving. Before stating the objectives of this thesis, I will give some background information regarding word problem solving and related factors.

Mathematical word problems and word problem solving

Word problems play a prominent role in both the educational practice of contemporary math approaches and in educational research. The term word problem is used to refer to any math exercise where significant background information on the problem is presented as text rather than in the form of mathematical notation. As word problems often involve a narrative of some sort, they are occasionally also referred to as *story problems* (Verschaffel, Greer, & De Corte, 2000). The literature generally makes a distinction between so-called *routine* and *non-routine* word problems (Pantziara, Gagatsis, & Elia, 2009; Schoenfeld, 1992). Routine word problems have a fixed problem structure and involve the application of routine calculations.

Routine word problems

Combine, change, and compare word problems are routine types of word problem that are commonly offered in elementary school.

In a combine word problem, reflected in the example below, a subset or superset must be computed given the information about two other sets. This type of word problem involves understanding part-whole relationships and knowing that the whole is equal to the sum of its parts (Cummins et al., 1988; Jitendra, 2002, Jitendra, DiPipi, & Perron-Jones, 2002).

> ***[Combine word problem]:***
> *Mary has 4 marbles. John has some marbles. They have 7 marbles altogether. How many marbles does John have?*

Change word problems are routine word problems in which a starting set undergoes a transfer-in or transfer-out of items, and the cardinality of a start set, transfer set, or a result set must be computed given information about two of the sets (Cummins et al.,1988; Jitendra et al., 2013). In other words, a change word problem starts with a beginning set in which the object identity and the amount of the object are defined. Then a change occurs to the beginning set that results in an 'ending set' in which the new amount is defined (Jitendra, 2002).

> ***[Change word problem]:***
> *Mary had 8 marbles. Then she gave some marbles to John. Now Mary has 3 marbles. How many marbles did she give to John?*

The last type of routine word problem that is investigated in many studies is a compare word problem. In compare word problems the cardinality of one set must be computed by comparing the information given about relative sizes of the other set sizes; one set serves as the comparison set and the other as the referent set. In this type of word problem, students often focus on relational terms like 'more than' or 'less than' to compare the two sets and identify the difference in value between the two sets (Cummins et al., 1988; Hegarty et al., 1995; Pape, 2003; Van der Schoot et al., 2009).

> ***[Compare word problem]:***
> *Mary has 5 marbles. John has 8 marbles. How many marbles does John have more than Mary?*

Non-routine word problems

In contrast with routine word problems, non-routine word 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 (Elia, Van den Heuvel-Panhuizen, & Kolovou, 2009). In other words, it is characteristic for non-routine word problems that they cannot be solved in a prescribed way. Solution strategies of non-routine word problems can, therefore, differ between each word problem that is solved. An example of a non-routine word problem is reported below.

> ***[Non-routine word problem]:***
> *A balloon first rose 200 meters from the ground, then moved 100 meters to the east, then dropped 100 meters. It then traveled 50 meters to the east, and finally dropped straight to the ground.*

How far was the balloon from its original starting point?

The solution of word problems

Whereas routine word problems like combine, change and compare problems, are frequently offered in the early grades of elementary school, students in the sixth grade are expected to solve a wide variety of non-routine word problems of increasing difficulty. Therefore, the research presented in this thesis is interested in both routine and non-routine word problems and examines students and teachers from early and later grades of elementary school.

Generally, the solution of word problems depends on two major phases: (1) *problem comprehension,* which involves the identification and representation of the problem structure of the word problem; and, (2) *problem solution,* which involves the determination of the used mathematical operations and the execution of these planned computations to solve the problem (Krawec, 2010; Lewis & Mayer, 1987).

A substantial amount of elementary school students has difficulties with solving word problems. This is not because of their inability to execute the planned mathematical computations, but a result of their difficulties with thoroughly understanding and representing the word problem text and distilling the correct mathematical operations to be performed (Carpenter, Corbitt, Kepner, Lindquist, & Reys, 1981; Cummins et al., 1988; Krawec, 2010; Lewis & Mayer, 1987; Van der Schoot et al., 2009). Hence, mistakes in word problem solutions frequently occur in the problem comprehension phase, rather than in the problem solution phase.

Visualization

The external (i.e., a gesture or drawing with paper and pencil) or internal (i.e., mental) construction of a visual representation is thought to be a powerful tool for overcoming the difficulties in understanding the word problem text. According to Hegarty and Kozhevnikov (1999), two types of visual representations can be distinguished: *pictorial* and *visual-schematic* representations.

Children who create pictorial representations tend to focus on the visual appearance of the given elements in the word problem. These representations consist of vivid and detailed visual images (Hegarty & Kozhevnikov, 1999; Presmeg, 1997, see Figure 1). However, several studies have reported that the production of pictorial representations is negatively related to word problem solving performance (Ahmad, Tarmizi, & Nawawi, 2010; Hegarty & Kozhevnikov, 1999; Kozhevnikov, Hegarty, & Mayer, 2002; Krawec, 2010; Van Garderen, 2006; Van Garderen & Montague, 2003). An explanation for this finding is that children who make pictorial representations fail to form a coherent visualization of the described problem situation and base their representations solely on a specific element or sentence in the word problem text (Hegarty & Kozhevnikov, 1999; Krawec, 2010; Van Garderen, 2006; Van Garderen & Montague, 2003).

Visual-schematic representations, on other hand, do contain a coherent image of the problem situation hidden in the word problem, including the relations between the solution-relevant elements (Edens & Potter, 2008; Hegarty & Kozhevnikov, 1999; Kozhevnikov et al., 2002; Van Garderen & Montague, 2003, see Figure 2). This explains why, in contrast to the production of pictorial representations, the production of visual-schematic representations is positively related to word problem solving performance (Hegarty & Kozhevnikov, 1999; Van Garderen, 2006; Van Garderen & Montague, 2003).

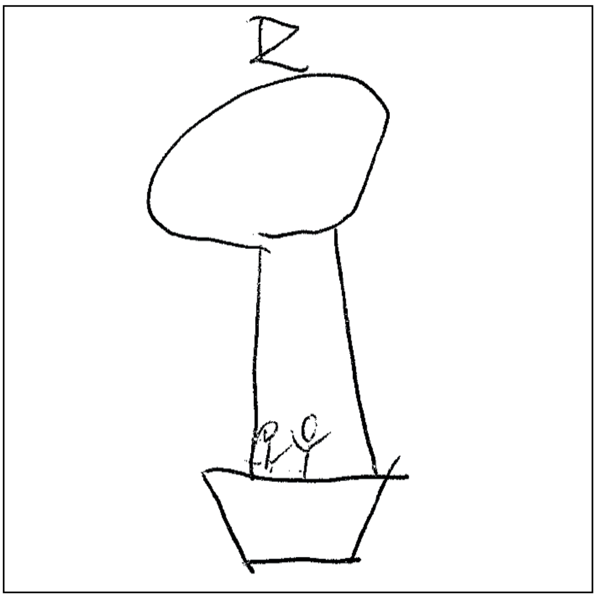


Figure 1. An example of a pictorial representation

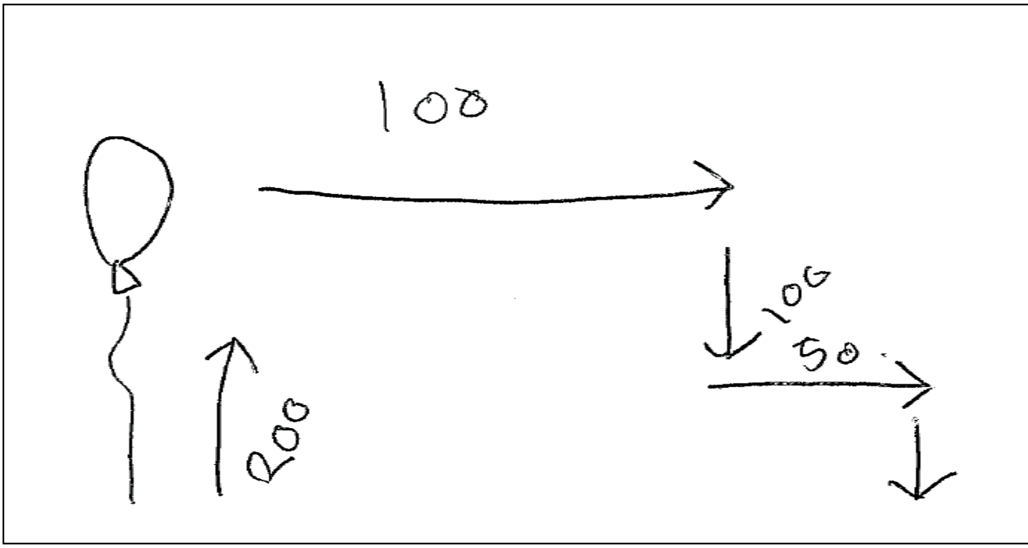


Figure 2. An example of a visual-schematic representation

The importance of visual-spatial skills in word problem solving

The scientific literature shows that spatial ability is a basic ability underlying mathematical word problem solving (e.g., Blatto-Vallee, Kelly, Gaustad, Porter, & Fonzi, 2007; Edens & Potter, 2008; Hegarty & Kozhevnikov, 1999). Spatial ability is related to students' word problem solving performance, as well as to the components/factors that influence this performance. Spatial skills are, for example, closely related to the production of visual-schematic representations, and children with good spatial skills have been found to be better able to make visual-schematic representations than children with poor spatial skills (e.g., Hegarty & Kozhevnikov, 1999; Krawec, 2010; Van Garderen, 2006; Van Garderen & Montague, 2003). These visual-schematic representations are in turn a factor affecting word problem solving performance. Although there are many definitions of what spatial ability is, it is generally accepted to be related to skills involving the retrieval, retention and transformation of visual information in a spatial context (Velez, Silver, & Tremaine, 2005). Especially the involvement of a specific spatial factor, i.e., *spatial visualization,* in making coherent visual-schematic representations has been made clear by several authors (Hegarty & Kozhevnikov, 1999; Krawec, 2010; Van Garderen, 2006; Van Garderen & Montague, 2003).

Besides indications that spatial ability plays an *indirect* role in word problem solving via the production of visual-schematic representations, several authors also report a *direct* relation between spatial ability and word problem solving (Battista, 1990; Blatto-Vallee et al., 2007; Booth & Thomas, 1999; Edens & Potter, 2008; Geary, Saults, Liu, & Howard, 2000). For example Blatto-Vallee et al., (2007), showed that spatial abilities explained almost 20% of unique variance in word problem solving performance.

Spatial ability and students' constructive play activities

Another, somewhat different way in which spatial ability relates to word problem solving is through its role in students' constructive play activities. Constructive play generally involves the manipulation, construction and motion of objects in space (i.e., rotating) (Caldera, Culp, O'Brian, Truglio, Alvarez, & Huston, 1999; Mitchell, 1973; Pomerleau, Malcuit, & Séguin, 1990). Constructive play activities that are related to performance on spatial tasks are Lego, Blocks, and jigsaw puzzles (Caldera et al., 1999; Levine, Ratkliff, Huttenlocher, & Cannon, 2012; Mitchell, 1973; Pomerleau et al., 1990).

However, previous studies have not reported a direct relation between constructive play and students' word problem solving performance. To fill this gap, the research presented in this thesis tries to gain more insight in the specific relation between spatial ability, word problem solving, and constructive play. Specifically, we investigated the mediating role of spatial ability in the relation between constructive play and word problem solving.

The importance of semantic-linguistic skills in word problem solving

Besides (visual-)spatial skills, several previous studies showed that semantic-linguistic skills (i.e., reading comprehension) are also closely related to word problem solving (Pape, 2004; Vilenius-Tuohimaa, Aunola, & Nurmi, 2008). General reading comprehension abilities are found to be important in dealing with semantic-linguistic word problem characteristics, such as the semantic structure of a word problem, the sequence of the known elements in the problem text, and the degree to which the semantic relations between the given and the unknown quantities of the problem are stated explicitly (De Corte, Verschaffel, & De Win, 1985). All these word problem characteristics have been shown to have an effect on children's solution processes (e.g., De Corte et al., 1985; De Corte & Verschaffel, 1987; Søvik, Frostrad, & Heggberget, 1999).

Word problem solving instruction

Many studies have concluded that students experience severe difficulties in solving word problems and accentuated the importance of skills that help students identify and represent the word problem text to generate a correct word problem solution (e.g. Carpenter et al., 1981; Cummins et al., 1988; Van Garderen, 2006). In spite of these findings, there is still a lack of instructional programs that take an evidence-based approach to word problem solving and are adapted to the educational practice of contemporary math education.

Cognitive Strategy Instruction (CSI), an instructional approach that focuses on explicit instruction in cognitive and metacognitive strategies that help students identify and represent the problem structure in order to improve their word problem solving performance, seems however promising (e.g., Jitendra, DiPipi, & Perron-Jones, 2002, Jitendra et al., 2013; Jitendra, & Star, 2012; Jitendra et al., 2009; Krawec, 2012; Krawec, Huang, Montague, Kressler, & Melia de Alba, 2013; Montague, Enders, & Dietz, 2011; Montague, Warger, & Morgan, 2000). A commonly investigated example of a CSI is schema-based instruction (SBI, Jitendra et al., 2002, 2009, 2012, 2013). Schema-based instruction moves away from keywords and superficial problem features and is more focused on helping children find the underlying problem structure. In SBI students are taught to identify and represent the problem structures of certain word problem types (by constructing a visual-schematic representation or diagram), and are encouraged to reflect on the similarities and differences between these problem types. The implementation of SBI in the curriculum of contemporary math education appears, however, to be challenging.

An alternative instructional approach, focused on the use of cognitive strategies, has been developed by Montague (2003) and is known as the *Solve It!* instructional program. The Solve It! program is a more heuristic approach that teaches students how to: (a) read the problem for understanding; (b) paraphrase the problem by putting it into their own words; (c) visualize the problem; d) set up a plan for solving the problem; (e) compute; and (f) verify the solution of the problem.

However, like SBI also the Solve It! method has some important restrictions. Firstly, the cognitive step in which students are requested to visualize the word problem seems to be defined too generally. Findings show that it is incorrect to assume that a student knows exactly what pictures to draw, when, and under what circumstances, and for which type of problems (Jitendra, Griffin, Haria, Leh, Adams,

& Kaduvettoor, 2007; Jitendra et al., 2009). Another problem with this step is that previous research showed that a visual representation should meet certain requirements (i.e., involve the correct relations between solution-relevant elements) and that not all types of visual representations facilitate the solution process of word problems (Krawec, 2010, Van Garderen & Montague, 2003). Secondly, the effectiveness of the Solve it! program has generally been examined in small groups of children with learning and mathematical disabilities (Jitendra et al., 2002, 2013; Krawec, 2010, 2012; Montague et al., 2000), and not in a regular classroom setting. In addition, the Solve It! method has only been implemented by researchers and not by teachers. Surprisingly, there is currently no comparable instructional support available for teachers in mainstream classrooms. This is an important omission, given that mainstream schools are becoming more inclusive, and that a greater number of students attending them have mild to severe learning difficulties (Jitendra & Star, 2012; Sharma et al., 2012). It would, therefore, help teachers if they had instructional approaches at their disposal that have been designed to teach skills important for word problem solving.

Bearing this state of affairs in mind this thesis examines the introduction of an innovative approach to the instruction of word problem solving in mainstream classrooms, and examines how teachers implemented that approach, with a focus on their use of visual representations.

THESIS OUTLINE

Objectives

Research on word problem solving is often focused on the performance of students and not on their comprehension of the word problem text. Difficulties with word problem solving can, however, often be ascribed to problems with the correct understanding of the word problem text. The research presented in this thesis is, therefore, focused on the component processes and skills that underlie the successful comprehension of word problems. In particular, we examined both students' use of visual representations and the quality of these visual representations. Moreover, we were interested in the extent to which different types of visual representation increase or decrease the chance of solving a word problem correctly. To this end, this thesis sets out to achieve the following two objectives.

The first objective is to examine students' performances, notably the extent to which students use different types of visual representations, and the role that spatial and semantic-linguistic skills play in the solving of routine and non-routine word problems in early (second) and later (sixth) grades of elementary school.

The second objective of this thesis is to investigate how teachers implement an innovative instructional approach – based on the didactical use of visual-schematic representation – in their own classroom teaching practice. This instructional approach requires teachers to use visual-schematic representations that visualize the problem structure in a diverse and flexible way as well as to vary the kinds of representations in a way that suits problem characteristics. Moreover, they are expected to model the representation process transparently, correctly and completely, as well as to construct visual representations that correctly and completely depict the relations between all the components relevant to the solution of the problem.

Approach

To achieve the objectives of this thesis we conducted five cross-sectional studies in the field of educational psychology in which both second (N = 47) and sixth grade (N = 128) elementary school students were examined. In addition, we conducted one study in which we investigated the way in which eight mainstream – sixth grade – teachers implemented a teaching intervention for supporting non-routine word problem solving.

Furthermore, we conducted a feasibility study in which we examined four second-grade students who performed poorly in word problem solving. The feasibility study has been included as an Appendix to the scientific part of the thesis: it has been included in order to give an illustration of a word problem solving instruction which could be suitable in the early grades of elementary school.

Chapter overview

The first three studies conducted for the research presented in this thesis focus on two component skills and two basic abilities of word problem solving. The component skills examined are: 1) the production of visual-schematic representation, and 2) relational processing (i.e., deriving the correct relations between solution-relevant elements of the word problem text base). The two basic abilities examined are: 1) spatial ability, and 2) reading comprehension. These component skills and the underlying basic abilities related to them belong to two different processing domains: the visual-spatial and the semantic-linguistic domain.

In ***Chapter 2*** a study is reported in which the two component skills and two basic abilities were investigated in one theoretical model (see Figure 3), in order to examine the extent to which they explain unique variance in sixth graders' word problem solving performances.

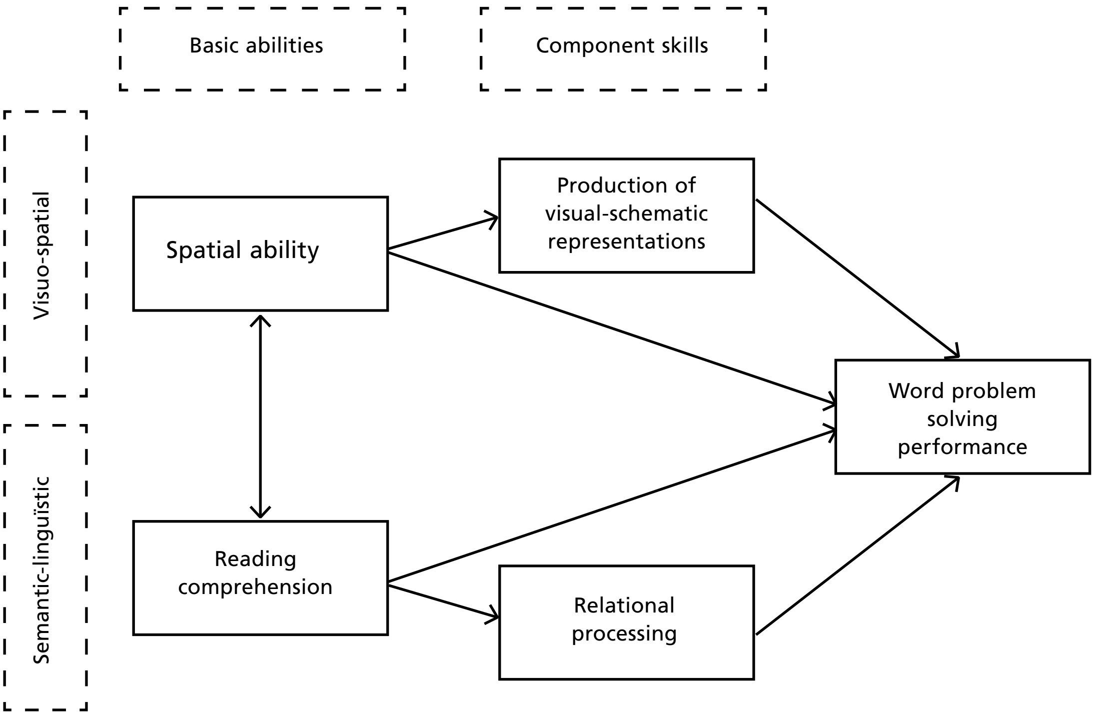


Figure 3. Path model with all hypothesized pathways

The purpose of the study reported in ***Chapter 3*** was to demonstrate that word problem solving instruction in Realistic Math Education seems to pay too little attention to the teaching of semantic-linguistic skills (i.e., reading comprehension) that allow sixth grade students to handle linguistic complexities in a word problem. The study attempted to show the importance of semantic-linguistic skills for the solution of semantically complex and less complex inconsistent compare word problems, in a group of successful and less successful sixth grade word problem solvers following the Dutch Realistic Math curriculum.

In ***Chapter 4*** the findings are examined of a study that investigated the importance of different types of visual representation, spatial ability, and reading comprehension on the word problem solving performance of sixth grade students. In contrast to previous studies, an item-level approach was used in this study rather than a test-level approach.

This change in statistical modeling generated a more thorough and sophisticated understanding of the process and enabled us to examine if and to what extent the production of a specific kind of visual representation affected the chance of successfully solving the word problem of which the visual representation had been made. Moreover, we wanted to examine if we were able to reproduce the findings of test-level analysis with regard to the importance of spatial ability and reading comprehension, by using an item-level analysis. This made it possible to identify any level of analysis discrepancies.

Chapter 5 is focused on the importance of spatial ability, and the role it plays in the relation between (early) constructive play activities and word problem solving performance of sixth grade elementary school students. The aim of the study described in this chapter was to investigate whether spatial ability acted as a mediator in the relation between constructive play and mathematical word problem solving performance in 128 sixth grade elementary school children.

The studies described in chapters 2 to 5 focused on the strategies, solution processes and performances on word problems of students in higher grades of elementary school (i.e., grade 6). Word problems are, however, already offered in the first grades of elementary school. Moreover, scientific research has shown that students from first and second grade of elementary school already experience difficulties solving word problems. Therefore, in ***Chapter 6*** a study that investigated the word problem solving performances of second grade elementary school students is reported. The findings of this study reveal a plausible reason for second grade students' differing performances on three commonly investigated routine word problem types, namely combine, change and compare problems.

The studies described in chapters 2 to 6 of this thesis are focused on the difficulties experienced by students in solving word problems. Their findings accentuate the importance of skills that help students to identify and represent the word problem text correctly in order to generate a deep understanding of the problem situation. Evidence-based word problem solving instructional programs that could help develop these skills are, however scarce, limited in scope and often not adapted to the educational practice of mainstream classrooms. In the study described in ***Chapter 7*** this is addressed by examining teachers' use of and their competence in making visual-schematic representations while executing an innovative word problem solving instruction in their own classrooms. The study was performed in the context of a teaching intervention that involved eight teachers who felt confident about teaching mathematics and using visual representations, and who were motivated to participate in and contribute to research in this area.

Chapter 8 reflects on the findings from the series of studies conducted as part of this research on word problem solving, and examines them in light of the objectives of this thesis. Furthermore, the implications of these findings for educational practice are considered and a list of recommendations for teacher training and teacher professionalization is presented and discussed. Finally, the main findings of this study are discussed in a broader perspective. That is, in the perspective of the current debate on the importance of building bridges between educational research and the educational practice. Central to this debate is the question whether the outcomes of educational research can be directly implemented in the classroom practice of students and teachers.

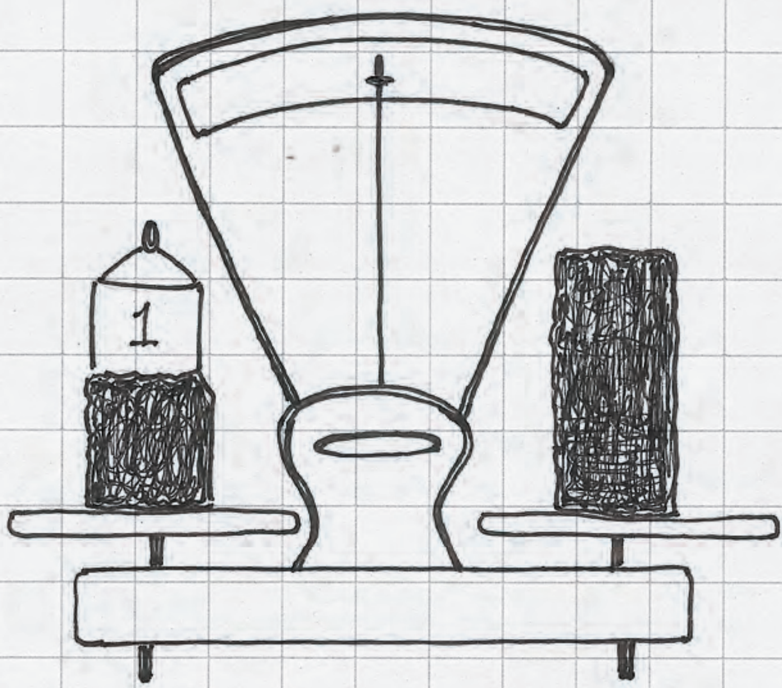
As already indicated above, besides the six studies discussed as part of the research presented in this thesis, ***Appendix I*** contains a feasibility study in which a word problem instruction based on the principles of the Solve it! method and schema-based instruction is evaluated. The aim of this feasibility study was to experiment with a word problem solving instruction, examine the extent to which it was (un)successful, and see whether second grade students were able to execute the cognitive steps of this instruction and improve their word problem solving performances on combine, change and compare problems. Although the feasibility study does not have significant scientific value, it does provide an example of a word problem solving instruction that can be given to students who are still in the early grades of elementary school.

2 What underlies successful word problem solving?

A path analysis in sixth grade students

Anton J. H. Boonen, Menno van der Schoot,
Floryt van Wesel, Meinou de Vries, & Jelle Jolles

Contemporary Educational Psychology
(2013), 38, 271-279



On one side of a scale there is a l kg weight and half a brick. On the other side there is one full brick. The scale is balanced. What is the weight of the brick?

ABSTRACT

Two component skills are thought to be necessary for successful word problem solving: (1) the production of visual-schematic representations and (2) the derivation of the correct relations between the solution-relevant elements from the text base. The first component skill is grounded in the visual-spatial domain, and presumed to be influenced by spatial ability, whereas the latter is seated in the semantic-linguistic domain, and presumed to be influenced by reading comprehension. These component skills as well as their underlying basic abilities are examined in 128 sixth grade students through path analysis. The results of the path analysis showed that both component skills and their underlying basic abilities explained 49% of students' word problem solving performance. Furthermore, spatial ability and reading comprehension both had a direct and an indirect relation (via the component skills) with word problem solving performance. These results contribute to the development of instruction methods that help students using these components while solving word problems.

INTRODUCTION

Mathematical word problem solving

Mathematical word problem solving plays a prominent role in contemporary mathematics education (Rasmussen & King, 2000; Timmermans, Van Lieshout, & Verhoeven, 2007). The term *word problem* is used to refer to any math exercise where significant background information on the problem is presented as text rather than in mathematical notation. As word problems often involve a narrative of some sort, they are occasionally also referred to as *story problems* (Verschaffel, Greer, & De Corte, 2000). An example of a word problem is given below (taken from Hegarty & Kozhevnikov, 1999):

> ***[Example 1]***
> *At each of the two ends of a straight path, a man planted a tree and then, every 5 meters along the path, he planted another tree. The length of the path is 15 meters. How many trees were planted?*

Students often experience difficulties in the understanding of the text of a word problem, rather than its solution (Carpenter, Corbitt, Kepner, Lindquist & Reys, 1981; Lewis & Mayer, 1987). Two component skills are thought to be necessary for successful word problem solving: (1) producing visual-schematic representations (e.g., Hegarty & Kozhevnikov, 1999; Krawec, 2010; Montague & Applegate, 2000; Van Garderen & Montague, 2003) and (2) relational processing, that is deriving the correct relations between the solution-relevant elements from the text base (e.g., Hegarty, Mayer, & Monk, 1995; Kintsch, 1998; Van der Schoot, Bakker-Arkema, Horsley, & Van Lieshout, 2009; Verschaffel, 1994; Verschaffel, De Corte, & Pauwels, 1992). These two component skills are presumed to explain unique variance in students' word problem solving performance and cover different processing domains (Hegarty & Kozhevnikov, 1999; Krawec, 2010; Van der Schoot et al., 2009). The production of visual-schematic representations is grounded in the visual-spatial domain (e.g., Hegarty & Kozhevnikov, 1999; Krawec, 2010; Mayer, 1985; Van Garderen, 2006), whereas relational processing is seated in the semantic-linguistic domain (e.g., Pape, 2003; Thevenot, 2010; Van der Schoot et al., 2009). These component skills, as well as the basic abilities which underlie each of these skills, are described below.

Component skill in the visuo-spatial domain: The production of visual-schematic representations

Rather than the superficial selection of numbers and relational keywords from the word problem text (often resulting in the execution of the wrong arithmetic operations), good word problem solvers generally construct a visual representation of the problem to facilitate understanding (e.g., Hegarty & Kozhevnikov, 1999; Krawec, 2010; Montague & Applegate, 2000; Van der Schoot et al., 2009). With this, the nature of these visual representations determines their effectiveness. According to Hegarty and Kozhevnikov (1999), two types of visual representations exist: *pictorial* and *visual-schematic* representations. Children who create pictorial representations tend to focus on the visual appearance of the given elements in the word problem. These representations consist of vivid and detailed visual images (Hegarty & Kozhevnikov, 1999; Presmeg, 1997). However, several studies have reported that the production of pictorial representations is negatively related to word problem solving performance (Ahmad, Tarmizi, & Nawawi, 2010; Hegarty & Kozhevnikov, 1999; Kozhevnikov, Hegarty, & Mayer, 2002; Krawec, 2010; Van Garderen, 2006; Van Garderen & Montague, 2003). An explanation for this finding is that children who make pictorial representations fail to form a coherent visualization of the described problem situation and base their representations solely on a specific element or sentence in the word problem text (Hegarty & Kozhevnikov, 1999; Krawec, 2010; Van Garderen, 2006; Van Garderen & Montague, 2003). Children who make visual-schematic representations do integrate the solution-relevant text elements into a coherent visualization of the word problem (e.g., Ahmad et al., 2010; Krawec, 2010; Van Garderen, 2006). This explains why, in contrast to the production of pictorial representations, the production of visual-schematic representations is found to be positively related to word problem solving performance (Hegarty & Kozhevnikov, 1999; Van Garderen, 2006; Van Garderen & Montague, 2003).

Basic ability in the visuo-spatial domain: Spatial abilities

The production of visual-schematic representations is influenced by spatial ability. Children with good spatial skills have been found to be better able to make visual-schematic representations than children with poor spatial skills (e.g., Hegarty & Kozhevnikov, 1999; Krawec, 2010; Van Garderen, 2006; Van Garderen & Montague, 2003). Although there are many definitions of what spatial ability is, it is generally accepted to be related to skills involving the retrieval, retention and transformation of visual information in a spatial context (Velez, Silver, & Tremaine, 2005). Especially the involvement of a specific spatial factor - that is, *spatial visualization* - in making coherent visual-schematic representations has been made clear by several authors (Hegarty & Kozhevnikov, 1999; Krawec, 2010; Van Garderen, 2006; Van Garderen & Montague, 2003). Spatial visualization refers to the ability to mentally manipulate objects (i.e., mental rotation; Kaufmann, 2007; Voyer, Voyer, & Bryden, 1995). In the present study, spatial ability refers to spatial visualization as described above.

Besides the role of spatial ability in word problem solving via the production of visual-schematic representations, several authors also report a *direct* relation between spatial ability and word problem solving (Battista, 1990; Blatto-Vallee, Kelly, Gaustad, Porter, & Fonzi, 2007; Booth & Thomas, 1999; Edens & Potter, 2008; Geary, Saults, Liu, & Howard, 2000; Hegarty & Kozhevnikov, 1999; Orde, 1997). Blatto-Vallee et al. (2007), for example, showed that spatial abilities explained almost 20% of unique variance in word problem solving performance. Casey and colleagues revealed that the direct role of spatial abilities in word problem solving lies in performing the actual mathematical operations and numerical reasoning (e.g., Casey, Andrews, Schindler, Kersh, Samper, & Copley, 2008; Casey, Nuttall, & Pezaris, 1997, 2001).

Component skill in the semantic-linguistic domain: Relational processing

Although the production of visual-schematic representations is a necessary condition for successful word problem solving, it is not always a sufficient condition (Kintsch, 1998; Pape, 2003; Van der Schoot et al., 2009), since children may be very well capable of forming a visual-*schematic* representation without being able to infer the *correct* relations between the solution-relevant elements from the word problem text (Coquin-Viennot & Moreau, 2003; Krawec, 2010; Thevenot, 2010). Relational processing in word problem solving can be effectively revealed in word problems in which the relational term maps onto non-obvious mathematical operations (De Corte, Verschaffel, & De Win, 1985; Kintsch, 1998; Thevenot, 2010; Thevenot & Oakhill, 2006, 2008; Van der Schoot et al., 2009). In word problems with an obvious mapping, it is sufficient to first select the numbers and relational terms from the text and then to directly translate these into a set of computations (Hegarty, Mayer, & Monk, 1995; Pape, 2003; Van der Schoot et al., 2009). However, in non-obvious word problems, other text elements are necessary for the construction of an effective mental model of the word problem including the appropriate relations between the key variables (De Corte et al., 1985; Thevenot, 2010; Thevenot & Oakhill, 2006, 2008; Van der Schoot et al., 2009). Consider, for example, the following word problem in which the relation term 'more than' primes an inappropriate mathematical operation:

> ***[Example 2]***
> *At the grocery store, a bottle of olive oil costs 7 euro.*
> *That is 2 euro 'more than' at the supermarket.*
> *If you need to buy 7 bottles of olive oil, how much will it cost at the supermarket?*

In this so-called inconsistent word problem (Hegarty, Mayer, & Green, 1992; Hegarty et al., 1995; Kintsch, 1998; Van der Schoot et al., 2009), the crucial arithmetic operation (i.e., 7-2) cannot be simply derived from the relational keyword ('more than'). Rather than making use of a superficial, direct-retrieval strategy (Giroux & Ste-Marie, 2001; Hegarty et al., 1995; Thevenot, 2010; Verschaffel, 1994; Verschaffel et al., 1992), problem solvers have to appeal to a problem-model strategy in which they translate the problem statement into a qualitative mental model of the base type of situation (in this case: a subtraction situation) that is hidden in the problem. Here, this translation requires the identification of the pronominal reference 'that is' as the indicator of the relation between the value of the first variable ('the price of a bottle of olive oil at the grocery store') and the second ('the price of a bottle of olive oil at the supermarket'). On the basis of the constructed mental model, problem solvers are then able to plan and execute the required arithmetic operations. Hence, inconsistent word problems are suitable to measure relational processing.

Basic ability in the semantic-linguistic domain: Reading comprehension

Previous studies have shown that the role of relational processing in word problem solving is influenced by a child's reading comprehension abilities (e.g., Lee, Ng, Ng, & Lim, 2004; Van der Schoot et al., 2009). For example, Lewis and Mayer (1987), Pape (2003), Van der Schoot et al. (2009), and Verschaffel et al. (1992) showed that children find it easier to convert the relation term ‘more than’ to a subtraction operation (as in the example above) than the relational term ‘less than’ to an addition operation. This effect has been explained by assuming that the semantic memory representation of ‘less than’ is more complex than that of ‘more than’, an effect which is known as the lexical marking principle (Clark, 1969). The reason behind this effect is that the marked relational term (‘less than’) and unmarked relational term (‘more than’) differ in their frequency of occurrence (French, 1979; Goodwin & Johnson- Laird, 2005; Schriefers, 1990). Whereas the marked term is used only in its contrastive, ‘negative’ sense (‘Peter has less marbles than David’), the unmarked term is used in two senses: the contrastive, ‘positive’ sense (‘Peter has more marbles than David’) but also a neutral, nominal sense (‘Does she have more than one child?’). For word problem solving, the implication is that the memory representation of ‘less than’ is more ‘fixed’ than the memory representation of ‘more than’ (Van der Schoot et al., 2009). Presumably, the fixedness of its memory representation hinders the problem solvers’ ability to reverse ‘less than’ in the inconsistent condition (in which it primes the inappropriate arithmetic operation).What is of relevance here is that processing a marked relational term such as ‘less than’ (or ‘times less than’) is found to be closely associated with reading comprehension abilities (Van der Schoot et al., 2009). In particular, overcoming its semantic complexity and performing the statement reversal are thought to be comprehension-related skills (Kintsch, 1998; Thevenot, 2010). Thus, in this study, reading comprehension is hypothesized to have an indirect effect on word problem solving performance via its influence on relational processing, that is, the mapping of mathematical terms onto mathematical operations (Lee et al., 2004).

However, previous studies have also demonstrated a direct effect between reading comprehension and word problem solving (Pape, 2004; Vilenius-Tuohimaa, Aunola, & Nurmi, 2008). Presumably, general reading comprehension abilities are important in dealing with the semantic-linguistic word problem characteristics such as the semantic structure of a word problem, the sequence of the known elements in the problem text, and the degree in which the semantic relations between the given and the unknown quantities of the problem are stated explicitly (De Corte et al., 1985). All these word problem characteristics have been shown to have an effect on children’s solution processes (e.g., De Corte et al., 1985; De Corte & Verschaffel, 1987; Søvik, Frostrad, & Heggberget, 1999).

Given that they are grounded in different processing domains (visual-spatial and semantic-linguistic), the two major component skills in word problem solving (production of visual-schematic representation and relational processing) are hypothesized to be unrelated in this study. Yet, the basic abilities which are presumed to underlie these component skills, respectively spatial ability and reading comprehension, are expected to be connected. This hypothesis is based on studies which indicate that both abilities share some cognitive elements like working memory (Ackerman, Beier, & Boyle, 2005; Hannon & Daneman, 2001; Shah & Miyake, 1996) and general intelligence (Ackerman et al., 2005; Keith, Reynolds, Patel, & Ridley, 2008), as well as on the large body of studies which accentuate the importance of spatial ability in the production of non-linguistic situation models during reading comprehension (Haenggi, Kintsch, & Gernsbacher, 1995; Kendeou, Papadopoulos, & Spanoudis, 2012; Kintsch, 1998; Phillips, Jarrold, Baddeley, Grant, & Karmiloff-Smith, 2004; Plass, Chun, Mayer, & Leutner, 2003). Nonetheless, we do not expect the relation between spatial ability and reading comprehension to bring about a direct relation between the two component skills. This expectation is based on the assumption that the direct relationship between these component skills is weak and will therefore vanish in the presence of (the relationship between) the basic abilities.

The present study

A path model for successful word problem solving is established on the basis of the two component skills and their underlying basic abilities as discussed above. The complete path model is represented in Figure 1. The upper part of the model involves constructs in the visuo-spatial domain - that is, visual-schematic representations and spatial ability - while the lower part involves constructs in the semantic-linguistic domain, that is, relational processing and reading comprehension. Of note is that within both domains direct and indirect paths are hypothesized. Furthermore, a correlation between both basic abilities is captured in the path model.

While all separate relations in our proposed model have been previously investigated in earlier studies, the present study is unique as it combines the component skills and basic underlying abilities from both processing domains in one model. The results obtained from this study can broaden our knowledge of the factors that are important for word problem solving and can provide an interesting starting point for an effective word problem solving instruction.

The aim of the present study is twofold:

Investigate whether the component skills and basic abilities in the two processing domains explain unique variance in students' word problem solving skills.

Examine the direct and indirect (via the component skills) effects of the basic abilities on word problem solving.

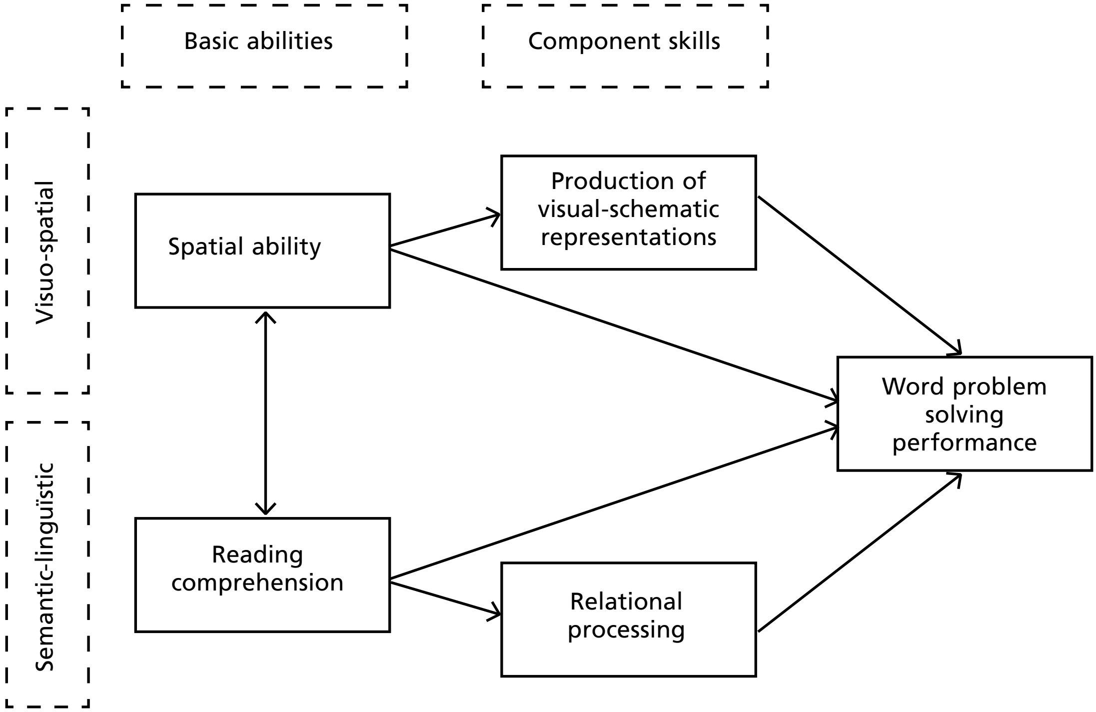


Figure 1. Path model with all hypothesized pathways

METHODS

Participants

The study contained data from 128 Dutch sixth grade students (64 boys, M_{age} = 11.73 years, SD_{age} = 0.43 years and 64 girls, M_{age} = 11.72 years, SD_{age} = 0.39 years) from eight elementary schools in The Netherlands. These eight schools were randomly drawn from a total of 20 elementary schools. Approximately 15 students of each of the eight elementary schools were selected on the basis of their proficiency score on the CITO Mathematics test (2008). The CITO Mathematics test is a nationwide standardized test (developed by the Institute for Educational Measurement) to follow students' general math ability during their elementary school career. On the basis of this test the students are equally divided in low, average and high math performers to obtain a representative sample. Parents provided written informed consent based on printed information about the purpose of the study.

Instruments and measurement procedure

The measurement instruments that were used in this study were administrated to the students by three trained independent research-assistants in two sessions of approximately 45 and 30 minutes.

Word problem solving performance

Word problem solving performance were examined with the Mathematical Processing Instrument (MPI), translated to Dutch. The MPI consisted of 14 word problems based on previous studies (Hegarty & Kozhevnikov, 1999; Van Garderen & Montague, 2003, see Appendix A). The internal consistency coefficient (Cronbach's alpha) of this instrument, measured in American participants, is .78 (Hegarty & Kozhevnikov, 1999). The Cronbach's alpha of the MPI in this study was .72. The word problems were printed on cards and presented in four different orders. All problems were read out loud to the students to control for differences in decoding skill. To prevent that the execution of the required arithmetic operations would be a determining factor in students' word problem solving, these operations were easy and could be solved by every student. Furthermore, students were allowed to solve each word problem within three minutes and during this time the experimenter did not speak to the student. To be sure that students had enough time to solve the word problems, a pilot study was conducted with five sixth grade students. The results of the pilot study showed that every student was able to solve each of the 14 items of the MPI within the required three minutes. The number of problems solved correctly was used as the dependent variable in the analysis.

Component skill in the visuo-spatial domain: Production of visual-(schematic) representations

After the three minutes of problem solving time, a short interview was held about the nature of the (mental) representation evoked by the word problem. The exact procedure of this interview is adapted from the study of Hegarty and Kozhevnikov (1999). We adjusted some questions of this interview procedure to make sure that children were not forced to make a visual representation, but used the strategy they preferred to solve the word problem (see Appendix B for the interview-format).

For each visual representation a score was obtained expressing whether the students had made a visual-schematic or a pictorial representation. These two representation categories are exemplified by the following word problem:

> ***[World problem 1]:***
> *A balloon first rose 200 meters from the ground, then moved 100 meters to the east, and then dropped 100 meters. It then traveled 50 meters to the east, and finally dropped straight on the ground. How far was the balloon from its original starting point?"*

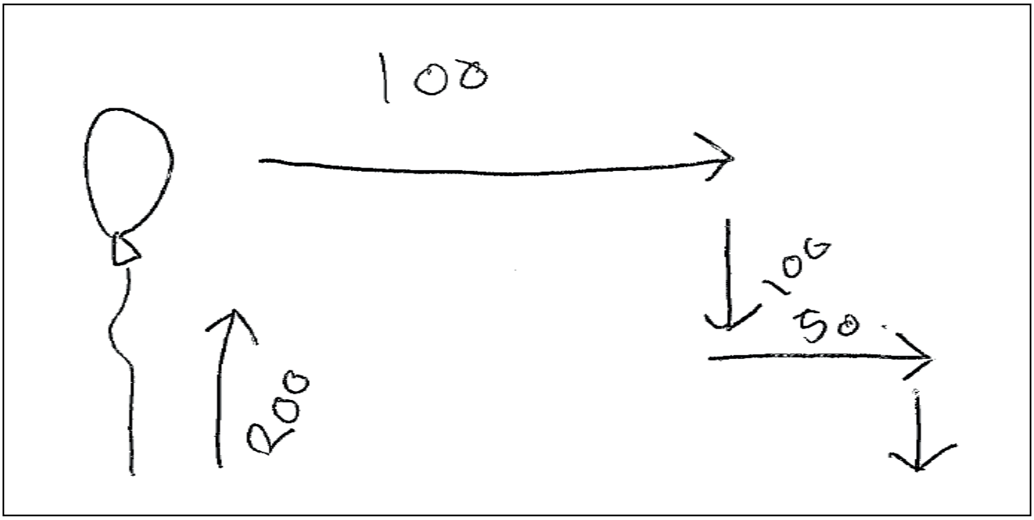


Figure 2. An example of a visual-schematic representation of word problem 1

A representation was coded as visual-schematic if students drew a diagram, used gestures showing spatial relations between elements in a problem in explaining their solution strategy, or reported a spatial image. Figure 2 shows an example of a visual-schematic representation.

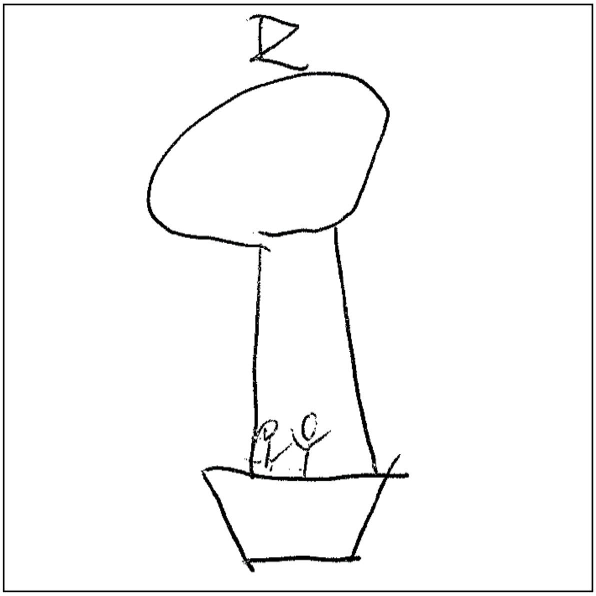


Figure 3. An example of a pictorial representation of word problem 1

A representation was coded as pictorial if the student drew an image of the objects and/or persons referred to in the problem, rather than the relations between them.
Figure 3 shows an example of a pictorial representation.
In total 612 representations were made by the students. All representations were coded by three independent coders. In the first coding session 32 representations were randomly selected and coded according to the two categories by all coders. The inter-rater reliability of these 32 coded representations was high (Cohen's Kappa (κ) = .88, Tabachnick & Fidell, 2006). Because the results of this coding session were good, the remaining representations were coded by all coders in the same way. Because we were only interested in the production of visual-schematic representations, the total number of visual-schematic representations made by each student was included in the analysis.

Component skill in the semantic-linguistic domain: Relational processing

To determine relational processing, i.e., the derivation of the correct relations between the solution-relevant elements from the text base of the word problem, we used the inconsistency task. The inconsistency task contained eight two-step compare problems consisting of three sentences, which were selected from the study of Hegarty et al. (1992) and translated into Dutch. The first sentence of each word problem was an assignment statement expressing the value of the first variable, that is, the price of a product at a well-known Dutch store or supermarket (e.g., At Albert Heijn a bottle of olive oil costs 4 euro). The second sentence contained a relational statement expressing the value of the second variable (i.e., the price of this product at another store or supermarket) in relation to the first (e.g., At Spar, a bottle of olive oil costs 3 euro more than at Albert Heijn). In the third sentence, the problem solver was asked to find a multiple of the value of the second variable (e.g., If you need to buy three bottles of olive oil, how much will you pay at Spar?). The answer to these word problems always involved first computing the value of the second variable (e.g., 4 + 3 = 7) and then multiplying this solution by the quantity given in the third sentence (e.g., 7 times 3 = 21). In this task, the consistency of the word problems was manipulated. Consistency refers to whether the relational term in the second sentence was consistent or inconsistent with the required arithmetic operation. A consistent sentence explicitly expressed the value of the second variable (V2) in relation to the first variable (V1) introduced in the prior sentence (At V2, product A costs N euro [more/ less] than at V1). An inconsistent sentence related the value of the second variable to the first by using a pronominal reference (This is N euro [more/ less] than at V2). Consequently, the relational term in a consistent word problem primed the appropriate arithmetic operation ('more than' when the required operation is addition, and 'less than' when the required operation is subtraction), and the relational term in an inconsistent word problem primed the inappropriate arithmetic operation ('more than' when the required operation is subtraction, and 'less than' when the required operation is addition). We controlled for difficulty in reading comprehension throughout the consistent and inconsistent word problems by balancing the number of unmarked ('more than') and marked ('less than') relational terms. As such, the relatively higher complexity that would have been introduced by an inconsistent item cannot be explained by any effect of markedness.

The numerical values used in the word problems were selected on basis of the following rules in order to control for the difficulty of the required calculations: (1) The answers of the first step of the operation were below 10, (2) The final answers were between the 14 and 40, (3) None of the first step or final answers contained a fraction of a number or negative number, (4) No numerical value occurred twice in the same problem, and (5) None of the (possible) answers resulted in 1. The numerical values used in consistent and inconsistent word problems were matched for magnitude.

For the analysis, we looked at the students' accuracy (i.e., the amount of correct answers) on the inconsistent word problems. The internal consistency coefficient of this measure in the present study was high (Cronbach's alpha = .90).

Basic ability in the visuo-spatial domain: Spatial ability

The Paper Folding task (retrieved from The Kit of Factor-Referenced Cognitive Tests; Ekstrom, French, Harman, & Derman, 1976) and the Picture Rotation task (Quaiser-Pohl, 2003) were standardized tasks used to measure spatial visualization. In the Paper Folding task, children were asked to imagine the folding and unfolding of pieces of paper. In each problem in the test, some figures were drawn at the left of a vertical line and there were others drawn at the right. The figures at the left of the vertical line represented a square piece of paper being folded. On the last of these figures one or two small circles were drawn to show where the paper had been punched. Each hole was punched throughout the thicknesses of paper at that point. One of the five figures at the right of the vertical line showed where the holes would be located when the paper was completely unfolded. Children had to decide which one of these figures was correct. This task took 6 minutes and had a sufficient internal consistency coefficient in the present study (Cronbach's alpha = .70). Figure 4 shows one of the 20 test items of the Paper Folding task.

In the Picture Rotation task children were asked to rotate a non-manipulated picture of an animal at the left of a vertical line. The three pictures at the right of the vertical line showed the rotated and/or mirrored image of that same animal. One of these three pictures was only rotated; two of these pictures were both rotated and mirrored. Children had to decide which of the three pictures was only rotated. This task took 1.5 minutes and its internal consistency coefficient in this study was high (Cronbach's alpha = .93). Figure 5 shows one of the 30 test items of the Picture Rotation task.

To obtain a general measure of spatial ability, the raw scores of

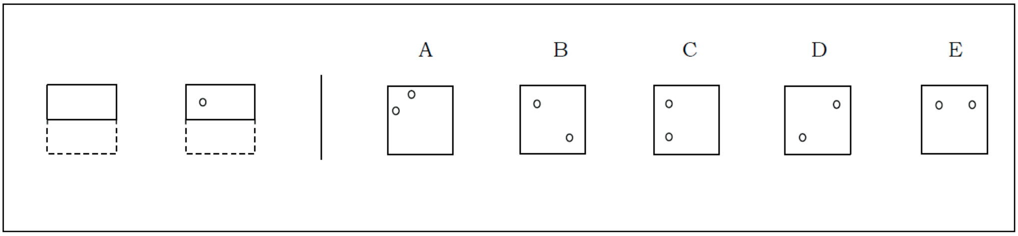


Figure 4. The Paper Folding task (Ekstrom, French, Harman, & Derman, 1976)

Figure 5. The Picture Rotation task (based on Quaiser-Pohl, 2003)

each of the spatial ability tasks were rescaled into a z-score. Subsequently, these z-scores were aggregated into an average z-score (*M* = .00, *SD* = .84).

Basic ability in the semantic-linguistic domain: Reading comprehension

The standardized CITO (Institute for Educational Measurement) Reading comprehension test (2010) was used to measure children's reading comprehension skills. Each test contains two modules, each consisting of a text and 25 multiple choice questions. The questions pertained to the word, sentence or text level and tapped both the text base and situational representation that the reader constructed from the text (e.g., Kintsch, 1988). Students' raw test scores on the 50 items were rescaled to a normed proficiency score. The proficiency scores (*M* = 42.06, *SD* = 14.06) made it possible to compare the results of the reading comprehension test with other versions of this test from other years. The internal consistency coefficient of this test in sixth grade students was high with a Cronbach's alpha of .89 (Weekers, Groenen, Kleintjes & Feenstra, 2011).

Data analysis

Path analyses using MPlus Version 4 (Muthén & Muthén, 2006) were performed to examine if the hypothesized model fitted the data.

The standard Maximum Likelihood (ML) method of estimating free parameters in structural equation models was used to asses model fit. In this procedure, a non-significant chi-square (X^2), a root-mean-square error of approximation (RMSEA) under .05, and a Comparative Fit Index (CFI) value above .95 together indicate a strong fit of the data with the model, while a RMSEA value under .08 and a CFI above .90 indicate an adequate fit (Hu & Bentler, 1999; Kline, 2005). Two path analyses were performed to examine the path model which fitted the data best.

First, the complete hypothesized model (see Figure 1) was tested, including the two component skills, their underlying basic abilities and their connection with word problem solving performance. This model was considered as the baseline model in the analyses. To examine the presence of mediation by the two component skills, the baseline model, including both direct and indirect effects, was tested against a second model containing only the direct effects (see Figure 6). If the second model had worse fit indices compared to the baseline model - based on a significant increase of the chi-square statistic (CMIN) -, mediation effects were present (Kline, 2005). The degree in which the effect is reduced is an indicator of the potency of the mediator (Preacher & Hayes, 2008). The value of this indirect effect was calculated with the following formula[1]:

$$B_{indirect} = B_{(path\ a)} * B_{(path\ b)}$$

followed by

$$B_{indirect} / B_{(total)}$$

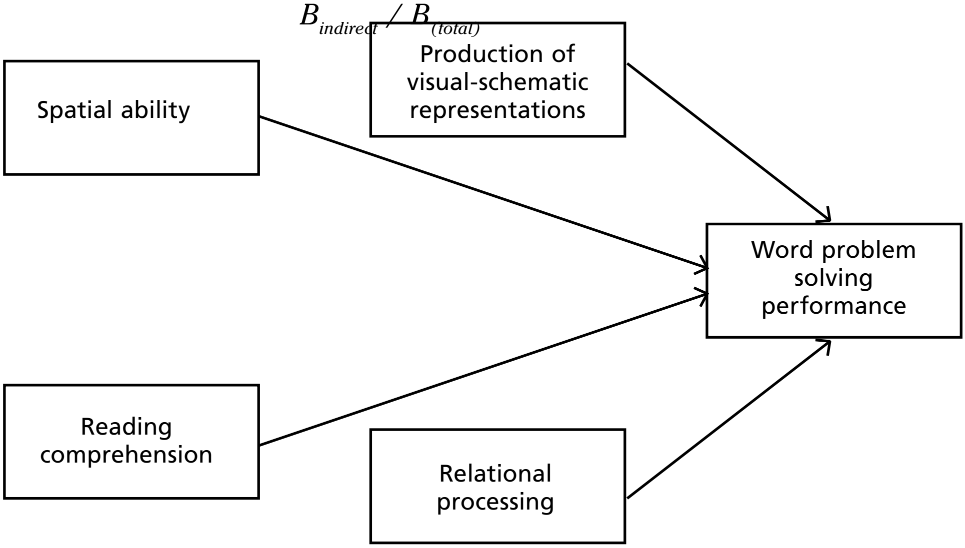


Figure 6. Model 2, including only the direct effects

2

RESULTS

Descriptive statistics

Table 1 presents the means and standard deviations of, and the correlations between, the five measures of this study. This table shows that the correlations between the measures are moderate, except for two correlations. The correlation between the production of visual-schematic representations and students' relational processing skills is negligible (r = .08). On the other hand, the correlation between spatial ability and word problem solving is strong (r = .59).

Table 1. *Intercorrelations, means, standard deviations for all measures*

Measure	1.	2.	3.	4.	5.	*M*	*SD*
1. Word problem solving performance	-					6.68	2.87
2. Relational processing	.37**	-				2.94	1.27
3. Production of visual-schematic representations	.44**	.08	-			2.13	2.45
4. Reading comprehension	.48**	.33**	.26**	-		42.06	14.06
5. Spatial ability (z-score)	.59**	.24*	.31**	.43**	-	.00	.84

** p < .01, ** p < .001*

Examining the complete hypothesized path model, including direct and indirect effects

The hypothesized path model is assessed with Maximum Likelihood estimation. The fit indices for this baseline model are good: X^2 (3) = 3.50, p = .32, CFI = .99 and RMSEA = .04.

Figure 7 shows the graphical representation of the hypothesized model, including the standardized parameter estimates. Table 2 shows the complete parameter estimates of the model. The path

[1.] B (path a): the unstandardized coefficient from spatial ability/reading comprehension to the production of visual-schematic representations/ relational processing.
B (path b): the unstandardized coefficient from the production of visual-schematic representations/ relational processing to word problem solving performance.
B (total): direct relation between spatial ability/ reading comprehension and word problem solving performance.

analysis shows that 49.1% (R^2 = .491) of the variance in word problem solving performance is explained by the production of visual-schematic representations (β = .27, p < .001), spatial ability (β = .39, p < .001), students' relational processing skills (β = .21, p < .001) and reading comprehension (β = .18, p < .05). This is a large effect size (Tabachnick & Fidell, 2006). Spatial ability (β = .31, p < .001) explains 9.6% (R^2 = .096) of the variance in the production of visual-schematic representations and reading comprehension (β = .34, p < .001) explains 11.2% (R^2 = .112) of the variance in relational processing. These two effect sizes can be categorized as medium (Tabachnick & Fidell, 2006). Finally, in line with our expectations, the correlation between spatial ability and reading comprehension is significant (r = .44, p < .001).

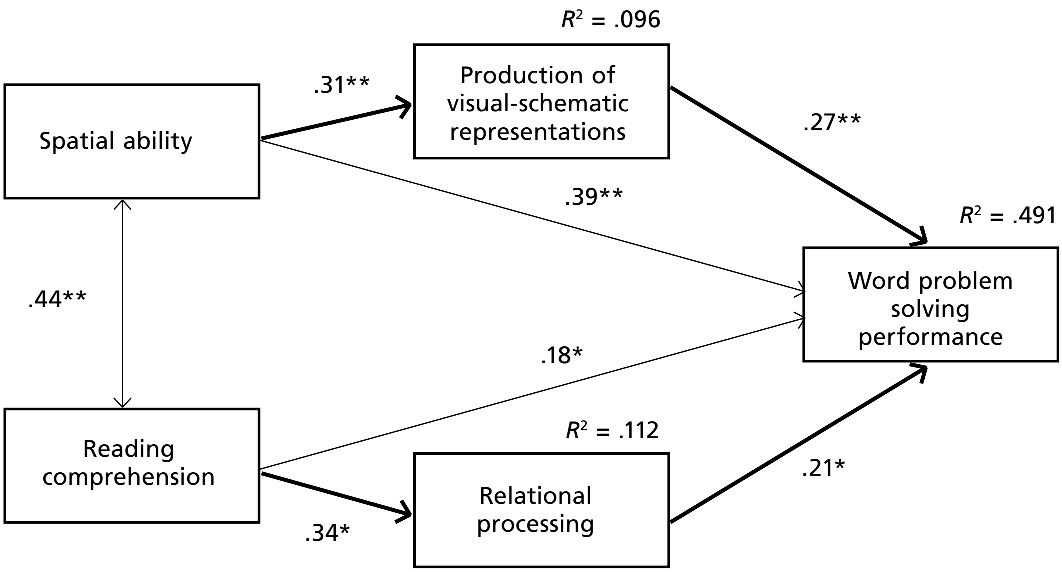


Figure 7. Hypothesized model, including the standardized estimates of the variables influencing word problem solving performance, the significant pathways are indicated with an asterisk, * p < .05, ** p < .001

Table 2. *Results from the path analysis, including unstandardized and standardized parameter estimates of the direct pathways*

Pathway		*B*	*SE*	β
Visual-schematic representations	Word problem solving performance	0.31**	.08	.27
Spatial ability	Visual-schematic representations	0.90**	.24	.31
Spatial ability	Word problem solving performance	1.30**	.25	.39
Relational processing	Word problem solving performance	0.45**	.15	.21
Reading comprehension	Relational processing	0.03**	.01	.34
Reading comprehension	Word problem solving performance	0.04**	.02	.18

* $p < .05$, ** $p < .001$

Testing mediation

In order to test the existence of mediation by the two component skills, the baseline model is tested against a second model, including only the direct effects (see Figure 6). If the baseline model fits the data better, mediation exists and both direct and indirect effects are present.

Also the second path model is assessed with Maximum Likelihood estimation. This model has a bad model fit: $X^2(6) = 54.22$, $p < .001$, CFI = .62, RMSEA = .25. Compared to the baseline model, the second model fits the data less adequately: CMIN (3) = 50.72, $p < .001$. This finding indicates that the model with both direct and indirect effects fits the data better than the model with only the direct effects. This means that – at least partial – mediation occurs. Thus, in line with our expectations, spatial ability and reading comprehension have both a direct and indirect relation with word problem solving. The value of the indirect effect of spatial ability can be calculated as follows:

$$B_{indirect} = B_{(a)} * B_{(b)} = 0.90 \text{ x } 0.31 = 0.279, \text{ and}$$
$$B_{indirect} / B_{total} = 0.279 / 1.30 = 0.21.$$

The value of the indirect effect of reading comprehension can be calculated in the same way:

$$B_{indirect} = B_{(a)} * B_{(b)} = 0.03 \text{ x } 0.45 = 0.014, \text{ and}$$
$$B_{indirect} / B_{total} = 0.014 / 0.04 = 0.34.$$

Thus, the production of visual-schematic representations explains 21% of the relation between spatial ability and word problem solving performance. On the other hand, relational processing explains 34% of the relation between reading comprehension and word problem solving performance.

DISCUSSION

This study examined the importance of two component skills - that is, the production of visual-schematic representations and relational processing - as well as their basic underlying abilities - that is, spatial ability and reading comprehension - for successful word problem solving (e.g., Hegarty & Kozhevnikov, 1999; Van der Schoot et al., 2009; Van Garderen, 2006). The uniqueness of this study lies in the fact that it is the first study that examined these constructs, tapping different processing domains (i.e., visuo-spatial and semantic-linguistic), in one hypothesized path model. Moreover, both direct and indirect effects of spatial ability and reading comprehension were investigated.

In line with previous research, the results of the path analyses showed that the two component skills (i.e., the production of visual-schematic representations and relational processing) explained unique variance in students' word problem solving performance (Hegarty & Kozhevnikov, 1999; Van der Schoot et al., 2009; Van Garderen, 2006). With respect to the direct and indirect effects of the component skills' underlying basic abilities, this study showed that 21% of the relation between spatial ability and word problem solving was explained by the production of visual-schematic representations. Furthermore, 34% of the relation between reading comprehension and word problem solving was explained by relational processing. Overall, the path model explained 49% of the variance in word problem solving.

Limitations

Two limitations of this study should be mentioned. The first limitation covers the instrument to determine the nature of the visual representations that were made. After each item of the MPI a short interview was held to establish (1) whether a visual representation was made and (2) whether this representation was pictorial or visual-schematic in nature. Although the most visual representations were made on paper during the task (*M* = 3.58), some representations were made mentally (*M* = 1.20). This means that, when the students were asked to describe and draw the pictures they had in their mind while solving the problem (see the interview procedure described in Appendix B), careful observations from the test assistants were essential to disclose these mental representations. Yet, they could not be completely sure if the representation drawn on a piece of paper (asked retrospectively) was an exact copy of the representation that was made in the head of the child during task performance. Videotapes of each test administration were used to facilitate the process of signaling the mental visual representations.

The second limitation pertains to the correlational nature of the data, which makes it impossible to draw conclusions about any causal relationships between basic abilities, component skills and word problem solving performance. The results of this study only show that these variables are associated with each other. Future experimental studies in which the component skills and basic abilities are manipulated, should make it possible to draw stronger conclusions concerning causal relationships between the processes which are involved in word problem solving.

Directions for future research

In future research the production of visual-schematic representations and relational processing should be examined in more detail to draw stronger conclusions. For example, we suggest to examine the production and characteristics of visual representations in the light of individual differences, i.e., differences between low, average and high achievers and/or boys and girls. Several authors have found differences between low, average and high achievers in their production of visual representations and word problem skills (e.g., Van Garderen, 2006; Van Garderen & Montague, 2003). In addition, the scientific literature gives indications that boys have better spatial skills than girls (e.g., Casey, Nuttall, Benbow, & Pezaris, 1995; Casey et al., 1997). Therefore, the production of visual-schematic representations might be a more naturally representation strategy for boys compared to girls.

The findings of this study are also interesting for educational practice. Follow-up studies should examine the effects of interventions in which elementary and secondary school students are taught to systematically build visual-schematic (mental) representations during math problem solving. Several studies have shown that it is more effective to teach children to make their own representations, instead of providing representations in advance (e.g., in the form of illustrations, Van Dijk, Van Oers, & Terwel, 2003; Van Dijk, Van Oers, Van den Eeden, & Terwel, 2003). The use of schema-based instruction in word problem solving (e.g., Jitendra, DiPipi, Perron-Jones, 2002; Jitendra & Hoff, 1996), where students have to map the information onto a relevant schematic diagram after identifying the problem type, might therefore be a less effective manner to increase word problem solving performance. Besides teaching students to produce visual-schematic representations, one should teach students to derive the correct relations between solution-relevant elements from the text base of the word problem. As reading comprehension is found to be essential for this component skill, word problem instruction should not only focus on the strategic aspects of word problem solving, but also on the more semantic-linguistic aspects.

APPENDIX 2.A WORD PROBLEMS OF THE MATHEMATICAL PROCESSING INSTRUMENT

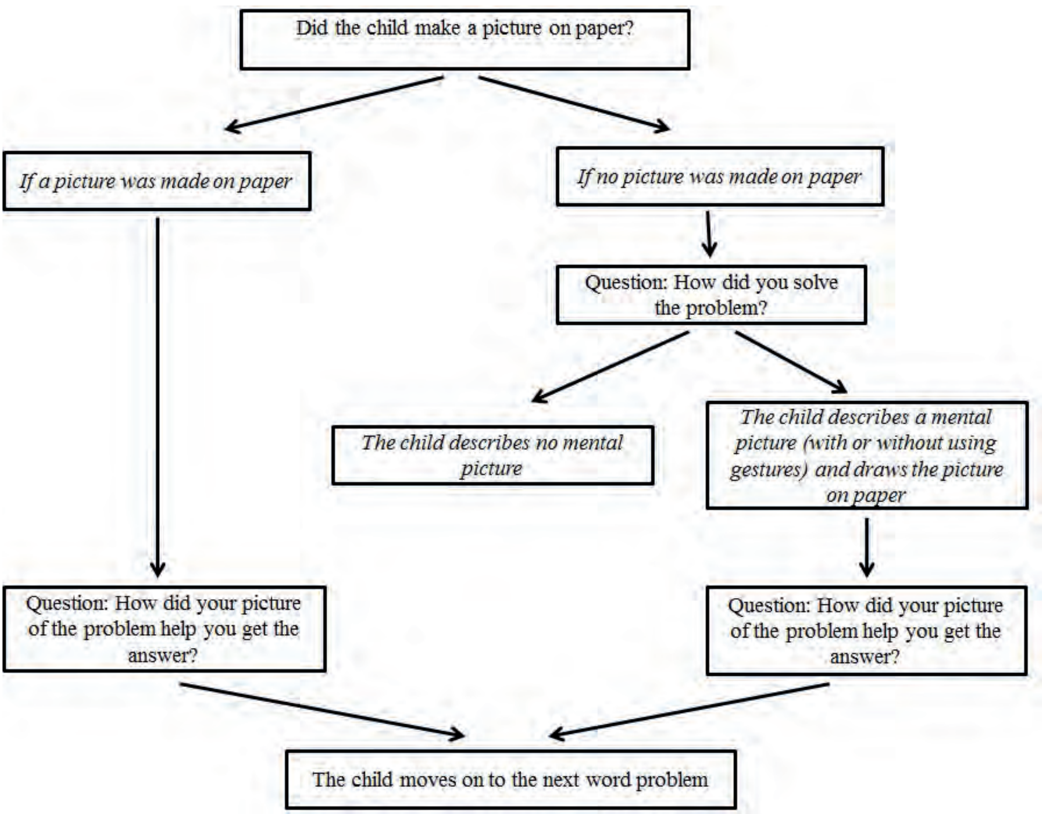
The word problems of the Mathematical Processing Instrument (Hegarty & Kozhevnikov, 1999):

1. At each of the two ends of a straight path, a man planted a tree and then every 5 meters along the path he planted another tree. The length of the path is 15 meters. How many trees were planted?
2. On one side of a scale there is a 1 kg weight and half a brick. On the other side there is one full brick. The scale is balanced. What is the weight of the brick?
3. A balloon first rose 200 meters from the ground, then moved 100 meters to the east, then dropped 100 meters. It then traveled 50 meters to the east, and finally dropped straight to the ground. How far was the balloon from its original starting point?
4. In an athletics race, Jim is four meters ahead of Tom and Peter is three meters behind Jim. How far is Peter ahead of Tom?
5. A square (A) has an area of 1 square meter. Another square (B) has sides twice as long. What is the area of B?
6. From a long stick of wood, a man cut 6 short sticks, each 2 feet long. He then found he had a piece of 1 foot long left over. Find the length of the original stick.
7. The area of a rectangular field is 60 square meters. If its length is 10 meters, how far would you have traveled if you walked the whole way around the field?
8. Jack, Paul and Brian all have birthdays on the 1st of January, but Jack is one year older than Paul and Jack is three years younger than Brian. If Brian is 10 years old, how old is Paul?
9. The diameter of a tin of peaches is 10 cm. How many tins will fit in a box 30 cm by 40 cm (one layer only)?
10. Four young trees were set out in a row 10 meters apart. A well was situated beside the last tree. A bucket of water is needed to water two trees. How far would a gardener have to walk altogether if he had to water the four trees using only one bucket?
11. A hitchhiker set out on a journey of 60 miles. He walked the first 5 miles and then got a lift from a lorry driver. When the driver dropped him he still had half of his journey to travel. How far had he traveled in the lorry?
12. How many picture frames 6 cm long and 4 cm wide can be made

from a piece of framing 200 cm long?
13. On one side of a scale there are three pots of jam and a 100 g weight. On the other side there are a 200 g and a 500 g weight. The scale is balanced. What is the weight of a pot of jam?
14. A ship was North-West. It made a turn of 90 degrees to the right. An hour later it made a turn through 45 degrees to the left. In what direction was it then traveling?

APPENDIX 2.B. INTERVIEW PROCEDURE MATHEMATICAL PROCESSING INSTRUMENT

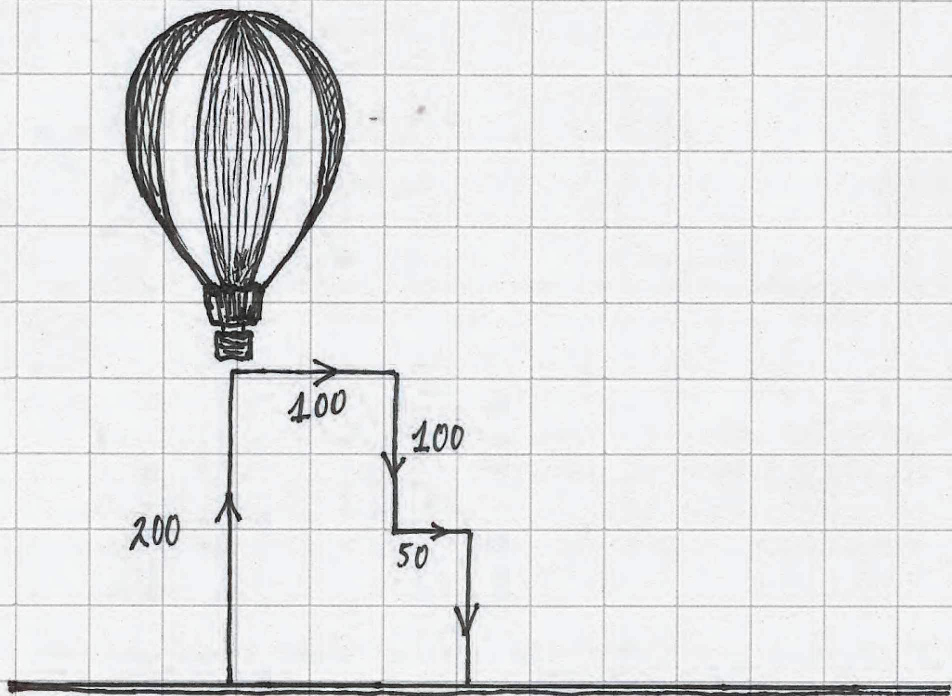
Interview procedure which was followed after each word problem on the Mathematical Processing Instrument.



3 Word problem solving in contemporary math education:

A plea for semantic-linguistic skills training

Anton J. H. Boonen, Björn B. de Koning, Jelle Jolles, & Menno van der Schoot
(under review)



A balloon first rose 200 meters from the ground, then moved 100 meters to the east, then dropped 100 meters. It then traveled 50 meters to the east, and finally dropped straight to the ground. How far was the balloon from its original starting point?

ABSTRACT

This study pointed out that word problem solving instruction in Realistic Math Education (RME) may have insufficient attention for teaching semantic-linguistic skills to handle semantic complexities in word problems. We investigated the performances of 80 sixth grade students, classified as successful and less successful word problem solvers based on a standardized mathematics test from the RME curriculum, on word problems that ask for both sophisticated representation skills and semantic-linguistic skills. The results showed that even successful word problem solvers had a low performance on semantically complex word problems, despite adequate performance on semantically less complex word problems. Less successful word problem solvers had low scores on both semantically simple and complex word problems. Results showed that reading comprehension was only related to the successful word problem solvers' performance on semantically complex word problems. On the basis of this study, we concluded that semantic-linguistic skills should be given a (more) prominent role during word problem solving instruction in RME.

INTRODUCTION

In the last decades, mathematical word problem solving has gained much attention from both researchers and educational practitioners (Boonen, Van der Schoot, Van Wesel, De Vries, & Jolles, 2013; Campbell, 1992; Depaepe, De Corte, & Verschaffel, 2010; Hegarty, Mayer, & Monk, 1995; Hajer, 1996; Hickendorff, 2011, 2013; Moreno, Ozogul, & Reisslein, 2011; Swanson, Lussler & Orosco, 2013). Mathematical word problems refer to mathematical exercises that present relevant information on a problem as text, rather than in the form of mathematical notation (Rasmussen & King, 2000; Timmermans, Van Lieshout, & Verhoeven, 2007). Hence, effectively solving a mathematical word problem is assumed to depend not only on students' ability to perform the required mathematical operations, but also on the extent to which they are able to accurately understand the text of the word problem (Hegarty et al., 1995; Jitendra & Star, 2012; Lewis & Mayer, 1987; Van der Schoot, Baker-Arkema, Horsley, & Van Lieshout, 2009). Both of these aspects are related in such a way that developing a deeper understanding of the text of the word problem serves as a crucial step before the correct mathematical computations can be performed. Hence, a key challenge for word problem solvers is to get an adequate understanding of the problem statement (Boonen et al., 2013; Lee, Ng, & Ng, 2009; Thevenot, 2010).

Two individual skills are relevant in this regard. First, an important factor contributing to a deeper understanding of the text of the word problem is the ability to construct a rich and coherent mental representation containing all (the relations between the) solution-relevant elements that are derived from the text base of the word problem (De Corte, Verschaffel, & De Win, 1985; Hegarty et al., 1995; Pape, 2003). That is, word problem solvers have to use a problem-model strategy in which they translate the problem statement into a qualitative mental representation of the problem situation hidden in the text (Pape, 2003; Van der Schoot et al., 2009). This mental representation subsequently allows them to make a solution plan and execute the required mathematical operations. Although successful word problem solvers appear to employ such a representational strategy, less successful problem solvers often adopt an impulsive, superficial direct translation strategy, in which they only focus on selecting the presented numbers that, in turn, form the basis for their mathematical calculations (Hegarty et al., 1995; Verschaffel, De Corte, & Pauwels, 1992).

The second important individual skill in word problem solving success substantiated by research evidence is the influence of a student's reading comprehension abilities (Boonen et al., 2013; Pape, 2004; van der Schoot et al., 2009). It has been suggested that reading comprehension abilities are especially helpful in dealing with semantic-linguistic word problem characteristics such as the sequence of the known elements in the text of the word problem, the degree to which the semantic relations between the given and unknown quantities of the problem are made explicit, and the relevance of the information in the text of the word problem (De Corte et al., 1985; De Corte, Verschaffel, & Pauwels, 1990; Marzocchi, Lucangeli, De Meo, Fini, & Cornoldi, 2002; Verschaffel et al., 1992).

Moreover, semantic-linguistic skills appear to be more important in overcoming such textual complexities than being able to apply sophisticated representation strategies (De Corte et al., 1985; 1990). This might explain why the use of a sophisticated mental representation strategy is not sufficient in all circumstances. That is, word problems containing semantically complex features require both accurate representation skills and reading comprehension skills, whereas for word problems with a lower semantic-linguistic complexity, sophisticated representational skills might be sufficient.

Teaching word problem solving in Realistic Mathematics Education

These findings suggest that, to teach students how to effectively solve mathematical word problems, sophisticated representational skills and semantic-linguistic skills should both be part of the mathematics education program. Particularly, paying attention to semantic-linguistic skills is relevant to help students improve their word problem solving success, as word problems become semantically more complex as students progress in their educational career, for example, when they make the transition to secondary education. Word problems offered in secondary school subjects like geometry, physics and biology, include more verbal information and generally contain more complex semantic-linguistic text features (Helwig, Rozek-Tedesco, Tindal, Heath & Almond, 1999; Silver & Cai, 1996).

The Netherlands, like many other countries, currently places great emphasis on the teaching of word problem solving in contemporary mathematics education (Elia, Van den Heuvel-Panhuizen, & Kovolou, 2009; Ruijssenaars, Van Luit, & Van Lieshout, 2004;). The teaching of mathematics in the Netherlands takes place within the context of a domain-specific instructional approach, called Realistic

Mathematics Education (RME, Van den Heuvel-Panhuizen, 2003), where the process of mathematical word problem solving plays an important role (Barnes, 2005; Hickendorff, 2011; Prenger, 2005; Van de Boer, 2003; Van den Heuvel-Panhuizen, 2005). Studies investigating the educational practice of RME show that the teaching of sophisticated representation strategies receives a lot of attention in word problem solving instruction (Elia et al., 2009; Van Dijk, Van Oers, & Terwel, 2003; Van den Heuvel-Panhuizen, 2003). However, the training of semantic-linguistic skills appears to be less explicitly trained in the instructional practice of RME, in spite of its proven importance in previous studies (e.g., De Corte et al., 1985, 1990; Hegarty et al., 1992). This is presumably because teachers may underestimate or are not aware of the importance of semantic-linguistic skills for solving word problems (Hajer, 1996; Van Eerde, 2009). Thus, the current approach to teaching word problem solving appears to emphasize the development of representation skills, but seems to pay less attention to the role of semantic-linguistic skills.

In this respect, educational practice regarding teaching word problem solving does not seem to be aligned with what is currently known from research about the factors involved in effective word problem solving. This study aims to provide evidence for the claim that semantic-linguistic skills receive little attention in word problem solving instruction in RME, thereby identifying an important area of concern with respect to the way word problem solving is currently taught in the Netherlands. To test this claim, we compared students' performance on word problems obtained while following the RME curriculum to their performances on an independent word problem solving task. First, we classified students as successful or less successful word problem solvers with the help of a mathematics test that is part of the RME curriculum, viz., the CITO (Institute for Educational Measurement) Mathematics test. This test can be considered a method-specific (i.e., RME-specific) mathematics test of students' word problem solving performance, as it builds upon the currently used instructional method for word problem solving. Hence, this test reflects the skills that students learn in the RME classroom, in order to solve word problems (Doorman, Drijvers, Dekker, Van den Heuvel-Panhuizen, De Lange, & Wijers, 2007; Hickendorff, 2011). Second, we examined students' performance on an independent word problem solving test, which contained either word problems that could be solved by only using a sophisticated mental representation strategy, or word problems that required them to also use their semantic-linguistic skills.[2]

Based on the assumption that word problem solving instruction in RME pays little attention to handling the semantic-linguistic features of the problem text, we hypothesized that it is likely that a key aspect that differentiates successful from less successful word problem solvers concerns their ability to construct a sophisticated mental representation of the problem text. Previous studies have shown that asking students to solve compare problems, especially inconsistent compare problems (see Example 1), is a suitable method for investigating whether or not they use a sophisticated representation strategy (e.g., Pape, 2003; Van der Schoot et al., 2009).

> ***[Example 1]***
> *At the grocery store, a bottle of olive oil costs 7 euro.*
> *That is 2 euro more than at the supermarket.*
> *If you need to buy 7 bottles of olive oil, how much will it cost at the supermarket?*

In this example, the translation process requires the identification of the pronominal reference 'that is' as the indicator of the relation between the value of the first variable ('the price of a bottle of olive oil at the grocery store') to the second ('the price of a bottle of olive oil at the supermarket'). This identification is necessary to become cognizant of the fact that, in an inconsistent compare problem, the relational term 'more than' refers to a subtraction operation rather than to an addition operation. So, inconsistent word problems create greater cognitive complexity than consistent word problems, requiring students to ignore the well-established association between *more* with increases, and addition and *less* with decreases and subtraction (Schumacher & Fuchs, 2012). Empirical evidence corroborates this interpretation by showing that word problem solvers make more (reversal) errors on inconsistent than on consistent word problems (i.e., consistency effect, Lewis & Mayer, 1987; Pape, 2003; Van der Schoot et al., 2009). Especially students who fail to build a high-quality mental representation of the problem statement, and thus immediately start calculating with the given numbers and relations, seem to be less

[2]This procedure provides an advantage over prior studies of, among others, Hegarty et al. (1995), Pape (2003), and van der Schoot et al. (2009), which typically use the main dependent variable of the study (i.e., problem solving success) as an outcome measure as well as a means to classify students into successful and less successful word problem solvers. The classification used in the present study, on the other hand, is based on an external, well-established measure of mathematical word problem solving, which is independent of the main dependent variable of the study (i.e., word problem solving success). This allows us to make more meaningful group comparisons.

successful on inconsistent word problems (Hegarty et al., 1995).

In the present study, we expected neither successful nor less successful problem solvers to experience difficulties with solving consistent compare word problems. However, we did assume that successful word problem solvers in the RME curriculum would experience less difficulties with correctly solving inconsistent compare problems as a result of their use of a sophisticated representation strategy (acquired during word problem solving instruction in RME), than less successful problem solvers who employ a more superficial problem solving approach (Van der Schoot et al., 2009; Verschaffel et al., 1992).

It is important to keep in mind that this only holds for consistent and inconsistent compare problems with a low semantic complexity; that is, problems that only tap into students' ability to construct a sophisticated mental representation. If the semantic complexity of compare problems increases, even students classified as successful word problem solvers (according to our classification based on the RME instruction) may come to experience difficulties with correctly solving inconsistent compare problems. In this case, correctly solving a word problem requires students to use both mental representational skills and semantic-linguistic skills, while word problem solving instruction in RME has provided students only with considerable training in the first of these two skills.

A relatively well-studied and accepted way to increase the semantic complexity of (inconsistent) compare problems is to manipulate the relational term (Lewis & Mayer, 1987; Van der Schoot et al., 2009). That is, we can increase the semantic complexity of a word problem by making a distinction between an unmarked ('more than'), and a marked ('less than'), relation term. Research has shown that students find it easier to convert the unmarked relational term 'more than' into a subtraction operation than the marked relational term 'less than' into an addition operation (Clark, 1969; Kintsch, 1998; Lewis & Mayer, 1987; Pape, 2003; Van der Schoot et al., 2009). The difficulties experienced with solving marked inconsistent word problems lie in the fact that these problems draw on students' use of a sophisticated representation strategy as well as on their semantic-linguistic skills. As the effect of semantic-linguistic complexity only starts to play a role when the problem statement has been mentally represented accurately, the influence of semantic-linguistic skills is restricted to the group of successful problem solvers. So, although our group of successful word problem solvers may use a sophisticated representation strategy, the lack of attention to semantic-linguistic skills in the educational practice of RME is likely to cause them to experience difficulties with correctly solving (semantically complex) marked inconsistent word problems.

According to several researchers, the extent to which successful word problem solvers might be able to overcome difficulties with correctly solving marked inconsistent word problems is related to their semantic-linguistic skills (e.g., Lee, Ng, Ng, & Lim, 2004; Van der Schoot et al., 2009). Translating a marked relational term like 'less than' into an addition operation is found to be closely associated with a general measure of semantic-linguistic skills, and with reading comprehension in particular (Lee et al., 2004; Van der Schoot et al., 2009). This suggests that reading comprehension skills, together with sophisticated representation skills, might be necessary to deal with semantically complex word problems. The present study therefore also takes into account students' general reading comprehension ability.

In sum, the present study aimed to test the claim that the current Dutch instructional approach used in RME pays limited attention to the semantic-linguistic skills that allow students to handle linguistic complexities in a word problem. To this end, we tested the following hypotheses:

We hypothesized that, as a result of difficulties with constructing a coherent mental representation of word problems, less successful word problem solvers in the RME curriculum would make more errors on both unmarked and marked inconsistent word problems than on unmarked and marked consistent word problems

We hypothesized that, as a result of paying insufficient attention to semantic-linguistic skills in the teaching of word problem solving, successful word problem solvers in the RME curriculum would experience difficulties with solving semantically complex, marked inconsistent word problems, but not with solving semantically less complex, unmarked, inconsistent word problems.

3. We hypothesized that, as a result of the alleged relation between reading comprehension ability and the ability to overcome the semantic-linguistic complexities of a word problem, a positive relation for successful problem solvers exists between reading comprehension ability and the number of correctly solved marked inconsistent word problems.

METHODS

Selection of participants

Data from 80 Dutch sixth-grade students (42 boys, M_{age} = 11.72 years, SD_{age} = 0.39 years and 38 girls, M_{age} = 11.71 years, SD_{age} = 0.41 years) from eight elementary schools in the Netherlands were collected. These students were almost equally divided in two groups (by means of the median split method) on the basis of their score on the CITO Mathematics test (2008). This selection procedure resulted in a group of less successful word problem solvers (N = 41) and a group of successful word problems solvers (N = 39). The CITO Mathematics test is a nationwide standardized test that reflects the way in which word problem solving is instructed in Realistic Mathematics Education. The test contains elements like *mental arithmetic* (addition, subtraction, multiplication and division), *complex applications* (problems involving multiple operations) and *measurement and geometry* (knowledge of measurement situations), all of which are offered as mathematical word problems. The internal consistency of this test was high (Cronbach's alpha = .95, Janssen, Verhelst, Engelen & Scheltens, 2010).

Parents provided written informed consent based on printed information about the purpose of the study.

Instruments and procedure

The two measurement instruments that were used in this study were administrated to the students by three trained independent research assistants in a session of approximately 45 minutes.

Inconsistency task

The inconsistency task contained eight two-step compare problems that were selected from the study of Hegarty et al. (1992) and translated into Dutch. All of the word problems consisted of three sentences. The first sentence of each compare problem was an assignment statement expressing the value of the first variable, namely the price of a product at a well-known Dutch store or supermarket (e.g., At Albert Heijn a bottle of olive oil costs 4 euro). The second sentence contained a relational statement, expressing the value of the second variable (i.e., the price of this product at another store or supermarket) in relation to the first (e.g., At Spar, a bottle of olive oil costs 3 euro more than at Albert Heijn). In the third sentence, the problem solver was asked to find a multiple of the value of the second variable (e.g., If you need to buy three bottles of olive oil, how much will you pay at Spar?). The answer to these compare problems always involved first computing the value of the second variable (e.g., 4 + 3 = 7), and then multiplying this solution by the quantity given in the third sentence (e.g., 7 times 3 = 21).

The eight compare problems were separated in four different word problem types by the crossing of two within-subject factors: *consistency* (consistent vs. inconsistent) and *markedness* (unmarked vs. marked). Consistency referred to whether the relational term in the second sentence was consistent or inconsistent with the required arithmetic operation. A consistent sentence explicitly expressed the value of the second variable (At Spar a bottle of olive oil costs 3 euro [more/less] than at Albert Heijn) introduced in the prior sentence (At Albert Heijn a bottle of olive oil costs 4 euro). An inconsistent sentence related the value of the second variable to the first by using a pronominal reference (That is 3 euro [more/ less] than at Albert Heijn). Consequently, the relational term in a consistent compare problem primed the appropriate arithmetic operation ('more than' when the required operation is addition, and 'less than' when the required operation is subtraction). The relational term in an inconsistent compare problem primed the inappropriate arithmetic operation ('more than' when the required operation is subtraction, and 'less than' when the required operation is addition). Markedness expressed the semantic complexity of the relational term. A marked relational term (i.e., less than) is semantically more complex than an unmarked relational term (i.e., more than).

The stimuli were arranged in four material sets. Each participant was presented with eight word problems, two from each word problem type. The order in which the word problems were presented in each set was pseudorandomized. Each set was presented to 20 participants. Across sets and across participants, each word problem occurred equally often in the unmarked/consistent, marked/ consistent, unmarked/inconsistent and marked/ inconsistent version to ensure full combination of conditions and materials. Across word problems, we controlled for the difficulty of the required calculations, and for the number of letters in the names of the variables (i.e., stores) and products. To ensure that the execution of the required arithmetic operations would not be a determining factor in students' word problem solving performance, the operations were selected on the basis of the following rules: (1) the answers to the first step of the operation were below 10; (2) the final answers were between 14 and 40; (3) none of the first steps or final answers contained a fraction of a number or negative number; (4) no numerical value occurred twice in the same problem; and (5) none of the (possible) answers were 1. The numerical values used in consistent and inconsistent problems of each word problem type were matched for magnitude (see Van der Schoot et al., 2009).

For the analyses, we looked at students' accuracy (i.e., the amount of correct answers) on each of the four word problem types: (1) unmarked/consistent; (2) marked/consistent; (3) unmarked/inconsistent; and (4) marked/inconsistent. The internal consistency of this measure in the present study was high (Cronbach's alpha = .90).

Reading comprehension

The (Grade 6 version of the) normed standardized CITO (Institute for Educational Measurement) Test for Reading Comprehension (2010) of the Dutch National Institute for Educational Measurement was used to assess children's reading comprehension level. This test is part of the standard Dutch CITO pupil monitoring system and is designed to determine general reading comprehension level in elementary school children. This test consists of two modules, each involving a text and 25 multiple choice questions. The questions pertained to the word, sentence or text level, and tapped both the text base and situational representation that the reader constructed from the text (Kintsch, 1998). On this test, children's reading comprehension level is expressed by a reading proficiency score, which, in this study, ranged from 15 to 95 (M = 40.51, SD = 13.94). The internal consistency of this test was high with a Cronbach's alpha of .89 (Weekers, Groenen, Kleintjes & Feenstra, 2011).

Data analysis

A 2 x 2 x 2 analysis of variance (ANOVA) was conducted with Consistency (consistent vs. inconsistent) and Markedness (unmarked vs. marked) as within-subject factors and Group (less successful vs. successful word problem solvers) as the between-subject factor. Follow-up tests were performed using paired sample t-tests.

In the present study, the role of reading comprehension in the four word problem types was examined by calculating the correlations (Pearson's r) between reading comprehension and the difference score between the unmarked inconsistent and consistent word problem types, and the correlation between reading comprehension and the difference score between the marked inconsistent and consistent word problem types. These difference scores reflect the differences in performance between the consistent and inconsistent word problem types, and can be taken as a measure of the extent to which students are able to construct a mental representation of the described problem situation. The lower the difference score, the less word problem solvers suffer from the inconsistency. The correlations were calculated for less successful and successful word problem solvers separately.

This approach deviates from, but provides an important advantage over, the study by Van der Schoot et al. (2009), who added reading comprehension as a covariate in the repeated measures ANOVA. That is, the results obtained by Van der Schoot et al. (2009) could provide only limited insight into the exact locus of the covariate's effect, as it was not known which group (less successful or successful word problem solvers) or in which word problem type (consistent unmarked/marked or inconsistent unmarked/marked) reading comprehension played a role. Moreover, it turns out that the repeated measures ANCOVA does change the main effects of the repeated measures compared to assessing the main effects via a simple repeated measures ANOVA (see Thomas, Annaz, Ansari, Scerif, Jarrold, & Karmiloff-Smith, 2009). So, the approach used in the present study enabled us to obtain more specific insight into the precise role of reading comprehension in word problem solving. All the analyses had an alpha of .05.

RESULTS

In Figures 1 and 2, word problem solving performance is presented as a function of consistency (consistent vs. inconsistent) and markedness (marked vs. unmarked) for less successful problem solvers (Figure 1), and for successful problem solvers (Figure 2), respectively.

Inspection of both figures shows that, as expected, the effects of consistency and markedness differed for less successful and successful word problem solvers. As shown in Figure 1, for less successful word problem solvers there was a consistency effect for both marked and unmarked word problems (Consistency: F (1,40) = 10.94, $p < .01$, $\eta_p^2 = .22$; Consistency x Markedness interaction: F (1,40) = 0.25, $p = .62$, $\eta_p^2 = .01$, indicating a large and small effect size respectively, according to Pierce, Block & Auguinis [2004]). So, less successful word problem solvers performed significantly lower on both the unmarked and marked inconsistent word problem types, compared to the consistent unmarked and marked word problem types (t (40) = 2.22, $p < .05$; t (40) = 3.02, $p < .01$ respectively).

However, as displayed in Figure 2, different findings were obtained in the group of successful problem solvers. In this group, the consistency effect was present for marked but absent for unmarked word problems (Consistency: F (1, 38) = 13.00, $p < .001$, $\eta_p^2 = .26$; Consistency x Markedness interaction: F (1, 38) = 16.03, $p < .001$, $\eta_p^2 = .30$, which can be considered to be large effects according to Pierce et al., [2004]). This indicates that successful word problem solvers performed significantly lower on marked inconsistent compared to marked consistent word problems (t (38) = 4.67, $p < .001$); whereas performance on unmarked consistent and unmarked inconsistent word problem types did not differ significantly (t (38) = 1.07, $p = .29$). This pattern of findings regarding the successful and less successful problem solvers was evidenced by a significant three-way interaction between consistency, markedness, and group (F (1,78) = 4.32, $p < .05$, $\eta_p^2 = .05$, indicating a medium-sized effect).

In sum, these findings show that less successful word problem solvers performed lower on both semantic-linguistically simple and complex word problems, whereas successful word problem solvers only performed lower when the word problem text contained complex semantic-linguistic features.

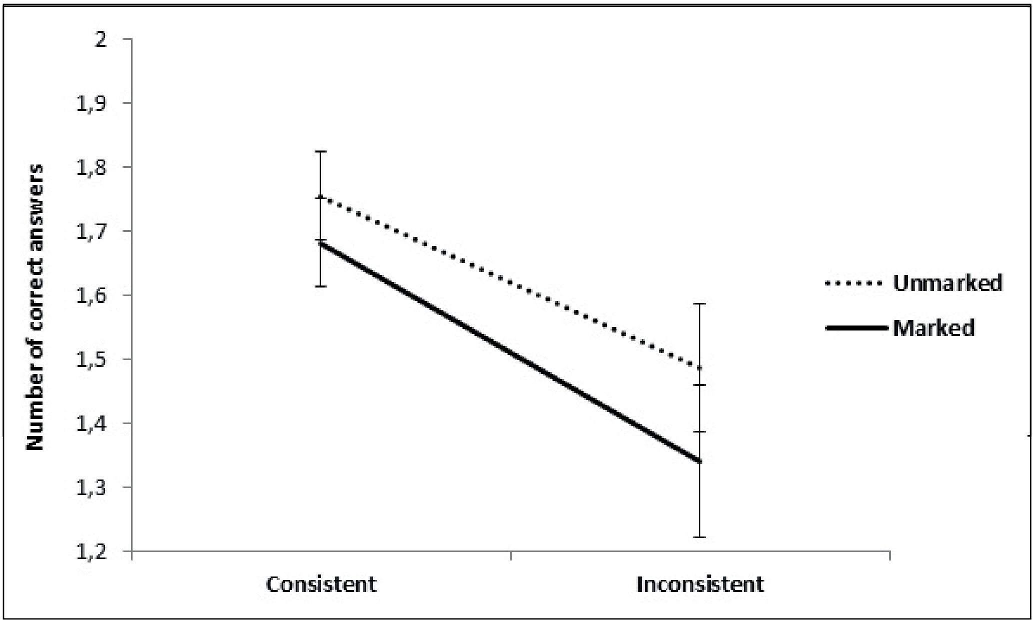


Figure 1. Less successful word problem solvers: Interaction effect Consistency x Markedness x Group

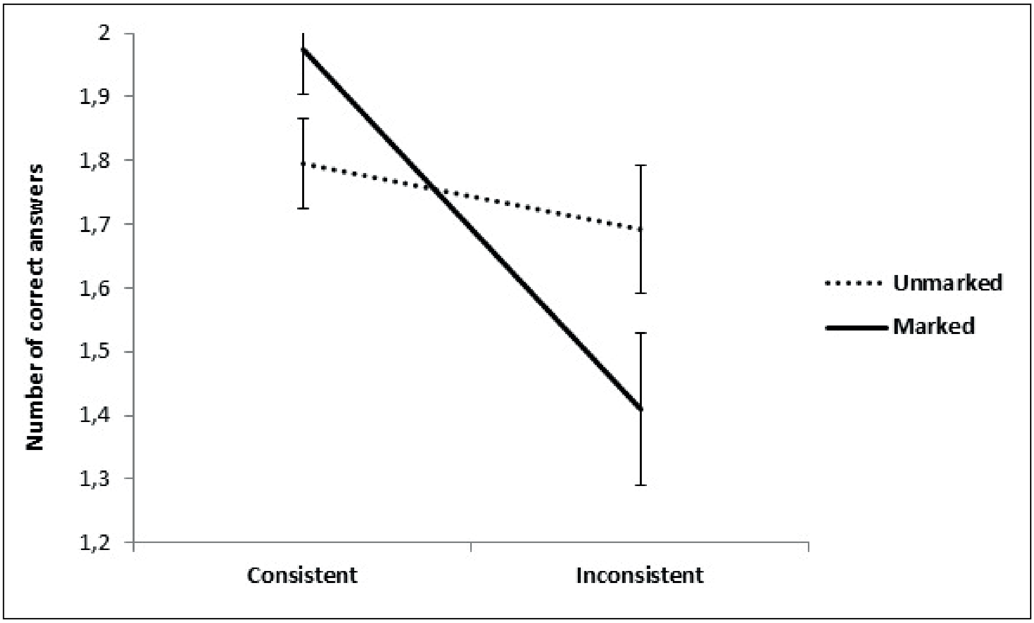


Figure 2. Successful word problem solvers: Interaction effect Consistency x Markedness x Group

Regarding the role of reading comprehension ability in word problem solving: overall successful word problem solvers (M = 46.42, SD = 2.66) scored significantly higher on the standardized reading comprehension test than less successful word problem solvers (M = 35.02, SD = 1.27), t (53.32) = 3.87, p < .001). To obtain more detailed insight into the role of reading comprehension skills in solving marked inconsistent word problems, reading comprehension ability was correlated with the difference scores (inconsistent - consistent) computed for the marked and unmarked word problem types.

In line with our expectations, the results of the correlational analyses show that only in the group of successful word problem solvers the difference score for the marked word problem type was significantly related to reading comprehension (Pearson's r = -.40, p < .05, r^2 = .16). That is, in the group of successful word problem solvers, a higher reading comprehension score was associated with a smaller difference score, which indicates that performance on marked word problems is higher for students who have higher reading comprehension abilities. This suggests that students with higher reading comprehension abilities appear to have a higher chance of overcoming problems with solving marked word problems.

Importantly, reading comprehension was not correlated with the successful word problem solvers' difference scores for unmarked word problems (r = -.27, p = .10). Furthermore, in the group of less successful word problem solvers, reading comprehension was also not correlated with the difference scores computed for either unmarked (r = -.04, p = .76) or marked word problems (r = -.04, p = .83).

DISCUSSION

This study set out to investigate the claim that the contemporary RME approach pays limited attention to the teaching of semantic-linguistic skills during word problem solving instruction. We therefore designed a study in which we not only manipulated the extent to which a sophisticated representation strategy was required, but also varied the semantic complexity of the word problems by using a marked (i.e., high semantic complexity) or unmarked (i.e., low semantic complexity) relational term in the word problem text. Moreover, we classified students as successful and less successful word problem solvers on the basis of their performance on an independent and well-established RME-specific mathematics test.

Using this classification procedure, it was hypothesized that less successful word problem solvers would experience difficulties with correctly solving inconsistent word problems irrespective of their semantic complexity (Hypothesis 1). This hypothesis was confirmed by our analyses, which showed that less successful word problem solvers performed poorly on both marked and unmarked inconsistent word problems. Successful word problem solvers, on the other hand, were able to effectively solve inconsistent word problems that had a low semantic complexity. This finding suggests that the sophisticated representation skills required to solve non-obvious word problems are adequately learned in the RME curriculum, at least by successful word problem solvers.

However, on semantically complex word problems even the successful problem solvers experienced difficulties, as indicated by the large number of errors they made on marked inconsistent word problems (Hypothesis 2). More concretely, successful word problem solvers found it more difficult to translate a marked relational term ('less than') into an addition operation, than to translate an unmarked relational term ('more than') into a subtraction operation.

These findings once again support prior observations that (subtle) semantic-linguistic elements of a word problem, more specifically the marked relational term, influence word problem solving success (Clark, 1969; Kintsch, 1998; Lewis & Mayer, 1987; Pape, 2003; Van der Schoot et al., 2009). Moreover, they are in line with empirical work reporting processing problems with marked terms, which is suggested to be caused by the semantic representation of negative poles like 'less than' being more fixed and complex, and therefore less likely to be reversed, than that of positive poles like 'more than' (e.g., Lewis & Mayer, 1987; for a detailed explanation of the underlying mechanism, see e.g. Clark, 1969). For example, earlier studies have shown that students are less able to recall marked terms accurately in memory tasks (Clark & Card, 1969), have slower naming responses for marked terms in naming tasks (Schriefers, 1990), have slower solution times for problems with marked adjectives in reasoning problems (French, 1979), and experience problems with reversing a marked inconsistent word problem (e.g., Pape, 2003; Van der Schoot et al., 2009).

Importantly, our results showed that even successful students appear to be insufficiently equipped with the semantic-linguistic skills required to solve semantically complex word problems correctly. Given the current classification procedure, it is possible that students were simply not taught the necessary amount of semantic-linguistic skills during word problem solving instruction. This reinforces our premise that the development of semantic-linguistic skills receives little attention in contemporary RME instruction, thereby identifying an important aspect of current teaching practice of word problem solving in RME that could be reconsidered.

Building upon prior studies (e.g., Lee et al., 2004; van der Schoot et al., 2009), another aim of this study was to investigate whether reading comprehension ability could help (successful) word problem solvers to overcome the semantically complex marked relational term in an inconsistent word problem. In line with our expectations, reading comprehension was positively related to the performance on marked (but not unmarked) inconsistent word problems for the group of successful word problem solvers; whereas for the less successful group no significant relations were found between reading comprehension and word problem solving (Hypothesis 3).

These results provide corroborating evidence that general reading comprehension skills play an important role in students' ability to correctly solve semantically complex word problems. Moreover, our findings represent an advance over prior work by more specifically delineating which types of word problems and for which students reading comprehension ability might have an effect. This study shows that reading comprehension skills are especially helpful when it comes to improving the performance on semantically complex word problems by successful word problem solvers (as classified by the RME mathematics test). This suggests that despite having acquired limited semantic-linguistic skills during word problem solving instruction in the RME curriculum, (successful) students have the ability to rely on their reading comprehension skills to effectively solve semantically complex word problems.

In conclusion, the present study showed that students who performed well on word problems offered in RME, and therefore were characterized as successful word problem solvers, did not necessarily correctly solve word problems on an independent word problem test that contained problems that are semantically complex, and hence require both representational skills and semantic-linguistic skills. These findings suggest that word problem solving instruction in the RME curriculum is insufficient in the sense that little emphasis is placed on the explicit teaching of semantic-linguistic skills. This conclusion is particularly relevant for the educational practice of RME. The main implication is that word problem solving instruction should place greater emphasis on teaching the semantic-linguistic skills that enable students to process the semantic complexities that appear in the word problem statement adequately.

It is important to start developing such skills early in elementary school, as word problems get semantically more complex as students progress in their educational career, for example when making the transition from elementary to secondary education (Helwig et al., 1999; Silver & Cai, 1996). Making teachers in RME aware of the possible imbalance between the amount of instruction time being devoted to the teaching of strategic representation skills and semantic-linguistic skills, and encouraging them to pay more attention to semantic-linguistic skills, would provide a good starting point. Moreover, it is useful to make a distinction between learning to process more subtle semantic-linguistic text features (like a marked relation term) and dealing with more general semantic text complexities (like the relevance of the information in the word problem text, the explicitness of the described relations, and the sequence of the known elements in the word problem text).

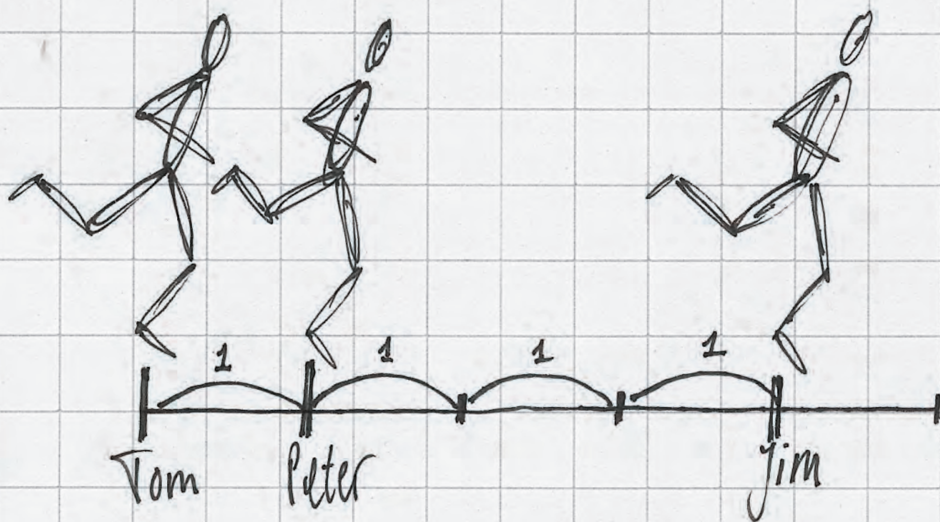
These and other practical aspects of the results, such as finding the optimal balance between the amount of instruction in strategic representational and semantic-linguistic skills, remain to be addressed in future research. Presumably, currently effective intervention programs that focus on both strategic representational and semantic-linguistic skills, such as schema-based instruction (e.g., Jitendra, Star, Rodriguez, Lindell, & Someki, 2011; Jitendra, DiPipi, & Perron-Jones, 2002), and the *Solve It!* instruction method (Krawec, Huang, Montague, Kressler, & Melia de Alba, 2013; Montague, Warger, & Morgan, 2000), could provide a fruitful starting point in pursuing this challenge.

4 The role of visual representation type, spatial ability, and reading comprehension in word problem solving:

An item-level analysis in elementary school children.

Anton J. H. Boonen, Floryt van Wesel, Jelle Jolles, & Menno van der Schoot

International Journal of Educational Research (2014), 68, 15-26



In an athletics race, Jim is four meters ahead of Tom and Peter is three meters behind Jim. How far is Peter ahead of Tom?

ABSTRACT

This study examined the role of visual representation type, spatial ability, and reading comprehension in word problem solving in 128 sixth-grade students by using primarily an item-level approach rather than a test-level approach. We revealed that compared to students who did not make a visual representation, those who produced an accurate visual-schematic representation increased the chance of solving a word problem correctly almost six times. Inaccurate visual-schematic and pictorial representations, on the other hand, decreased students' chance of problem solving success. Noteworthy, reading comprehension was related to word problem solving at the test-level but not at the item-level. In interpreting the results, we advocate the use of item-level analyses since they are able to disclose such level-of-analysis discrepancies.

INTRODUCTION

Mathematical word problem solving has received a lot of attention in the scientific literature (e.g., Boonen, Van der Schoot, Van Wesel, De Vries, & Jolles, 2013; Cummins, Kintsch, Reusser, & Weimer, 1988; Hegarty & Kozhevnikov, 1999; Pape, 2003). Theoretical models converge on the idea that word problem solving is mainly composed of two phases: (1) the problem representation phase, which involves the identification and representation of the problem situation which is "hidden" in the word problem text, and (2) the problem solution phase, which includes the planning and execution of the required mathematical computations (e.g., Hegarty, Mayer, & Monk, 1995; Krawec, 2010; Lewis & Mayer, 1987; for an overview of the most significant theories on word problem solving, see Kintsch, 1998; Kintsch & Greeno, 1985). These models have led to the conclusion that students often struggle with solving word problems even when they perform competently on the computations required to solve these problems (Cummins et al., 1988; Lewis & Mayer, 1987; Schumacher & Fuchs, 2012). One of the problem solving skills children have been found to have difficulties with is the ability to generate an adequate visual representation of a word problem (Hegarty & Kozhevnikov, 1999; Van Garderen, 2006). However, we believe that the test-level-based approach which was used in quite a lot of the previous studies has some drawbacks. Therefore, the present study took a new, item-level, approach to gain a more complete understanding of the role of (different types of) visual representations in word problem solving.

In previous studies, a positive relationship between visual representation type and word problem solving performance has been established by calculating correlations between the total amount of (specific) visual representations produced and the total amount of correctly solved word problems (e.g., Blatto-Vallee, Kelly, Gaustad, Porter, & Fonzi, 2007; Guoliang & Pangpang, 2003; Hegarty & Kozhevnikov, 1999; Krawec, 2010, 2012; Van Garderen, 2006; Van Garderen & Montague, 2003). However, calculating the correlation between two sum scores is an example of a test-level approach entailing limitations to be considered in this study. In particular, for three reasons, this correlation does not necessarily demonstrate that a certain type of visual representation has actually resulted in the correct answer to the word problem for which the visual representation was made. First, we have to consider that the established correlation between the total amount of (specific) visual representations and the total amount of correctly solved word problems might be explained by an underlying latent factor, such as a general measure of intelligence or cognitive ability (Boonen et al., 2013; Keith, Reynolds, Patel, & Ridley, 2008). Second, the relation might be explained by a mediating variable, for instance the capability of students to derive the correct mathematical operations from the visual representation and to determine the order in which these operations should be executed (i.e., the solution planning phase; see Mayer, 1985; Krawec, 2010). Finally, we have to acknowledge the possibility that a certain type of visual representation was made, but not used for answering the corresponding word problem. Ignoring this latter point may lead to an error of reasoning known as the *ecological fallacy,* where a researcher makes an inference about an individual based on aggregated data for a group (Lichtman, 1974; Robinson, 2009). For the purpose of the current study, it is important to recognize that a similar mistake is made when conclusions are drawn about performance on a specific item of a test (i.e., at the “individual” level) based on statistical analyses on the sum score of that test (i.e., at the “group” level). These considerations limit the conclusions we can draw with respect to the established relationship between visual representations and word problem solving performance.

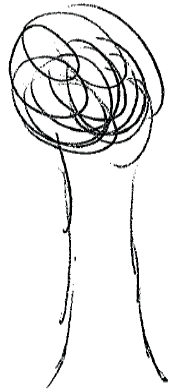
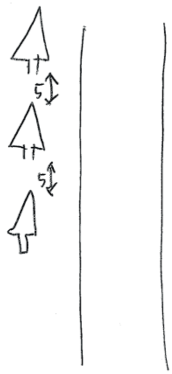
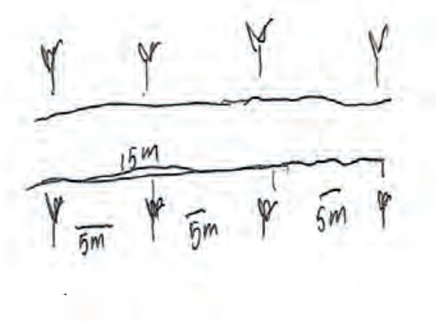
Hence, we can conclude that the way in which the relation between (type of) visual representation and word problem solving performance was investigated in past research does not provide a decisive answer to the question if, and to what extent, visual representation type affects the chance of producing a correct solution to the word problem for which the representation was made. Taking the abovementioned considerations into account, we opted to investigate the importance of different types of visual representation for word problem solving success of students at the item-level rather than at the test-level. To achieve this, a change in statistical modeling was necessary. Previous studies used the sum scores at the test-level (i.e., the total amount of correctly answered word problems and the total amount of visual representations produced), which are continuous in nature and suitable for a linear regression model. However, in our study we used the scores on the item-level. These scores are categorical in nature and suitable for a logistic regression model, viz., for each item the answer is either correct or incorrect (dichotomous). Likewise, each visual representation which is evoked can be classified in one class of a set of categories (see below). Hence, our change in approach from the test-level to the item-level also meant that we opted for a logistic regression model instead of a linear regression model.

The item-level approach thus gave us the opportunity to examine if, and to what extent, the production of a visual representation affected the chance of successfully solving the word problem for which the visual representation was made (henceforth referred to as the *chance of problem solving success*). However, the chance of solving a word problem successfully is thought to be largely dependent on the *type* of visual representation that is produced (Hegarty & Kozhevnikov, 1999). In the present study three different types of visual representation were distinguished: pictorial representations, inaccurate visual-schematic representations, and accurate visual-schematic representations.

Generally, the production of pictorial representations involves the construction of vivid and detailed images (Hegarty & Kozhevnikov, 1999; Van Garderen, 2006). We expected to find that pictorial representations negatively affect the chance of problem solving success, as these representations merely concern images that encode the visual appearance of objects and persons described, and thus are irrelevant for the actual solution process (Hegarty & Kozhevnikov, 1999; Mayer, 1998; Van Garderen, 2006; Van Garderen & Montague, 2003). In line with the "seductive details"-effect (Sanchez & Wiley, 2006), we hypothesized that forming a pictorial representation would divert the problem solvers' attention away from constructing a coherent (mental) model of the word problem, including the appropriate relations between the key variables. Visual-schematic representations, on other hand, do contain a coherent image of the problem situation hidden in the word problem, including the relations between the solution-relevant elements (Edens & Potter, 2008; Hegarty & Kozhevnikov, 1999; Kozhevnikov, Hegarty & Mayer, 2002; Van Garderen & Montague, 2003). In contrast to the previously mentioned literature, however, in our study we made a distinction between two different types of visual-schematic representations. We hypothesized that only *accurate* visual-schematic representations would increase the chance of problem solving success, as in this visual representation type problem solvers infer the correct relations between the solution-relevant elements from the text base of the word problem and integrate them into a coherent visualization of the problem situation (Krawec, 2010). In *inaccurate* visual-schematic representations, these relations are also included but, in contrast to accurate visual-schematic representations, they are either incorrectly drawn or partly missing. It follows that this may put problem solvers on the wrong track in solving the problem (Krawec, 2010). Therefore, we expected to find that inaccurate visual-schematic representations would actually decrease the chance of problem solving success. An example of each visual representation type is given in Table 1.

Table 1. *Examples of the different types of visual representations*

[Word problem example]
At each of the two ends of a straight path, a man planted a tree and then every 5 meters along the path he planted another tree. The length of the path is 15 meters. How many trees were planted?

Pictorial representation	**Inaccurate visual-schematic representation**	**Accurate visual-schematic representation**
		

At this point, it is important to recognize that both the accurate and inaccurate visual-schematic representation category represent a mixture of internal visualization (mental imagery) and external visualization (gestures, drawing) efforts. Although these different approaches seem to share a common processing mechanism (Leutner, Leopold, & Sumfleth, 2009; Schnotz & Kürschner, 2008), there may also be differences. For example, mental imagery requires participants to keep information active in working memory, while creating an external representation may "offload" cognitive processing (Leutner et al., 2009; Schnotz & Kürschner, 2008). Also, drawing pictures on paper while comprehending the (structure of the) word problem might help less successful problem solvers in the process of building effective problem representations (Bryant & Tversky, 1999; Leutner et al., 2009). In particular, drawing construction is thought to facilitate the metacognitive monitoring processes involved in comprehending (word problem) text (e.g., Van Meter, 2001; Van Meter & Garner, 2005). In this study, we therefore examined whether the supposed differences between internal and external visual representations affected the chance of problem solving success.

In addition to the type (accurate visual-schematic vs. inaccurate visual-schematic vs. pictorial) and locus (internal vs. external) of visual representations, we also looked at basic cognitive abilities underlying word problem solving success. Previous research, using a test-level approach, has shown that particularly two basic cognitive abilities are related to students' word problem solving performance: 1) spatial ability, covering the visual-spatial processing domain (Boonen et al., 2013; Hegarty & Kozhevnikov, 1999; Van Garderen, 2006); and 2) reading comprehension, covering the semantic-linguistic processing domain (Bernardo, 1999; Van der Schoot et al., 2009; Vilenius-Tuohimaa, Aunola, & Nurmi, 2008). Based on the increasingly accepted idea that word problem solving taps into both domains (Boonen et al., 2013; Krawec, 2010, 2012), we aimed to replicate the findings of prior studies regarding the importance of spatial ability and reading comprehension by using analyses at the item-level rather than at the test-level. In doing so, we were able to overcome the aforementioned limitations regarding test-level analyses and prevent ourselves from falling for the ecological fallacy. As a consequence, we could contribute to a more complete understanding of the importance of both these abilities.

METHODS

Participants

One hundred twenty-eight Dutch sixth-grade students (64 boys, M_{age} = 11.73 years, SD_{age} = 0.43 years and 64 girls, M_{age} = 11.72 years, SD_{age} = 0.39 years) from eight elementary schools across the Netherlands took part in this study. The students were selected on the basis of their performance on the CITO Mathematics test (2008) so as to obtain a representative and balanced sample of low, average, and high math performers. The CITO Mathematics test is a nationwide standardized test (developed by the Institute for Educational Measurement) used to monitor students' general math ability during their elementary school career. Parents provided written informed consent based on printed information about the purpose of the study.

Instruments and measurement procedure

Below, we describe the measures we used to assess word problem solving performance, the production of different visual representation types, spatial ability, and reading comprehension.

Word problem solving performance

Word problem solving performance was examined with the Mathematical Processing Instrument (MPI; Hegarty & Kozhevnikov, 1999; Van Garderen & Montague, 2003, see Appendix A) which was translated into Dutch (Boonen et al., 2013). The internal consistency (Cronbach's alpha) of this instrument, which consists of 14 non-routine word problems, was reported as high (.78) in a sample of American participants (Hegarty & Kozhevnikov, 1999). The Cronbach's alpha in this study was .72. The word problems were printed on cards and presented in four different orders. All problems were read out loud to the students to control for differences in decoding skill. To exclude the possibility that the execution of the required arithmetic operations would be a determining factor in students' word problem solving, the operations were easy and could be solved by every student. Furthermore, students were required to solve each word problem within three minutes and during this time the experimenter did not speak to the student. To be sure that students had enough time to solve the word problems, a pilot study was conducted with five sixth-grade students. The results of the pilot study showed that every student was able to solve each of the 14 items of the MPI within the three minutes provided (see Boonen et al., 2013).

Representation type

After solving each word problem on the MPI, a short interview was held about the nature of the visual representation which was evoked. The exact procedure of this interview was adapted from the study of Hegarty and Kozhevnikov (1999; see Appendix B for the interview-format). For each item, a score was derived reflecting (1) the type of the visual representation which was made, and (2) whether the representation was external (i.e., a gesture or drawing with paper and pencil) or internal (i.e., mental). The category "no visual representation" was assigned when the student did not use any visual representation to solve the word problem. When the student made a visual representation, three categories were distinguished: accurate visual-schematic representations, inaccurate visual-schematic representations, or pictorial representations. These visual representation types are clarified for the following word problem:

> *"A balloon first rose 200 meters from the ground, then moved 100 meters to the east, then dropped 100 meters. It then traveled 50 meters to the east, and finally dropped straight to the ground. How far was the balloon from its original starting point?"*

A visual representation was coded as accurate visual-schematic if students drew an image or diagram, used gestures, or reported a mental image, thereby specifying/including the correct relations between the solution-relevant elements in a problem. Figure 1 shows an example of an accurate visual-schematic representation.
A visual representation was coded as inaccurate visual-schematic if relations were included in the representation, but one or more

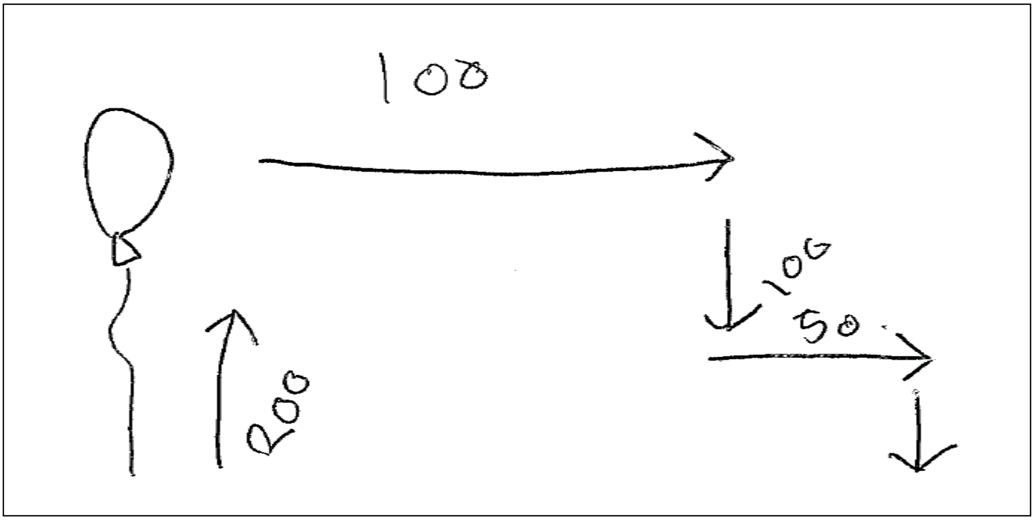


Figure 1. An example of an accurate visual-schematic representation

relations were missing or incorrectly specified (for clarity's sake: the terms "accurate" and "inaccurate" thus do not refer to whether the answer is correct or incorrect but rather to whether or not the child inferred the correct relations between the solution-relevant elements from the word problem text). Figure 2 shows an example of an inaccurate visual-schematic representation.

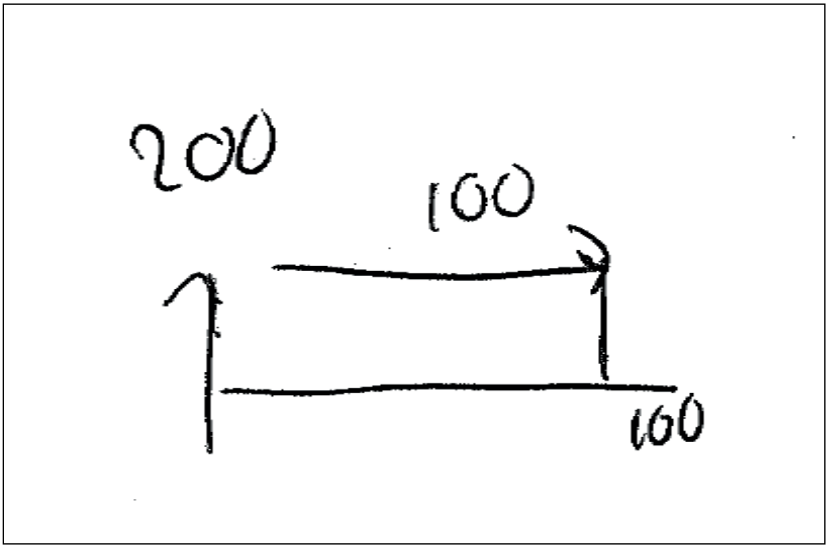


Figure 2. An example of an inaccurate visual-schematic representation

A visual representation was coded as pictorial if a student reported or drew an image of the objects and/or person(s) referred to in the problem. Thereby, the relevant criterion was that he or she focused solely on the external appearance of the objects and/or persons, without paying attention to the structure of the problem situation described in the text. Figure 3 shows an example of a pictorial representation.

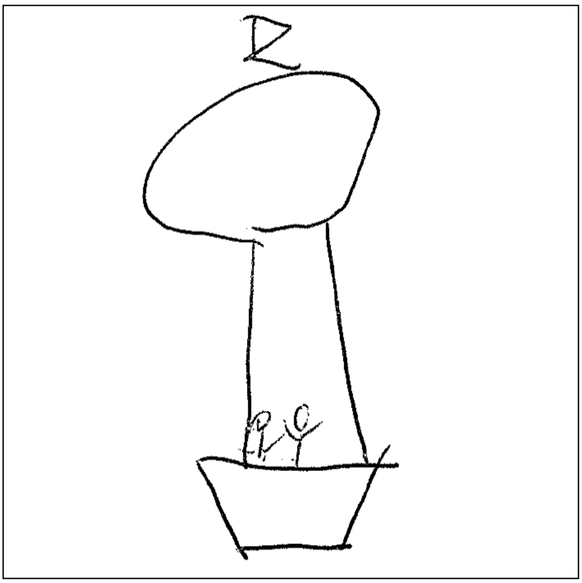


Figure 3. An example of a pictorial representation

Students made a total of 625 visual representations. All visual representations were coded by three independent coders. In the first coding session 32 representations were randomly selected and coded into the three categories by all coders. The inter-rater reliability of these 32 coded representations was high (Cohen's Kappa (κ) = .88; Tabachnick & Fidell, 2006). Because the results of this coding session were satisfactory, the remaining visual representations were coded by all coders in the same way.

4

Spatial ability

The Paper Folding task (retrieved from The Kit of Factor-Referenced Cognitive Tests; Ekstrom, French, Harman, & Derman, 1976) and the Picture Rotation task (based on Quaiser-Pohl, 2003) are standardized tasks used to measure spatial visualization (see Boonen et al., 2013).

In the Paper Folding task, students were asked to imagine the folding and unfolding of pieces of paper. Each problem of the test consisted of several figures drawn left and right of a vertical line. All figures represent a squared piece of paper on which one or two small circles were drawn to show where the paper had been punched. The figures on the left side represented folded pieces of paper. In the last figure on this side holes were punched throughout the thicknesses of the paper. On the right side of the vertical line five figures showing where the holes will be located when the paper was completely unfolded were presented. Students had to decide which one of these figures was correct with reference to the folded piece of paper on the left side. Figure 4 shows one of the 20 test items of the Paper Folding task. This task took 6 minutes and had a sufficient internal consistency coefficient in the present study (Cronbach's alpha = .70).

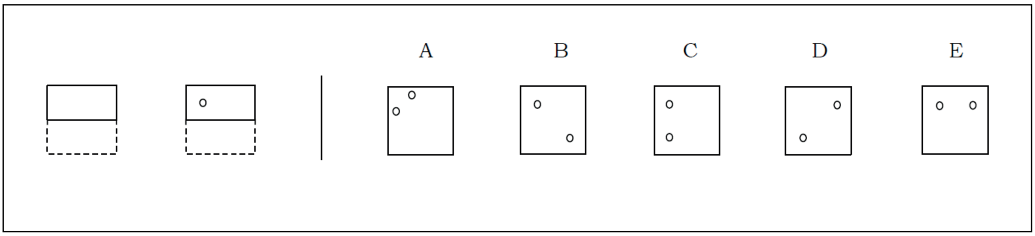


Figure 4. The Paper Folding task (Ekstrom, French, Harman, & Derman, 1976)

In the Picture Rotation task students were asked to rotate a non-manipulated picture of an animal presented at the left side of a vertical line. The three pictures at the right side of the vertical line showed rotated and/or mirrored images of that same animal. Two of these pictures were both rotated and mirrored and one was only rotated. Students had to decide which of the three pictures was only rotated. Figure 5 shows one of the 30 test items of the Picture Rotation task. This task took 1.5 minutes and its internal consistency coefficient in this study was high (Cronbach's alpha = .93).

Figure 5. The Picture Rotation task (based on Quaiser-Pohl, 2003)

To obtain a general measure of spatial ability, the raw scores on each of the spatial ability tasks were rescaled into a z-score. Subsequently, these z-scores were aggregated into an average z-score (M = .00, SD = .84).

Reading comprehension

The (Grade 6 version of the) nationally normed standardized Reading Comprehension test of the Dutch National Institute for Educational Measurement (CITO, 2010) was used to assess children's reading comprehension level. This test is part of the standard Dutch CITO pupil monitoring system and is designed to determine general reading comprehension level in elementary school children. This test consists of two modules, each involving a text and 25 multiple choice questions. The questions pertained to the word, sentence, or text level and tapped both the text base and situational representation that the reader constructed from the text (Kintsch, 1998). On this test, children's reading comprehension level is expressed by a proficiency score. These proficiency scores (M = 42.06, SD = 14.06, range = 15.00 to 95.00) made it possible to compare the results of the reading comprehension test with other versions of this test from other years. The internal consistency of this test was high with a Cronbach's alpha of .89 (Weekers, Groenen, Kleintjes & Feenstra, 2011).

Analyses

To replicate the findings of previous studies (e.g., Hegarty and Kozhevnikov, 1999; Van Garderen, 2006; Van Garderen & Montague, 2003) at the test-level, descriptive statistics of, and correlations between, the key measures of the study were calculated. However, the main aim of the present study was to examine the extent to which the type and locus of representation, as well as spatial ability and reading comprehension skill, affected the chance of problem solving success at the item-level. As noted previously, this required using a different model for statistical analyses. First, a chi-square test was performed to examine the association between representation type (no visual representation vs. pictorial representation vs. accurate visual-schematic representation vs. inaccurate visual-schematic representation) and problem solving success (correct vs. incorrect answer). Subsequently, we performed a logistic regression analysis to examine (i.e., quantify) the extent to which the type and locus (internal vs. external) of representation, spatial ability and reading comprehension affected the chance of problem solving success. In this analysis, the category "no visual representation" served as the reference category.

RESULTS

Descriptive statistics and correlations

Table 2 shows the means, standard deviations and correlations of the different measures used in this study at the test-level: (1) the number of word problems for which a correct answer was given (word problem solving performance); (2) the number of word problems for which no visual representations were made; (3) the number of word problems for which accurate visual-schematic representations were made; (4) the number of word problems for which inaccurate visual-schematic representations were made; (5) the number of word problems for which pictorial representations were made; (6) the averaged spatial ability scores; and (7) the reading comprehension proficiency scores. The correlations showed that the production of accurate visual-schematic representations showed a medium to large positive correlation with word problem solving performance ($r = .44$, $p < .001$, Cohen, 1992). Furthermore, there was a medium negative correlation between the production of pictorial representations and word problem solving performance ($r = -.27$, $p < .001$). The production of inaccurate visual-schematic representation was, however, not significantly correlated with word problem solving performance ($r = -.04$, $p = .67$). Finally, spatial ability and reading comprehension both showed a medium to large positive correlation with word problem solving performance (respectively: $r = .59$, $p < .001$; $r = .45$, $p < .001$) and the production of accurate visual-schematic representations (respectively: $r = .31$, $p < .001$; $r = .23$, $p < .05$).

Association between representation type and chance of problem solving success

Students did not make a visual representation in 1167 out of the total amount of 1792 word problem items. From the remaining 625 visual representations that were produced, 279 representations could be allocated to the accurate visual-schematic representation type, 257 to the inaccurate visual-schematic representation type, and 89 to the pictorial representation type.

The results of the chi-square test showed that there was significant statistical dependence (i.e., association) between visual representation type and problem solving success: X^2 (3) = 220.10, $p < .001$ (see Table 3). The standardized residuals revealed that students who made accurate visual-schematic representations produced a correct answer more frequently (st.res = 8.5) than students who did not make visual representations (st. res. = -0.6), students who made inaccurate representations (st. res. = -5.7), and students who made pictorial representations (st. res. = -3.3). In addition, not making a visual representation resulted in as many correct as incorrect answers (st. res. correct = -.06; st. res. incorrect = .5). Finally, students who made inaccurate visual-schematic representations (st. res. = 5.4) and, to a lesser extent, those who made pictorial representations (st. res. = 3.2) answered the word problem incorrectly more frequently

Table 2. *Descriptive statistics and correlations between word problem solving performance, types of (visual) representations, spatial ability, and reading comprehension*

	M	*SD*	1.	2.	3.	4.	5.	6.	7.
1. Word problem solving performance	6.68	2.87	-						
2. No visual representation	9.13	3.68	-.20*	-					
3. Accurate visual-schematic representation	2.13	2.45	.44**	-.79**	-				
4. Inaccurate visual-schematic representation	2.06	1.77	-.04	-.73**	.25**	-			
5. Pictorial representation	0.69	1.06	-.27**	-.43**	.00	.30**	-		
6. Spatial ability (z-score)	.00	.85	.59**	-.22**	.31**	.11	-.16	-	
7. Reading comprehension	42.50	14.80	.45**	-.11	.23**	.01	-.17	.35**	-

** p < .05, ** p < .001*

Table 3. *Chi-square test: Correct/Incorrect solution x Type of representation*

			Solution		Total
			Incorrect	Correct	
Representation	No visual representation	Count	623	544	1167
		Std Residual	.5	-.6	
	Accurate visual-schematic representation	Count	48	231	279
		Std. Residual	-8.1	8.5	
	Inaccurate visual-schematic representation	Count	197	60	257
		Std. Residual	5.4	-5.7	
	Pictorial representation	Count	68	21	89
		Std. Residual	3.2	-3.3	
Total		Count	936	856	1792

The role of representation type, spatial ability, and reading comprehension

A logistic regression analysis was performed to examine whether the type of representation, the locus of representation, spatial ability, and reading comprehension differentially affected the chance of problem solving success. Table 4 shows the results of this analysis. In this table, the unstandardized (*B*) and standardized (Exp(B)) regression coefficients of *Model 1* reflected the extent to which representation type affected the chance of solving a word problem successfully (with “no visual representation” as the reference category, R^2 = .18). So, compared to the situation in which no visual representation was made, the production of an accurate visual-schematic representation increased the chance that a word problem was solved correctly 5.85 times (accurate visual-schematic representation vs. no visual representation, B = 1.77, Exp(B) = 5.85, SE = .17, p < .001). In contrast, the production of inaccurate visual-schematic representations decreased the chance that a word problem was solved correctly 2.94 times (i.e., 1/0.34), and pictorial representations decreased the chance that a word problem was solved correctly 2.78 times (i.e., 1/0.36; inaccurate visual representation vs. no visual representation: B = -1.09, Exp(B) = 0.34, SE = .17, p < .001; pictorial representation vs. no visual representation: B = -.1.02, Exp(B) = 0.36, SE = .26, p < .001). *Model 2 s*hows the results of the logistic regression analysis taking into account whether the visual representation was made externally (with paper and pencil) or internally (mentally) (R^2 = .19). The results showed that the locus of the visual representation did not affect the chance of solving a word problem correctly (internal vs. external: B = 0.43, Exp(B) = 1.53, *S.E.* = .27, p = .11). Subsequently, *Model 3* shows the results of the analysis after adding spatial ability and reading comprehension (R^2 = .25) to the set of predictors. In line with our expectations, an increase of 1 unit in spatial skills increased the chance of solving a word problem correctly 1.66 times (B = 0.51, Exp(B) = 1.66, SE = .08, p < .001). On the other hand, a 1 unit increase in reading comprehension skills increased the chance of solving a word problem correctly 1.02 times (B = 0.02, Exp(B) = 1.02, SE = .00, p < .05). Follow-up logistic regression analyses, performed for each representation type separately, revealed that the importance of spatial ability in problem solving success varied between the representation types. Spatial ability appeared to be the most powerful predictor of problem solving success when an accurate visual-schematic representation was produced (Exp(B) = 2.22), whereas the role of spatial ability in pictorial representation was the lowest (Exp(B) = 1.13). In contrast to spatial ability, the relevancy of reading comprehension in problem solving success did not differ between the different representation types (Exp(B) = 1.02).

Table 4. *Logistic regression analysis performed on variables associated with word problem solving*

Predictors	Model 1 (R^2 = .18)				Model 2 (R^2 = .19)				Model 3 (R^2 = .25)			
	B	Exp(*B*)	*SE*	*Wald*	*B*	Exp(*B*)	*SE*	*Wald*	*B*	Exp(*B*)	*SE*	*Wald*
Constant	-0.17	0.84	.06	8.22*	-1.17	0.31	.26	20.61**	-1.83	0.16	.33	30.07**
Accurate visual-schematic vs. no visual representation	1.76	5.85	.17	104.84**	2.45	11.56	.26	87.97**	2.28	9.80	.27	69.69**
Inaccurate visual-schematic vs. no visual representation	-1.09	0.34	.17	42.81**	-0.52	0.59	.22	5.57*	-0.73	0.48	.24	9.64*
Pictorial representation vs. no visual representation	-1.02	0.36	.26	15.11**	-0.58	0.56	.29	4.04*	-0.50	0.61	.30	2.83
Internal visual representation vs external visual representation					0.43	1.53	.27	2.54	0.23	1.26	.28	0.68
Spatial ability									0.51	1.66	.08	46.43**
Reading comprehension									0.02	1.02	.00	4.31*

* *p* < .05, ** *p* <.001

DISCUSSION

In this study, we examined the importance of visual representation type, spatial ability, and reading comprehension in word problem solving from an item-level approach rather than from the test-level approach often used in previous studies. This implied a move from a continuous outcome variable (MPI sum score) to a dichotomous outcome variable (items being correct or incorrect), and from correlations and linear regression models to chi-square tests and logistic regression models. This change in statistical modeling provided a more thorough and sophisticated understanding of representational, spatial and reading comprehension skills in word problem solving.

First, we performed a chi square test to examine the association between different types of representations and the chance of successfully solving the word problem for which the representation was made. The results of this test reinforced our decision to distinguish three instead of two types of visual representations. To be more specific, we demonstrated that only the production of accurate visual-schematic representations was more frequently associated with a correct than with an incorrect answer to a word problem. In contrast, not making a visual representation resulted in as many correct as incorrect answers to a word problem. Finally, inaccurate visual-schematic and pictorial representations were even found to be more frequently associated with an incorrect answer to a word problem.

Subsequently, we performed logistic regression analyses to quantify the extent to which the different representation types affected the chance of problem solving success. The results of these analyses showed that when students made an accurate visual-schematic representation this increased the chance that they solved the word problem correctly almost six times. Probably, this is due to the fact that this type of visual representation contains a complete and coherent image of the problem situation, including the correct relations between the key variables. In contrast, the production of inaccurate visual-schematic representations and pictorial representations *decreased* the chance of problem solving success, respectively 2.94 and 2.78 times. As pictorial representations merely concern images of the visual appearance of objects or persons described in the word problem (for example the image of the tree depicted in Table 1), they probably took the problem solvers' attention away from constructing a coherent model of the word problem, including the appropriate relations between the solution-relevant elements contained in it (Hegarty & Kozhevnikov, 1999; Van Garderen, 2006; Van Garderen & Montague, 2003). Although inaccurate visual-schematic representations do include these relations, they are either incorrectly drawn or missing. As a consequence, this type of representation may have put problem solvers on the wrong track when solving a word problem. In sum, we have offered a possible explanation of why pictorial and inaccurate visual-schematic representations were counterproductive in word problem solving. In addition to the type of visual representation, we looked at the locus of the visual representations which were made, and found no differences between internal and external visual representations with regard to the chance of problem solving success. Presumably, creating a representation of a word problem externally (with paper and pencil) and internally (mentally) both rely on the same basic processing mechanism underlying word problem solving (Leutner et al., 2009). This suggests that it is the content of the visual representation which matters, not the medium or locus of the representation.

Furthermore, besides contributing to a better understanding of the effects of the type and locus of representations on the chance of problem solving success, we tried to reproduce, at the item-level, the findings of previous studies using a test-level approach concerning the importance of spatial ability and reading comprehension in word problem solving (Bernardo, 1999; Boonen et al., 2013; Hegarty & Kozhevnikov, 1999; Van der Schoot et al., 2009; Van Garderen, 2006). In line with these earlier findings, the current study showed that spatial ability is a significant and relevant basic ability which increases the chance of solving a word problem successfully (Blatto-Vallee et al., 2007; Boonen et al., 2013; Hegarty & Kozhevnikov, 1999). In particular, our findings clarified the importance of spatial ability for the accurate visual-schematic representation type. Probably, this is due to the fact that the key elements and relations are encoded in these representations in a spatial manner (Hegarty & Kozhevnikov, 1999; Krawec, 2010; Van Garderen, 2006). However, our findings showed that the extent to which reading comprehension skills increase the chance of problem solving success is very limited. The results of the logistic regression analyses showed that although reading comprehension was a significant predictor in the model (due to the large number of items involved), the relevancy of its contribution was negligible (i.e., reading comprehension increased the chance of problem solving success only 1.02 times). Our item-level finding that reading comprehension was not a relevant factor contradicts the test-level findings from this study (r = .45) as well as previous studies demonstrating that reading comprehension and word problem solving performance are related. In other words, a relation between reading comprehension and word problem solving found at the test-level does not imply that reading comprehension positively affects the chance of problem solving success at the item-level.

The question remains how this shift in analysis from the test-level to the item-level leads to such different results. First, the finding that reading comprehension is important for word problem solving at the test-level may be explained by assuming that both abilities have a common underlying latent factor. Keith et al. (2008), for example, showed that both reading comprehension skill and word problem solving skill load high on a general (latent) measure of intelligence. Second, the reason that reading comprehension skills did not contribute to problem solving success at the item-level may have to do with the characteristics of the word problem items themselves. Previous studies showed that reading comprehension skills are particularly important in dealing with semantically complex word problem characteristics like the sequence of the known elements in the text of the word problem or the degree to which the semantic relations between the given and unknown quantities of the problem are made explicit (De Corte, Verschaffel, & De Win, 1985; De Corte, Verschaffel, & Pauwels, 1990; Marzocchi, Lucangeli, De Meo, Fini, & Cornoldi, 2002; Verschaffel, De Corte, & Pauwels, 1992). The items which comprised the current word problem solving test (i.e., the MPI) presumably did not involve these kind of semantic complexities, as a consequence of which children did not have to deploy their reading comprehension skills or only to a limited extent. So, one of the strengths of this study is that it demonstrated that some basic skill (i.e., reading comprehension) can be found to be related to performance on an academic test (i.e., word problem solving test) at the test-level but not at the item-level. We believe that, here but also in general, it is important to be aware of this possibility and to identify more fully the circumstances under which such level-of-analysis discrepancies may occur, so that we are better equipped to avoid the ecological fallacy.

In considering the weight that should be given to the above conclusions, we would like to discuss three constraints of the study. First, our focus was on self-generated representations rather than on receiving representations (i.e., pictorial/graphical support) provided externally. That is, the primary aim of the study was to investigate the extent to which children are able to increase their chance of problem solving success by "self-producing" visual representations. This is in contrast to much of the previous literature, which largely deals with the impact of different types of illustrations, provided by the experimenter or another external source, on word problem solving (e.g., Berends & Van Lieshout, 2009; Dewolf, Van Dooren, Cimen & Verschaffel, 2014). Second, we studied word problem solving using non-routine rather than routine problems. By definition, routine word problems contain a generic problem pattern or semantic structure which characterize many addition and subtraction problems. For example, word problems presented in the basic methods of mathematics education often involve a "change", "grouping" or "comparison" story situation (Cummins et al., 1988; Jitendra et al., 2009; Jitendra, George, Sood, & Price, 2010). Therefore, solving a routine word problem above all requires identifying the base type of problem situation which is "hidden" in the word problem text (Jitendra, 2002; Jitendra et al., 2009, 2010). In contrast, non-routine problems do not map onto relevant existing schemas. As a consequence, they have no standard procedure of solving, but instead require more heuristic-based strategies. In addition, and more relevant to this study, non-routine problems are more challenging because generic and familiar problem structures are easier to visually represent than "one-of-a-kind" problem structures (Elia, Van den Heuvel-Panhuizen, & Kovolou, 2009). This does not mean, however, that our findings are not generalizable to routine word problems. Our conclusions regarding the importance of accurate visual-schematic representations are also applicable to routine word problem solving. The only difference in favor of routine word problem solving, though, is that once a problem solver has learned how to create an accurate visual-schematic representation for, for example, a "comparison" problem type, this knowledge is expected to easily transfer to other "comparison" problem types. Third, we relied on verbal protocols to reveal whether, and, if so, what type of representations children created during word problem solving. As we have known for a long time, verbal protocols may constitute unreliable data (e.g., Nisbett & Wilson, 1977). Therefore, in future research, more sophisticated measures should be used to examine the role of visual representations in word problem solving in addition to retrospective verbal reporting. For example, eye-tracking methods represent a powerful "online" method of assessing representational processes during word problem solving (Van der Schoot et al., 2009).

The findings of this study are not only theoretically relevant but they also have valuable implications for the practice of elementary school mathematical education. First of all, it is striking that, in general, students made use of a visual representation to represent and solve a word problem on only 35% of the occasions. Apparently, students found it difficult, or were not effectively taught, to create visual representations during mathematical word problem solving. This is worrisome given the main outcome of this study showing that, compared to a situation in which no visual representation was made, students who made visual-schematic representations increased the chance of solving a word problem correctly almost six times. At least, this was found to be true for accurate visual-schematic representations. Inaccurate visual-schematic representations, on the other hand, decreased the chance of problem solving success. This substantiates the conclusion that inferring the appropriate relations between the key variables from the text base of the word problem (i.e., relational processing;, see Boonen et al., 2013) is a crucial aspect in the production of visual-schematic representations. Although previous instructional approaches on word problem solving accentuated the importance of visualizing the word problem (Jitendra et al., 2009; Montague, Enders, & Dietz, 2011; Montague, Warger, & Morgan, 2000), we conclude that a simple "make a picture"-strategy that is proposed in several instructional programs may be formulated too broadly, and does not give a clear indication of the specific requirements to which an effective visual representation should comply (Montague, 2003; Montague et al., 2000).

APPENDIX 4.A. WORD PROBLEMS OF THE MATHEMATICAL PROCESSING INSTRUMENT

The word problems on the Mathematical Processing Instrument (Hegarty & Kozhevnikov, 1999):

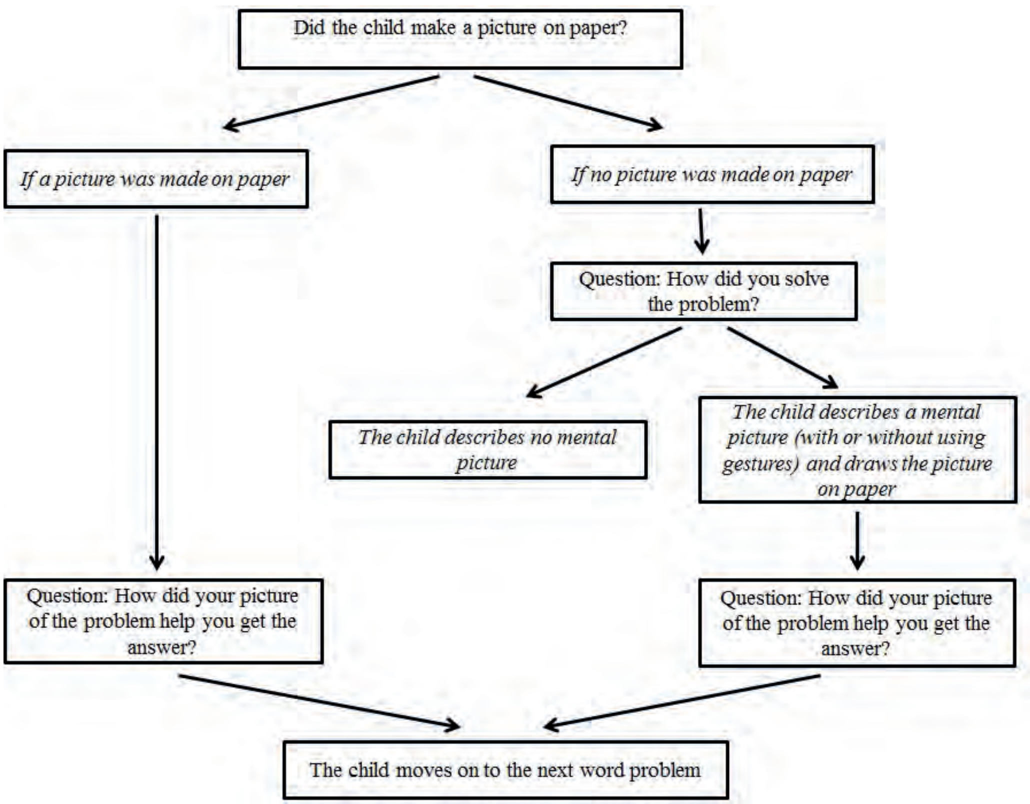
1. At each of the two ends of a straight path, a man planted a tree and then every 5 meters along the path he planted another tree. The length of the path is 15 meters. How many trees were planted?
2. On one side of a scale there is a 1 kg weight and half a brick. On the other side there is one full brick. The scale is balanced. What is the weight of the brick?
3. A balloon first rose 200 meters from the ground, then moved 100 meters to the east, then dropped 100 meters. It then traveled 50 meters to the east, and finally dropped straight to the ground. How far was the balloon from its original starting point?
4. In an athletics race, Jim is four meters ahead of Tom and Peter is three meters behind Jim. How far is Peter ahead of Tom?
5. A square (A) has an area of 1 square meter. Another square (B) has sides twice as long. What is the area of B?
6. From a long stick of wood, a man cut 6 short sticks, each 2 feet long. He then found he had a piece of 1 foot long left over. Find the length of the original stick.
7. The area of a rectangular field is 60 square meters. If its length is 10 meters, how far would you have traveled if you walked the whole way around the field?
8. Jack, Paul and Brian all have birthdays on the 1st of January, but Jack is one year older than Paul and Jack is three years younger than Brian. If Brian is 10 years old, how old is Paul?
9. The diameter of a tin of peaches is 10 cm. How many tins will fit in a box 30 cm by 40 cm (one layer only)?
10. Four young trees were set out in a row 10 meters apart. A well was situated beside the last tree. A bucket of water is needed to water two trees. How far would a gardener have to walk altogether if he had to water the four trees using only one bucket?
11. A hitchhiker set out on a journey of 60 miles. He walked the first 5 miles and then got a lift from a lorry driver. When the driver dropped him he still had half of his journey to travel. How far had he traveled in the lorry?
12. How many picture frames 6 cm long and 4 cm wide can be made

from a piece of framing 200 cm long?

13. On one side of a scale there are three pots of jam and a 100 g weight. On the other side there are a 200 g and a 500 g weight. The scale is balanced. What is the weight of a pot of jam?
14. A ship was North-West. It made a turn of 90 degrees to the right. An hour later it made a turn through 45 degrees to the left. In what direction was it then traveling?

APPENDIX 4.B. INTERVIEW PROCEDURE MATHEMATICAL PROCESSING INSTRUMENT

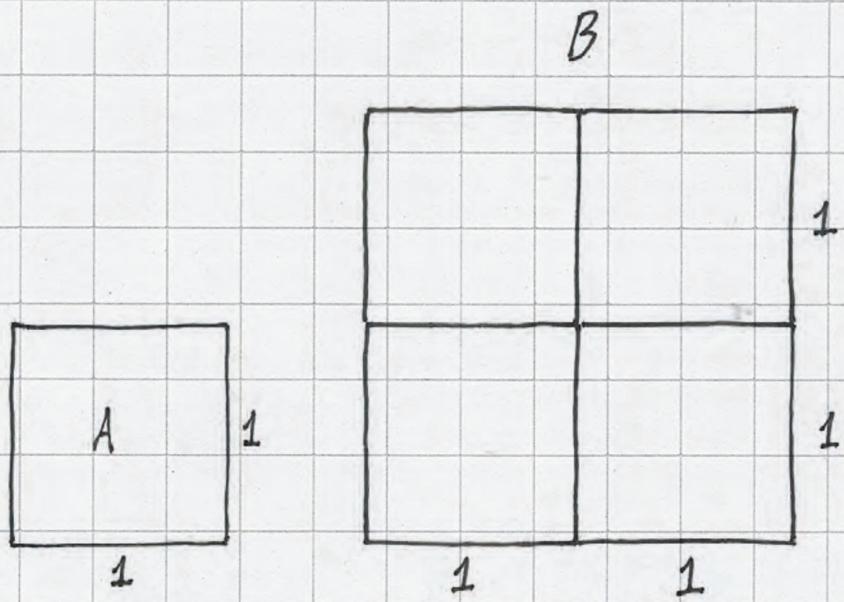
Interview procedure which was followed after each word problem on the Mathematical Processing Instrument.



5 The relation between children's constructive play activities, spatial ability and mathematical word problem solving performance:

A mediation analysis in sixth grade students

Meike Oostermeijer, Anton J. H. Boonen & Jelle Jolles
Frontiers in Psychology (2014), 5:782



A square (A) has an area of 1 square meter. Another square (B) has sides twice as long. What is the area of B?

ABSTRACT

The scientific literature shows that constructive play activities are positively related to children's spatial ability. Likewise, a close positive relation is found between spatial ability and mathematical word problem solving performances. The relation between children's constructive play and their performance on mathematical word problems is, however, not reported yet. The aim of the present study was to investigate whether spatial ability acted as a mediator in the relation between constructive play and mathematical word problem solving performance in 128 sixth grade elementary school children. This mediating role of spatial ability was tested by utilizing the current mediation approaches suggested by Preacher and Hayes. Results showed that 38.16% of the variance in mathematical word problem solving performance is explained by children's constructive play activities and spatial ability. More specifically, spatial ability acted as a partial mediator, explaining 31.58% of the relation between constructive play and mathematical word problem solving performance.

INTRODUCTION

In its home and school environment, almost every child is involved in playing with Legos, blocks and jigsaw puzzles. The term *constructive play*, which has a central role in this study, is often used to categorize these play activities. Constructive play generally involves the manipulation, construction and motion of objects in space (i.e., rotating) (Caldera, Culp, O'Brian, Truglio, Alvarez, & Huston, 1999; Mitchell, 1973; Pomerleau, Malcuit, & Séguin, 1990). The aim of the present study is to examine the link between children's constructive play activities and two interrelated factors, namely spatial ability and mathematical word problem solving performance. Although a positive relation between constructive play and spatial ability is reported by several authors (e.g., Bjorklund & Brown, 1998; Levine, Ratkliff, Huttenlocher, & Cannon, 2012) as well as a positive relation between spatial ability and mathematical word problem solving performance (Beentjes, 2008; Blatto-Vallee, Kelly, Gaustad, Porter, & Fonzi, 2007; Casey, Andrews, Schindler, Kersh, Samper, & Copley, 2008), a relation between constructive play and mathematical word problem solving performance is barely investigated. A possible reason for this absence is that spatial ability acts as a mediator in the relation between children's constructive play activities and their performances on mathematical word problems. The present study is primarily focused on testing this mediating role of spatial ability.

The relation between constructive play and spatial ability

The majority of the studies that examined constructive play has focused on its relation with (the development of) spatial ability (e.g., Bjorklund & Brown, 1998; Grimshaw, Sitarenios, & Finegan, 1995; Levine et al., 2012). Spatial ability involves the ability to represent, modify, generate and recall symbolic, non-linguistic information (Hegarty & Waller, 2005; Linn & Petersen, 1985; Tracy, 1987). Generally, three categories of spatial ability are distinguished in the literature, namely spatial perception, spatial visualization, and mental rotation (Hegarty & Waller, 2005; Linn & Petersen, 1985). Spatial perception involves determining spatial relationships with respect to the orientation of one's own body, in spite of distracting information. Spatial visualization is commonly associated with tasks that involve complicated, multistep manipulations of spatially presented information. Mental rotation includes the ability to mentally remember and subsequently rotate an object in the space (Hegarty & Waller, 2005; Linn & Petersen, 1985). Numerous studies have demonstrated that constructive play activities contribute to the development of spatial ability, in specific mental rotation (Bjorklund & Brown, 1998; Brosnan, 1998; Caldera et al., 1999; Grimshaw et al., 1995; Levine et al., 2012; Tracy, 1987; Wolfgang, Stannard, & Jones, 2001). In the present study spatial ability is, therefore, referred to as the performance on mental rotation tasks.

According to the scientific literature, constructive play activities like Legos, blocks and jigsaw puzzles exert the most influence on spatial ability (Caldera et al., 1999; Levine et al., 2012; Mitchell, 1973; Pomerleau et al., 1990). For example, evidence shows that the more children play with Legos, the more they improve in their spatial skills (Brosnan, 1998; Wolfgang, Stannard, & Jones, 2003). Besides playing with Legos, also block play has shown a positive relation with children's spatial ability (Caldera et al., 1999; Sprafkin, Serbin, Denier, & Connor, 1983; Tracy, 1987). Preschool children that are more interested in block play and reproducing complex block models perform better on spatial ability tasks. Also jigsaw puzzles are examined in relation with spatial ability. Recent research of Levine et al. (2012) has revealed that the frequency of playing with jigsaw puzzles contributed to the development of spatial ability. Jigsaw puzzles appear to appeal to both the mental and physical rotation of the pieces to fit them into different places.

The relation between spatial ability and mathematical word problem solving

Besides the positive relation between constructive play and children's spatial ability, a positive relation between spatial ability and mathematical ability, particularly mathematical word problem solving, is also reported in several studies (Beentjes, 2008; Blatto-Vallee, Kelly, Gaustad, Porter, & Fonzi, 2007; Casey, Pezaris, & Nutall, 1992; Guay & McDaniel, 1977; Hegarty & Kozhevnikov, 1999; Kozhevnikov, Hegarty, & Mayer, 2002; Lean & Clements, 1981; Tracy, 1987). Blatto-Vallee et al. (2007) showed, for example, that spatial ability explained almost 20% of unique variance in mathematical word problem solving performance. Casey and colleagues reported that the direct role of spatial ability in mathematical word problem solving lies in performing the actual mathematical operations and numerical reasoning (e.g., Casey et al., 2008; Casey, Nuttall, & Pezaris, 1997, 2001). Other studies have shown the importance of spatial ability in the production of visual-schematic representations (e.g., Hegarty & Kozhevnikov, 1999; Krawec, 2010; Van Garderen, 2006). In order to facilitate the understanding of the text base of a mathematical word problem, one has to make a coherent visual representation of the essential information of the problem. These visual representations include the spatial relations between solution-relevant elements of the word problem text (e.g., Hegarty & Kozhevnikov, 1999; Thevenot & Oakhill, 2006; Thevenot, 2010; Van Garderen, 2006; Van Garderen & Montague, 2003). To be able to make these types of representations, spatial ability is needed. So, children with good spatial skills are better able to make visual-schematic representations than children with poor spatial skills (e.g., Hegarty & Kozhevnikov, 1999; Krawec, 2010; Van Garderen, 2006; Van Garderen & Montague, 2003). The production of visual-schematic representations is found to be positively related to the performance on mathematical word problems (Krawec, 2010; Van Garderen, 2006; Van Garderen & Montague, 2003).

The present study

In summary, the scientific literature reports a positive relation between children's constructive play activities and spatial ability. Spatial ability increases as children engage more in playing with Legos, blocks and jigsaw puzzles (Hegarty & Waller, 2005; Linn & Petersen, 1985; Tracy, 1987). Moreover, the positive relation between spatial ability and mathematical word problem solving performance is also commonly investigated (Casey et al., 1997, 2001, 2008; Hegarty & Kozhevnikov, 1999). A relation between constructive play and mathematical word problem solving performance is, however, not reported yet. A limited amount of studies have shown a positive relation between constructive play and more general math skills (Caruso, 1993; Serbin, & Connor, 1979; Wolfgang et al., 2001). For example, the studies of Wolfgang et al. (2001) and Beentjes (2008) revealed that block play among preschoolers was a predictor of later school achievement in mathematics, when controlled for IQ and gender. All these studies did, however, not have a focus on mathematical word problem solving in particular. To our knowledge, this is one of the first studies that investigates the relation between constructive play and mathematical word problem solving performance with spatial ability serving as a mediator. According to the statistical literature, a mediator explains the relation between the independent and the dependent variable. Rather than hypothesizing a direct causal relationship between the independent variable and the dependent variable, a mediational model hypothesizes that the independent variable causes the mediator variable, which in turn causes the dependent variable (for more information see Hayes, 2009; Preacher & Hayes, 2008; Shrout & Bolger, 2002).

The mediating role of spatial ability in the relation between constructive play and mathematical word problem solving performance is reflected in the hypothetical (mediation) model reported in Figure 1.

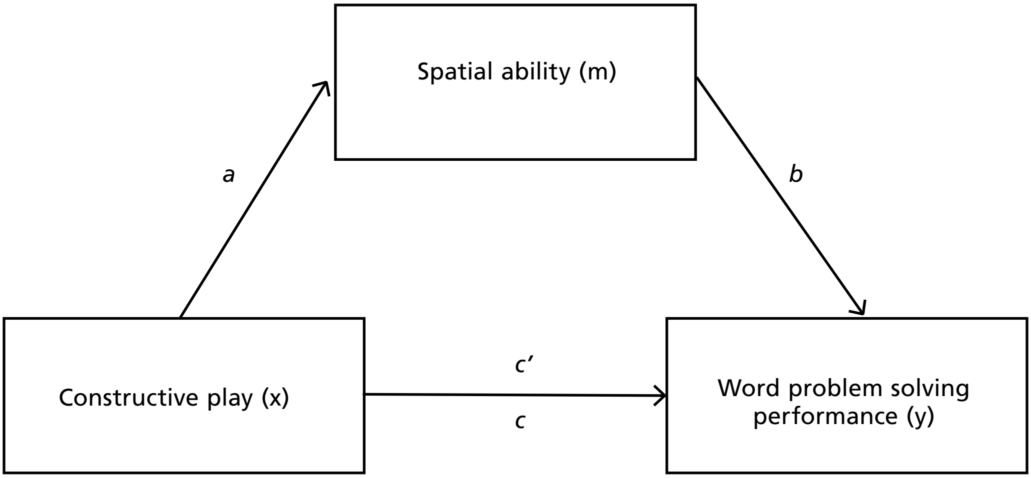


Figure 1. Hypothesized mediation model including the independent variable (i.e., constructive play, x), mediator (i.e., spatial ability, m), and dependent variable (i.e., word problem solving performance, y)

As studies have demonstrated that there is a difference in the extent in which boys and girls are engaged in constructive play activities (see e.g., Serbin & Connor, 1979; Scholten, 2008; Tracy, 1987), sex is added as a covariate to the mediation model.

METHODS

Participants

This study contained data from 128 Dutch sixth grade children (64 boys, M_{age} = 11.73 years, SD_{age} = 0.43 years and 64 girls, M_{age} = 11.72 years, SD_{age} = 0.39 years) from eight elementary schools in The Netherlands. Parents/caretakers provided written informed consent based on printed information about the purpose of the study.

Instruments and measurement procedure

Children's mathematical word problem solving performance and spatial ability were administered by three trained independent research-assistants in a session of approximately 25 minutes. Constructive play was examined with a questionnaire filled out by one of the parents/caretakers.

Mathematical word problem solving performance

Mathematical word problem solving performance was examined with the Mathematical Processing Instrument (MPI), which was first translated to Dutch. The MPI consisted of 14 mathematical word problems based on previous studies (Hegarty & Kozhevnikov, 1999; Van Garderen & Montague, 2003, see Appendix A). The internal consistency (Cronbach's alpha) of this instrument, measured in American participants, is .78 (Hegarty & Kozhevnikov, 1999). The Cronbach's alpha of the MPI in this study was .72. The word problems were printed on cards and presented in four different orders. All problems were read out loud to the children to control for differences in decoding skill. Furthermore, children were allowed to solve each word problem within a maximum of three minutes and during this time the experimenter did not speak to the child. To be sure that children had enough time to solve the mathematical word problems, a pilot study was conducted with five sixth grade students. The results of the pilot study showed that every child was able to solve each of the 14 items of the MPI within the required three minutes. The number of mathematical word problems solved correctly was used as the dependent variable in the analyses.

Constructive play

In order to determine to what extent children show constructive play behavior, a short questionnaire was forwarded to their parents/caretakers. They were asked to indicate on a 4 point Likert scale (1 = never, 4 = often) to what extent their child has undertaken the, for this study, representative constructive play activities (i.e., playing with Legos, blocks and jigsaw puzzles). The internal consistency of this questionnaire was sufficient (Cronbach's alpha = .71). A sum score was created by adding the scores on the three, for this study representative, constructive play activities. The higher the sum score, the more the student is involved in constructive play activities. The sum score was added as the independent variable in the analyses.

Spatial ability
The Picture Rotation task (Quaiser-Pohl, 2003) is a standardized task that was used to measure mental rotation. In the Picture Rotation task children were asked to rotate a non-manipulated picture of an animal at the left of a vertical line. The three pictures at the right of the vertical line showed the rotated and/or mirrored image of that same animal. One of these three pictures was only rotated; two of these pictures were both rotated and mirrored. Children had to decide which of the three pictures was only rotated. Children had 1.5 minutes to finish this task. The internal consistency of this measure in the present study was high (Cronbach's alpha = .93). Figure 2 shows one of the 30 test items of the Picture Rotation task. The accuracy on this task was used as the mediator in the analyses.

Data analysis

The mediating effect of constructive play on word problem solving performance via spatial ability was tested using bootstrap methods (Hayes, 2009; Shrout & Bolger, 2002). Bootstrap method has been validated in the literature and is preferred over other methods in assessing the existence of mediation among variables. Preference based on the fact that other methods for testing indirect effects

Figure 2. The Picture Rotation Task (based on Quaiser-Pohl 2003)

assume a standard normal distribution when calculating the *p*-value for the indirect effect, whereas bootstrapping does not assume normality of the sampling distribution. In addition, bootstrap method repeatedly samples from the data set, estimating the indirect effect with each resampled data set. This process is repeated thousands of times, producing bias-correct accelerated confidence intervals for the indirect effect (Preacher & Hayes, 2008).

RESULTS

Descriptive statistics

Table 1 presents the correlations between, and the means and standard deviations of, the four measures of this study. This table shows that the correlations between the constructive play, spatial ability and word problem solving performance are moderate to strong. No significant correlation is found between sex and the three key measures of this study.

Table 1. *Intercorrelations, means, standard deviations, ranges of the measures of this study*

	1.	2.	3.	4.	*N*	*M*	*SD*	*Range*
1. Constructive play	-				73	7.22	2.22	9.00
2. Spatial ability	.25*	-			128	13.25	7.45	27.00
3. Word problem solving performance	.35**	.55**	-		128	6.68	2.87	14.00
4. Sex	-.16	-.14	-.16	-	128	-	-	-

* $p < .05$, ** $p < .001$

Investigating the mediating role of spatial ability

Mediation was tested by regressing the dependent variable (i.e., word problem solving performance) on spatial ability in the presence of constructive play and sex. Analyses utilizing the bootstrap method (5.000 bootstrap samples were used) confirmed the existence of a mediation effect of constructive play on word problem solving performance via spatial ability (see Table 2). However, the results showed that there is a partial, but not complete mediation, because the measured effect between constructive play and word problem solving performance is not zero upon fixing the mediator variable (i.e., spatial ability, Preacher & Hayes, 2008). The value of the indirect effect of spatial ability can be calculated as follows:

$B_{\text{indirect}} = B_{(a)*} B_{(b)} = 0.75 \text{ x } 0.16 = 0.12$, and
$B_{\text{indirect}} / B_{\text{total}} = 0.12/ 0.38 = 0.3158.$

Thus, spatial ability explained 31.58% of the relation between constructive play and students' mathematical word problem solving performance. The absence of a zero in the confidence interval for the indirect pathways indicated that the indirect effect was significantly different from zero at $p < .05$, two tailed.

The complete model (including constructive play, spatial ability and sex) explained 38.16% of the variance in students word problem solving performance ($R^2 = .38$), which is a large effect (Fairchild, Mackinnon, Taborga, & Taylor, 2009; Green & Salkind, 2008).

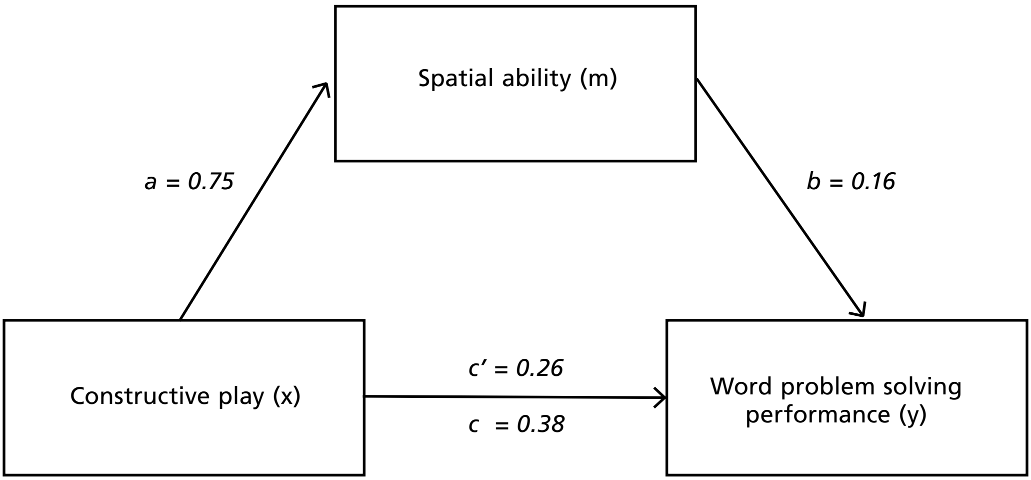


Figure 3. Results of the mediation analysis (N = 128). Sex was included in the equations as a statistical control but is not presented for reasons of clarity

DISCUSSION

The purpose of the present study was to investigate if spatial ability acts as a mediator in the relation between constructive play and mathematical word problem solving performance in sixth grade elementary school children. To our knowledge, this is one of the first studies that examined the mediating role of spatial ability in

Table 2. *Parameter estimates of the model examining the mediation role of spatial ability in the relation between constructive play and word problem solving performance.*

Model	Estimate	*SE*	*p*	CI (lower)	CI (upper)
Model without mediator					
Intercept	4.57	1.11	< .01	2.36	6.78
CP → WPS *(c)*	0.38	0.14	< .01	0.11	0.66
Sex → WPS	-1.44	0.60	< .05	-2.63	-0.24
R^2 (y, x)	.19				
Model with mediator					
Intercept	2.86	1.04	< .01	0.78	4.94
Model 1: SP as dependent variable					
CP → SP *(a)*	0.75	0.41	< .05	-0.06	1.56
Sex → SP	-3.33	1.79	.07	-6.90	0.25
Model 2: WPS as dependent variable					
SP → WPS *(b)*	0.16	0.04	< .001	0.09	0.23
CP → WPS *(c')*	0.26	0.12	< .05	0.02	0.51
Sex → WPS	-0.89	0.54	.10	-1.97	0.18
Indirect effects *(a x b)*	0.12	0.08	< .05	0.128	0.27
R^2 (m, x)	.10				
R^2 (y, m, x)	.38				

Note: regression weights *a, b, c, and c'* are illustrated in Figure 3. R^2(y, x) is the proportion of variance in y explained by x, R^2 (m, x) is the proportion of variance in m explained by x and m. the 95% CI for *a* x *b* is obtained by the bias-corrected bootstrap wit 5000 resamples. CP (Constructive play) is the independent variable (x), SP (Spatial ability) is the mediator (m), and WPS (Word problem solving performance) is the outcome (y). CI (lower = lower bound of 95% confidence interval; CI (upper) = upper bound.

this particular relation. In previous studies, relations between constructive play and spatial ability (e.g., Bjorklund & Brown, 1998; Brosnan, 1998), and between spatial ability and mathematical word problem solving performance (e.g., Blatto-Vallee et al., 2007; Hegarty & Kozhevnikov, 1999; Van Garderen & Montague, 2003) are reported. The relation between constructive play and mathematical word problem solving performance has however not been established yet.

The results of this study showed that spatial ability acted as a partial mediator in the relation between constructive play and children's mathematical word problem solving performance. This implies that children who were frequently engaged in constructive play in their past have better spatial skills and, as a result, show a higher performance on mathematical word problems. The variables in this study (i.e., constructive play, spatial ability and sex) explained 38.16% of the variance in performance on solving mathematical word problems. Furthermore, 31.58% of the relation between constructive play and mathematical word problem solving performance is explained by spatial ability.

Note that the findings of this study support the assessment of a mediating effect based on current recommendations using bootstrap approaches (Hayes, 2009; Shrout & Bolger, 2002)[3].

Limitations

Three limitations of the study should be mentioned. The first limitation of this study included the fact that only one task was used in the analyses to measure spatial ability (i.e., mental rotation). Ideally, method triangulation should be applied. The use of more tasks allows a more reliable measurement of the construct 'spatial ability' and reduces the chance of possible measurement errors (Wooldering, 2009). The second limitation pertains to the correlational nature of the data, which made it impossible to draw conclusions about any causal relationships between constructive play, spatial ability and mathematical word problem solving performance. The results of this study only showed that these variables were associated with each other. Future experimental studies in which the variables will be manipulated, should make it possible to draw stronger conclusions concerning causal relationships between the aspects which are important in mathematical word problem solving. The last limitation covers the way in which the constructive play activities of the children were administered. In the present study, a third party (i.e., the parents), filled out the questionnaires regarding the extent to which children show constructive play behavior. Although parents were able to provide a reliable image of the

[3.] This assessment of mediation is also support by the statistical approach that Baron & Kenny (1986) used in their research.

constructive play activities in which their children are/were involved, in future studies it would be even more reliable to directly observe these play activities.

Implications and directions for future research

The present study contributed to the increasing amount of scientific literature regarding the processes that are involved in learning mathematics, particularly mathematical word problem solving. An interesting focus of future research is to investigate the existence of individual differences in the specific relations between the three key variables of this study (i.e., constructive play, spatial ability, and mathematical word problem solving performance). Although not supported by the results of the present study, several authors have demonstrated that boys and girls differ in the extent in which they engage in constructive play (Serbin & Connor, 1979; Scholten, 2008; Tracy, 1987). That is, boys tend to play more with so-called 'masculine' or constructive toys, like Legos and blocks, than girls (Serbin & Connor, 1979; Tracy, 1987). Because the scientific literature gives no clear indications that sex differences exist in spatial ability (e.g., McGee, 1979; Voyer, Voyer, & Bryden, 1995), examining the mediating role of spatial ability for both boys and girls separately might be an interesting topic for follow-up studies.

The results of this study also have a strong practical relevance. Parents/caretakers should be aware of the importance of constructive play activities in childhood. According to the findings of this study, activities like playing with Legos, blocks and jigsaw puzzles, are positively related to students' spatial skills, which, in turn, is positively related to their performance on mathematical word problems. Parents/caretakers should, therefore, create opportunities to play with constructive toys. Also elementary school teachers should provide constructive learning material to their children and stimulate to use it by giving them appropriate instruction. Finally, this research accentuated the importance of spatial ability in mathematical word problem solving performance. In line with previous research (e.g., Hegarty & Kozhevnikov, 1999; Van Garderen, 2006), spatial ability was found to play a key role in solving mathematical word problems, especially in the production of visual-schematic representations. The training of spatial skills and the development of visual-schematic representations should, therefore, have a prominent role in word problem solving instruction of elementary school mathematics education.

APPENDIX 5.A.

Word problems of the Mathematical Processing Instrument

The mathematical word problems on the Mathematical Processing Instrument (Hegarty & Kozhevnikov, 1999):

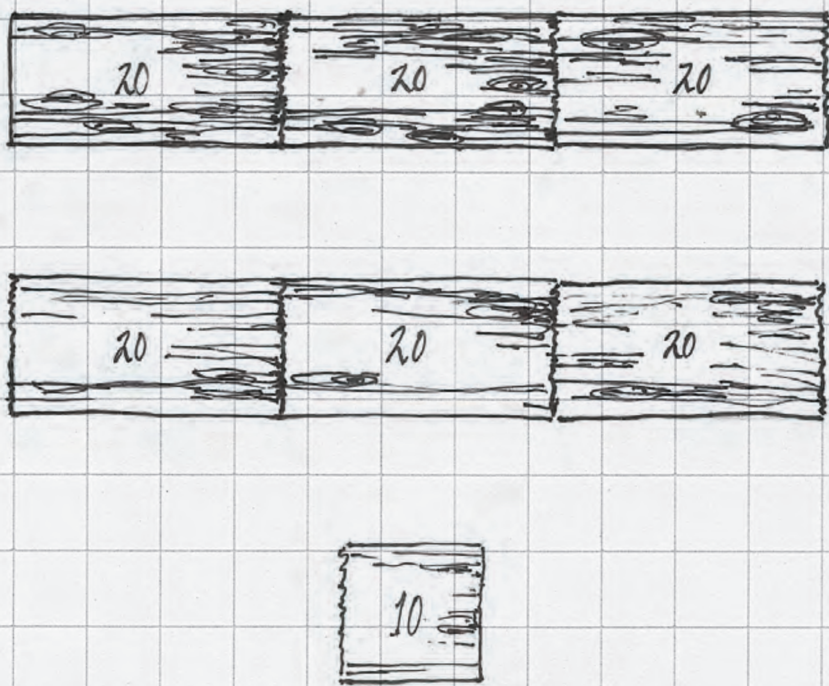
1. At each of the two ends of a straight path, a man planted a tree and then every 5 meters along the path he planted another tree. The length of the path is 15 meters. How many trees were planted?
2. On one side of a scale there is a l kg weight and half a brick. On the other side there is one full brick. The scale is balanced. What is the weight of the brick?
3. A balloon first rose 200 meters from the ground, then moved 100 meters to the east, then dropped 100 meters. It then traveled 50 meters to the east, and finally dropped straight to the ground. How far was the balloon from its original starting point?
4. In an athletics race, Jim is four meters ahead of Tom and Peter is three meters behind Jim. How far is Peter ahead of Tom?
5. A square (A) has an area of 1 square meter. Another square (B) has sides twice as long. What is the area of square B?
6. From a long stick of wood, a man cut 6 short sticks, each 2 feet long. He then found he had a piece of 1 foot long left over. Find the length of the original stick.
7. The area of a rectangular field is 60 square meters. If its length is 10 meters, how far would you have traveled if you walked the whole way around the field?
8. Jack, Paul and Brian all have birthdays on the 1st of January, but Jack is one year older than Paul and Jack is three years younger than Brian. If Brian is 10 years old, how old is Paul?
9. The diameter of a tin of peaches is 10 cm. How many tins will fit in a box 30 cm by 40 cm (one layer only)?
10. Four young trees were set out in a row 10 meters apart. A well was situated beside the last tree. A bucket of water is needed to water two trees. How far would a gardener have to walk altogether if he had to water the four trees using only one bucket?
11. A hitchhiker set out on a journey of 60 miles. He walked the first 5 miles and then got a lift from a lorry driver. When the driver dropped him, he still had half of his journey to travel. How far had he traveled in the lorry?
12. How many picture frames 6 cm long and 4 cm wide can be made from a piece of framing 200 cm long?

13. On one side of a scale there are three pots of jam and a 100 g weight. On the other side there are a 200 g and a 500 g weight. The scale is balanced. What is the weight of a pot of jam?
14. A ship was sailing North-West. It made a turn of 90 degrees to the right. An hour later it made a turn of 45 degrees to the left. In what direction was it then traveling?

6 Second grade elementary school students' differing performance on combine, change and compare word problems

Anton J. H. Boonen, & Jelle Jolles

(under review)



From a long stick of wood, a man cut 6 short sticks, each 2 feet long. He then found he had a piece of 1 foot long left over. Find the length of the original stick.

ABSTRACT

The purpose of the present study was to investigate the word problem solving skills of 47 second-grade students by examining how they performed on combine, change and compare word problems. The results of the repeated measures ANOVA showed that students scored significantly lower on compare problems than on combine and change word problems. Based on the results of this study, we disproved the hypothesis that the students in our sample experienced more difficulties in compare problems as a result of the so-called consistency effect; in fact they performed equally well on inconsistent and consistent compare problems. The findings indicate that the core problem which the students experience might be associated with the fact that they have difficulty in general with processing relational terms like 'more than' and 'less than'. Future studies should, therefore, provide more insight into the reasons why compare problems in particular cause so many difficulties for both young and older students. This information would be helpful when it comes to developing more adequate word problem instructions that can be implemented in the curriculum of contemporary math education.

INTRODUCTION

According to Realistic Math Education (RME), mathematics should be connected to realistic (verbal) contexts, stay close to children, and be relevant to society (Van den Heuvel-Panhuizen, 2003). Math problems in RME (and other contemporary math approaches) are, therefore, generally presented as text rather than in a numerical format. However, students have been shown to experience more difficulties with solving these so-called word problems already in the first grades of elementary school (Cummins, Kintsch, Reusser, & Weimer, 1988; Hegarty, Mayer, & Green, 1992; Hegarty, Mayer, & Monk, 1995). This discrepancy between performance on verbal and numerical format problems strongly suggests that factors other than calculation ability contribute to children's word problem solving success (Boonen, Van der Schoot, Van Wesel, De Vries, & Jolles, 2013; Cummins et al, 1988; Van der Schoot, Bakker-Arkema, Horsley, & Van Lieshout, 2009).

Many previous studies report that an important factor in how students perform on word problems is their comprehension of the text of a word problem (Boonen et al., 2013; Cummins et al., 1988; Krawec, 2010, 2012; Lewis & Mayer, 1987). The comprehension of a word problem mainly concerns the identification of (verbal and numerical) relations between the elements that are relevant for the solution, as these are used in the construction of a visual representation that reflects the structure of the word problem (Tolar, Fuchs, Cirino, Fuchs, Hamlett, & Fletcher, 2012; Cummins et al., 1988; Hegarty et al., 1995; Pape, 2003). More specifically, the verbal and numerical information that is relevant for the solution of the word problem should be connected and included in a visual representation, in order to clarify the problem situation described in the word problem (Boonen et al., 2013; Hegarty & Kozhevnikov, 1999; Thevenot, 2010).

In the early grades of elementary school, three types of word problems are frequently offered to the students, namely, combine, change and compare problems. These three specific types of word problems play a key role in several scientific studies investigating students' word problem solving performance (Cummins et al., 1988; Jitendra, 2002; Jitendra et al., 2013; Tolar et al., 2012). Because combine, change and compare problems also play a central role in the present study, an explanation of each of these types of word problems is given below.

In the combine word problem, reflected in the first word problem example, a subset or superset must be computed given the information about two other sets.

[Combine word problem]
Mary has 3 marbles. John has 5 marbles. How many marbles do they have altogether?

This type of problem involves understanding part-whole relationships and knowing that the whole is equal to the sum of its parts (Cummins et al., 1988; Jitendra, 2002, Jitendra, DiPipi, & Perron-Jones, 2002). Figure 1 reflects a possible way in which the problem structure of a combine problem can be represented.

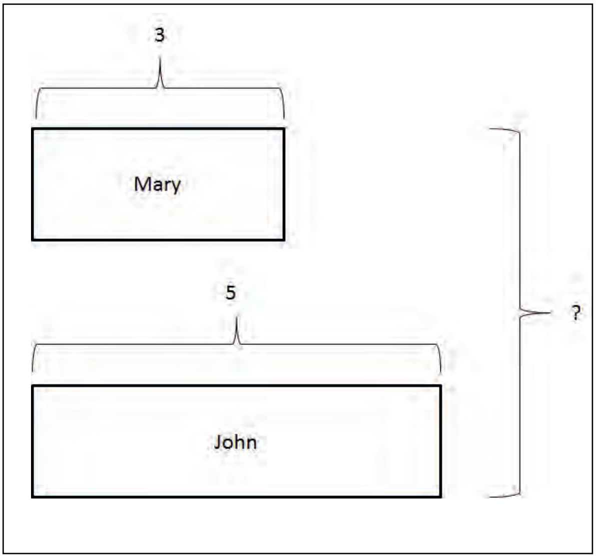


Figure 1. Visual-schematic representation of the problem structure of a combine problem

The second commonly investigated type of word problem is a change problem (see word problem below).

[Change word problem]
Mary had 3 marbles. Then John gave her 5 marbles. How many marbles does Mary have now?

Change problems are word problems in which a starting set undergoes a transfer-in or transfer-out of items, and the cardinality of a start set, transfer set, or a result set must be computed given information about two of the sets (Cummins et al.,1988; Jitendra et al., 2013). In other words, a change problem starts with a beginning set in which the object identity and the amount of the object are defined. Then a change occurs to the beginning set that results in an 'ending set' in which the new amount is defined (Jitendra, 2002). In Figure 2 an appropriate visual-schematic representation of the

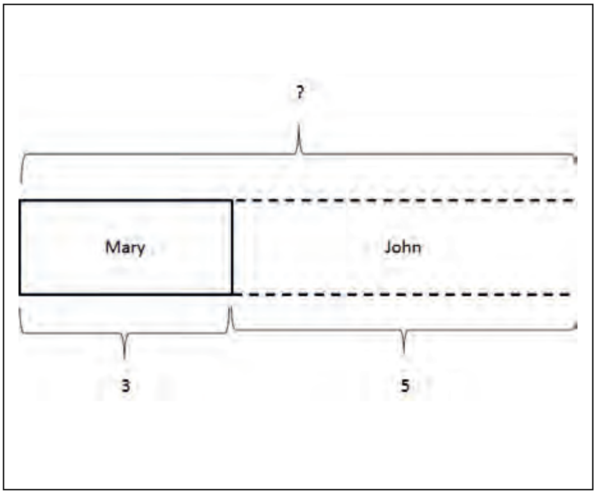


Figure 2. Visual-schematic representation of the problem structure of a change problem

problem structure of a change problem is given.
The last word problem type that is investigated in many studies is a compare problem (see word problem below).

> ***[Compare word problem]***
> *Mary has 5 marbles. John has 8 marbles. How many marbles does John have more than Mary?*

In compare problems the cardinality of one set must be computed by comparing the information given about relative sizes of the other set sizes; one set serves as the comparison set and the other as the referent set. In this type of word problem, students often focus on relational terms like 'more than' or 'less than' to compare the two sets and identify the difference in value between the two sets (Cummins et al., 1988; Hegarty et al., 1995; Pape, 2003; Van der Schoot et al., 2009). Figure 3 reflects the visual-schematic representation of the problem structure of a compare problem.

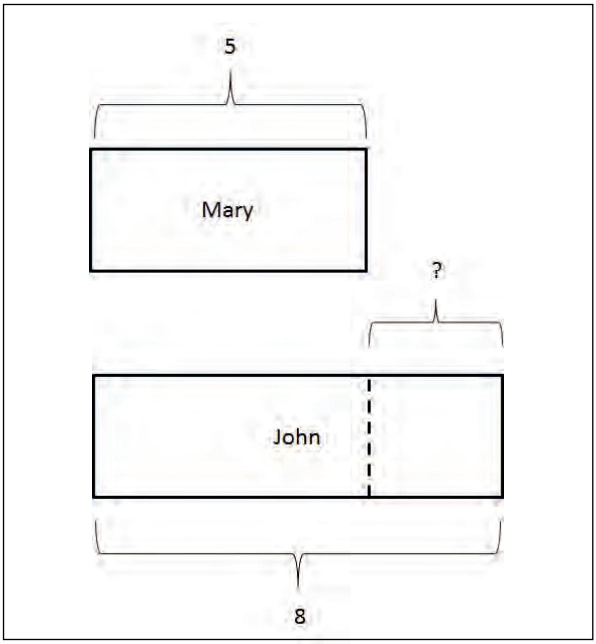


Figure 3. Visual-schematic representation of a compare problem

Research by Cummins et al., performed in the nineteen eighties, showed that first grade students rarely make errors on combine and change word problems, but that difficulties often arise when these students have to solve compare problems (Cummins et al., 1988). More recent studies mainly focused on how older students, namely sixth and seventh grade students (Pape, 2003; Van der Schoot et al., 2009), and undergraduates (Hegarty et al., 1992, 1995) performed compare problems.

The first explanation offered for these difficulties with solving compare problems is the hypothesis that young students have not yet understood that the quantitative difference between the same sets can be expressed in parallel ways with both the terms *more* and *fewer.* Their lack of knowledge and experience with the use of language to describe relations between quantities could underlie their relatively poor performance in solving compare problems. Notably, the lack of knowledge about the symmetry of language in the case of quantitative comparisons makes it difficult for young students to perform the translation procedure correctly (d'Ailly, Simpson, & McKinnon, 1997).

A second possible explanation for this difficulty with compare problems might be the extent to which the semantic relations between the given and unknown quantities of the problem are made explicit (De Corte, Verschaffel, & De Win, 1985; De Corte, Verschaffel, Pauwels, 1990; Marzocchi, Lucangeli, De Meo, Fini, & Cornoldi, 2002; Verschaffel, De Corte, & Pauwels, 1992).

A third frequently investigated hypothesis that might explain the difficulties with solving compare problems involves examining whether the relational keyword of the compare problem ('more than' or 'less than') is consistent or inconsistent with the required mathematical operation (Van der Schoot et al., 2009). In so-called inconsistent compare problems (Hegarty et al., 1992, 1995; Kintsch, 1998; Van der Schoot et al., 2009), the crucial mathematical operation cannot be simply derived from the relational keyword ('more than'). The relational term in an inconsistent compare problem primes an inappropriate mathematical operation, e.g., the relational term 'more than' evokes an addition operation, while the required operation is subtraction. This accounts for the difficulty with solving inconsistent compare problems. The finding that students make more errors on inconsistent than on consistent compare problems is referred to as the 'consistency effect' (Lewis & Mayer, 1987; Pape, 2003; Van der Schoot et al., 2009). Interestingly, the consistency effect has until now only been examined in students in higher elementary school grades and at university (Pape, 2003; Van der Schoot et al., 2009). This raises the question whether the difficulties that young students (i.e., in lower grades of elementary school) experience with solving compare problems are confined to inconsistent compare problems, as is often the case with older students. Or, do young students experience difficulty in general with processing the verbal information contained in a compare problem?

The present study

Combine, change and compare problems are more frequently offered in the early grades than in the later grades of elementary school. It is therefore valuable to investigate how young elementary school children's performance on combine and change problems differs from their performance on compare problems. The only previous study on this topic was conducted in the nineteen eighties (viz., the leading but somewhat outdated study by Cummins et al. 1988); hence, it is relevant to evaluate whether these findings are still valid after 25 years which have seen significant adaptations in school curricula as well as changes in society.

We hypothesized that students will perform poorer on compare problems than on combine and change problems, and examined whether young students experience more difficulties with solving compare problems because of a consistency effect. Based on the findings of previous studies of older students (e.g., Pape, 2003; Van der Schoot et al., 2009), we hypothesized that also younger students would make more errors on inconsistent compare problems than on consistent compare problems.

METHODS

Participants

Forty-seven second-grade students (26 boys, 21 girls) from two classes from a mainstream elementary school in the Netherlands participated in this study. The mean chronological age of the students was 89 months (*SD* = 4 months; range: 79 – 96 months). Parents provided written informed consent based on printed information about the purpose of the study.

Instruments and procedure

Word problem solving performance

Students' performances on the three different types of word problems (combine, change, & compare problems) were examined with an 18-item Word problem solving test (taken from Cummins et al., 1988, see Table 1). The WPS test was divided into two subtests containing nine word problems (three of each type of word problem). The items of each WPS subtest were presented on a different page and administered by the teacher in two classroom sessions of approximately 30 minutes. Each word problem was read out loud twice to the students to control for differences in decoding skills. After reading the word problem, students had to solve the word problem within three minutes and during this time the teacher did not speak to the student.

To examine the consistency effect in the compare word problems, both consistent and inconsistent compare problems were offered to the students. Consistency referred to whether the relational term ('more than' or 'less than') in the word problem was consistent or inconsistent with the required mathematical operation. The relational term in a consistent compare problem primed the appropriate mathematical operation (e.g., 'more than' when the required operation is addition, and 'less than' when the required operation is subtraction). The relational term in an inconsistent compare problem primed the inappropriate mathematical operation ('more than' when the required operation is subtraction, and 'less than' when the required operation is addition). The internal consistency (Cronbach's alpha) of the WPS test, measured in this study, was high (Cronbach's alpha = .82).

Table 1. *The 18 items of the Word problem solving test (taken from Cummins et al., 1988)*

Problem type	*Word problem*
Combine	1. Mary has 2 marbles. John has 5 marbles. How many marbles do they have altogether?
	2. Mary and John have some marbles altogether. Mary has 2 marbles. John has 4 marbles. How many marbles doe they have altogether?
	3. Mary has 4 marbles. John has some marbles. They have 7 marbles altogether. How many marbles does John have?
	4. Mary has some marbles. John has 6 marbles. They have 9 marbles altogether. How many marbles does Mary have?
	5. Mary and John have 8 marbles altogether. Mary has 7 marbles. How many marbles does John have?
	6. Mary and John have 4 marbles altogether. Mary has some marbles. John has 3 marbles. How many marbles does Mary have?
Change	
	1. Mary had 3 marbles. Then John gave her 5 marbles. How many marbles does Mary have now?
	2. Mary had 6 marbles. Then she gave 4 marbles to John. How many marbles does Mary have now?
	3. Mary had 2 marbles. Then John gave her some marbles. Now Mary has 9 marbles. Hoe many marbles did John give to her?
	4. Mary had 8 marbles. Then she gave some marbles to John. Now Mary has 3 marbles. How many marbles did she give to John?
	5. Mary had some marbles. Then John gave her 3 marbles. Now Mary has 5 marbles. How many marbles did Mary have in the beginning?
	6. Mary had some marbles. Then she gave 2 marbles to John. Now Mary has 6 marbles. How many marbles did she have in the beginning?
Compare	
	1. Mary has 5 marbles. John has 8 marbles. How many marbles does John have more than Mary?*
	2. Mary has 6 marbles. John has 2 marbles. How many marbles does John have less than Mary?
	3. Mary has 3 marbles. John has 4 marbles more than Mary. How many marbles does John have?
	4. Mary has 5 marbles. John has 3 marbles less than Mary. How many marbles does John have?
	5. Mary has 9 marbles. She has 4 marbles more than John. How many marbles does John have?*
	6. Mary has 4 marbles. She has 3 marbles less than John. How many marbles does John have?*

Note: inconsistent compare problems are indicated with an asterisk.

6

Data analysis

To examine students' performance on the three types of word problems, a repeated measures analysis of variance (ANOVA) with type of word problem (combine, change and compare) as within subject factor was performed. Follow-up tests were performed using paired sample t-tests. Subsequently, a one sample t-test was performed to examine the existence of a consistency effect; the performance on consistent compare problems was compared with the performance on inconsistent word problems. In all analyses we tested with an alpha of .05. Effect sizes (partial eta-squared [η_p^2]) were computed to estimate the practical significance of the effects.

RESULTS

Performance on combine, change and compare word problems

Results of the repeated measures ANOVA demonstrated a significant main effect of word problem type, $F(2,92) = 12.90$, $p < .001$, $\eta_p^2 = .36$, indicating a large effect size, see Pierce, Block, and Aguinis (2004). Figure 4 shows the accuracy on each of the three types of word problems (combine word problems, $M = 4.89$, $SD = 1.46$; change problems, $M = 4.85$, $SD = 1.43$; compare problems, $M = 3.81$, $SD = 1.53$). In line with our expectations, second grade students scored significantly lower on compare word problems than on combine (t (46) = 4.69, $p < .001$) and change (t (46) = 4.90, $p < .001$) word problems. No differences in students' performance on combine and change problems existed (t (46) = 0.27, $p = .79$).

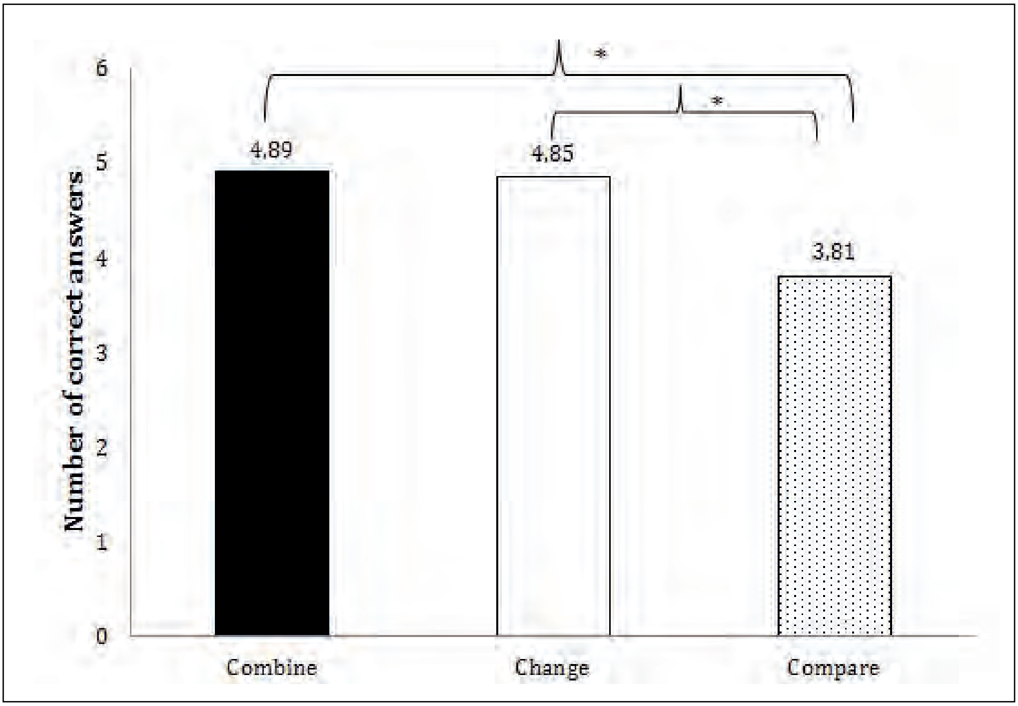


Figure 4. Results repeated measures ANOVA: students' accuracy on each word problem type, * $p < .001$

A consistency effect in compare word problems

The one sample t-test on students' performance on compare problems revealed no main effect of consistency, $t(46) = .15$, $p = .88$, indicating that a consistency effect was absent in our sample. This finding showed that students performed equally on consistent ($M = 1.91$, $SD = 0.83$) and inconsistent ($M = 1.89$, $SD = 0.98$) compare problems.

DISCUSSION

The present study aimed to provide clues as to why young, second grade students in particular experience more difficulty with compare problems than with combine and change problems. As expected, second grade students made more errors on compare problems than on the other two types of word problems. Importantly, a consistency effect in compare problems was not found in our study; the second grade students performed equally well on inconsistent and consistent compare problems[4]. A difficulty in general with processing relational terms like 'more than' and 'less than', is a plausible explanation for the lack of the consistency effect in our findings. Students in the lower elementary school grades might not yet possess the conceptual knowledge required to fully understand compare problems and this might explain their difficulties with solving this particular type of word problem (Cummins et al., 1988; Riley, Greeno, & Heller, 1983). They apparently do not have the knowledge to comprehend and process the linguistic input of a compare problem and recall the appropriate problem structure (Koedinger & Nathan, 2004). For example, a child may understand the part-whole relationship of a combine problem, but not yet understand how the comparative verbal form (e.g., how many more Xs than Ys) maps onto the sets (Cummins et al., 1988). As already mentioned by d'Ailley et al., (1997) students might have difficulties understanding the fact that the quantitative difference between the same sets could be expressed in parallel ways with both the terms 'more' and 'fewer'. Hence, the poorer performance on compare problems, which was found in this study, might be explained by a lack of knowledge about the symmetry of language in the case of quantitative comparisons; this makes it more difficult for young students to perform the translation procedure correctly.

Future studies should, for example, examine the reasons why compare problems in particular cause so many difficulties in young students, and evaluate the possible influence of the development of higher language skills. Research has indicated that the comprehension and processing speed of complex language (i.e., students mastery of relational terms which describe linguistic relations between elements that are relevant for the solution) continue to develop beyond childhood and into adolescence (Wassenberg, 2007; Wassenberg et al., 2008).

As scientific research during the last decades has shown that the difficulties students experience when solving compare problems remain stable over time (Cummins et al., 1988; Pape, 2003; Van der Schoot et al., 2009) another important topic for future research

[4]. Because just a small number of items (i.e., three consistent and three inconsistent compare problems) were included the statistical analysis, this finding should be interpreted with caution.

would be to focus on the development of effective word problem solving instruction. Adequate word problem solving instructional programs that teach students to solve these types of problems are still limited, or they have not been implemented in the educational practice of elementary schools.

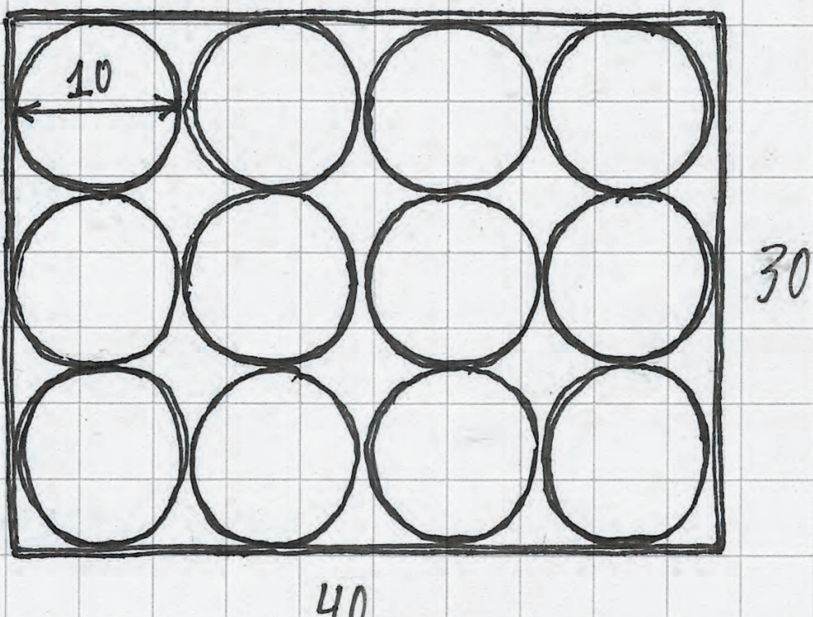
Instructional approaches, like *Schema-Based Instruction* and the *Solve It!* method, that focus on explicit instruction in cognitive and metacognitive strategies to help students identify and represent the problem structure and improve their word problem solving performance, seem promising (e.g., Jitendra, DiPipi, & Perron-Jones, 2002, Jitendra et al., 2013; Jitendra, & Star, 2012; Jitendra et al., 2009; Krawec, 2010; Montague, Warger, & Morgan, 2000). These instructional approaches move away from keywords and superficial problem features and focus more on helping children find the underlying problem structure.

In SBI, for example, students are taught to identify and represent the problem structures of certain types of word problems (i.e., combine, change and compare problems) and are encouraged to reflect on the similarities and differences between these problem types (Jitendra et al. 2002, 2009). However, the instructional programs SBI and Solve it! are generally only used by researchers. Therefore, educational practice in regular elementary school classrooms might be improved if teachers were to implement and work with word problem instruction as well. One of the main hurdles encountered during the implementation of these instructional approaches is that they require greater effort and good classroom management skills (Montague et al., 2000). This is an important reason why the effectiveness of SBI and Solve it! has until now been mainly investigated in small groups of children with learning and mathematical disabilities (Jitendra et al., 2002, 2013) and not in a regular classroom setting. Therefore, before these kinds of instructional programs can be implemented in the curriculum of contemporary math education, it is essential that they are made easy to understand for both students and teachers, and that they can be implemented with a relatively small amount of effort.

7 It's not a math lesson - we're learning to draw!

Teachers' use of visual representations in instructing word problem solving in sixth grade of elementary school

Anton J. H. Boonen, Helen. C. Reed,
Judith. I. Schoonenboom, & Jelle Jolles
(submitted)



The diameter of a tin of peaches is 10 cm. How many tins will fit in a box 30 cm by 40 cm (one layer only)?

ABSTRACT

Non-routine word problem solving contributes to the mathematical development of elementary school students. Teachers' modelling of word problems by constructing visual-schematic representations is found to be a helpful tool. Existing instructional programs, based on heuristic approaches, do not provide sufficient guidance for students to produce visual-schematic representations. The goal of the present study was to examine *teachers' use of visual-schematic representations* when implementing a teaching intervention for supporting non-routine word problem solving. The eight participating teachers in this study were, after a short training, able to produce visual-schematic representations during their instruction. However, teachers seemed to base their use of these representations on personal preferences rather than on an optimal fit with the word problem characteristics. This should be a key aspect for teacher training.

INTRODUCTION

[Word problem example]:
On one side of a scale there are three pots of jam and a 100g weight. On the other side there are a 200g and a 500g weight. The scale is balanced. What is the weight of a pot of jam?

In contemporary math education, word problems like the one above are frequently offered to students. Learning to solve these so-called *non-routine word problems* is an essential feature of mathematical development (Depaepe, De Corte, & Verschaffel, 2010; Jiminéz & Verschaffel, 2014; Swanson, Lussier, & Orosco, 2013). In this study, non-routine word problems are defined as challenging problems set in realistic contexts, that require understanding, analysis and interpretation. They are not simple computational tasks embedded in words; they require an appropriate selection of strategies and decisions that lead to a logical solution (Van Garderen & Montague, 2003).

Students' difficulties in solving non-routine word problems are common in contemporary math classrooms (Boonen, Van der Schoot, Van Wesel, De Vries, & Jolles, 2013; Hegarty & Kozhevnikov, 1999; Krawec, 2010; Van Garderen & Montague, 2003) and are widely recognised by both researchers and teachers (Boonen, Van Wesel, Jolles, & Van der Schoot, 2014; Prenger, 2005; Van Garderen, 2006; Van Garderen & Montague, 2003). For this reason, instructional programs have been developed that provide support in word problem solving for low-performing students (see Jitendra, 2002; Jitendra et al., 2013; Montague, 2003; Montague, Warger, & Morgan, 2000). These programs have primarily been used in research settings involving researchers working with individuals or small groups. Surprisingly, there is currently no comparable instructional support available for teachers in mainstream classrooms, as far as we know. This is an important lacuna, given that mainstream schools are becoming more inclusive, with greater numbers of students with mild to severe learning difficulties in classrooms (Jitendra & Star, 2012; Sharma, Loreman, & Forlin, 2012). Dutch teachers, for example, label, on average, a quarter of their students as students with special educational needs in mainstream primary education (Van der Veen, Smeets, & Derriks 2010). It would therefore be of benefit to teachers if they had instructional approaches at their disposal designed to support learning in areas that many students find difficult - in this case, non-routine word problem solving.

Within this context, the present study introduces an innovative approach to instructing word problem solving - based on the use of visual representations - during whole-class teaching in mainstream sixth grade classrooms, and examines how teachers implement that approach. In sixth grade, students are expected to be able to solve a wide variety of non-routine problems of increasing difficulty. The challenge in enabling students to tackle such problems is considerable, so instructional support at this grade level is particularly appropriate.

What is difficult about solving non-routine word problems?

Non-routine word problems cannot be solved using any fixed algorithmic method or set of prescribed procedures (Elia, van den Heuvel-Panhuizen, & Kovolou, 2009; Pantziara, Gagatsis, & Elia, 2009). Rather, word problem solving depends on two major phases, namely problem comprehension and problem solution. *Problem comprehension* involves: (1) understanding the problem text (i.e., what question is to be answered); (2) identifying the relevant numerical and linguistic components of the problem (e.g., key words as 'more than', 'less than'); (3) identifying the spatial relations between these components; and (4) representing the components and the spatial relations between them (i.e., the problem structure) in a complete and coherent way. *Problem solution* involves determining the mathematical operations (i.e., addition, subtraction, multiplication and/or division) to be applied to the identified numerical components and executing these computations to solve the problem (Krawec, 2010; Lewis & Mayer, 1987). Solving non-routine word problems thus involves carrying out and integrating several cognitive activities that involve a non-trivial amount of related information.

Research shows that the difficulties experienced by many students in solving word problems arise not from their inability to execute computations, but from difficulties in understanding the problem text, identifying solution-relevant components and the relations between them, and representing the problem structure (Boonen et al., 2013; Carpenter, Corbitt, Kepner, Lindquist, & Reys, 1981; Cummins, Kintsch, Reusser & Weimer, 1988; Krawec, 2010; Lewis & Mayer, 1987). Hence, erroneous word problem solutions are frequently a consequence of errors in the problem comprehension phase rather than in the problem solution phase.

Previous research into supporting word problem solving

From the perspective presented above, providing support for the problem comprehension phase should be particularly beneficial to students' problem solving performance. Yet, in educational practice, word problem solving in the classroom as well as teacher training often focuses on the solution phase. Existing research-based programs developed to support word problem solving of lowperforming students incorporate both problem solving phases in the form of prescribed cognitive strategies (i.e., reading and understanding the problem, analyzing the information presented, developing logical solution plans, evaluating solutions) presented in a sequence of steps (Jitendra & Star, 2012; Jitendra et al., 2009; Krawec, 2012; Montague et al., 2000).

In these programs, support for the problem comprehension phase typically includes strategies for understanding the problem text (e.g., by paraphrasing the text and underlying relevant information) and strategies for identifying and representing the underlying problem structure by means of a *visual representation.* The assumption is that a visual representation should clarify the problem structure by making the numerical, linguistic and spatial relations between solution-relevant elements visible, which consequently facilitates understanding of the problem and identification of the computations to be performed (Boonen et al., 2014; Krawec, 2010, 2012). Thus, using visual representations during problem comprehension could be an effective way to support word problem solving (Van Garderen & Montague, 2003).

Two current research-based approaches to using visual representations in the problem comprehension phase can be distinguished. The first is to provide students with specific visual representations for specific types of problem, namely *routine* word problems. Students are then encouraged to reflect on the similarities and differences between problem types and the corresponding visual representations (Jitendra, Dipipi, & Perron-Jones, 2002; Jitendra & Star, 2012; Jitendra et al., 2009). Although this approach can be successful for teaching students how to tackle routine problems with an identical structure (e.g., 'compare' problems as: *Mary has 5 marbles. John has 8 marbles. How many more marbles does John have than Mary?)*, it is usually not possible to match non-routine word problems to a fixed representation type. Teaching students to use only one specific visual representation for each type of problem is, moreover, risky, as it may lead to an inflexible and rigid use of representation strategies (Jitendra, Griffin, Haria, Leh, Adams, & Kaduvettoor, 2007; Van Dijk, Van Oers, & Terwel, 2003; Van Dijk, Van Oers, Terwel, & Van den Eeden, 2003a).

The second approach – typically for *non-routine* word problems – defines visual representation in general heuristic terms: this should be done either mentally or by making a drawing. Indeed, it is characteristic for non-routine problems that they cannot be represented in a prescribed way; in such conditions, the use of a heuristic approach may seem appropriate. It is a false assumption, however, that students know how to translate such a heuristic into a useful visual representation. Students generally do not know what to draw, when, under which circumstances and for which types of problems (Jitendra et al., 2007; Jitendra et al., 2009).

In summary, there are important shortcomings in current approaches to using visual representations to support comprehension of non-routine word problems. On the one hand, approaches that teach fixed visual representations for problem solving are insufficient to deal with the situation-specific structure of non-routine problems. On the other hand, heuristic approaches do not provide sufficient guidance for students to produce useful visual representations.

Visual representations for comprehension of non-routine word problems

These shortcomings could be addressed by teaching students to produce visual representations that accurately depict the situation-specific structure of non-routine word problems. Such representations should present a complete and coherent model of the relations between all solution-relevant problem components. We refer to these further as *accurate visual-schematic representations*. It is important to note that such representations can incorporate standard mathematical models, such as a bar model, pie chart or number line, but that they often comprise freely constructed drawings. Examples of an accurate bar model and 'own' construction are shown in Table 1(a) and (b) respectively for the weighing scale problem presented earlier.

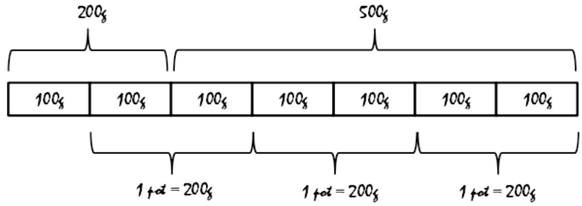
Research shows that accurate visual-schematic representations facilitate problem comprehension, help to identify the required calculation processes and thereby contribute to successful problem solving (see e.g., Boonen et al., 2013, 2014; Hegarty & Kozhevnikov, 1999; Van Garderen, 2006; Van Garderen & Montague, 2003). For example, the visual-schematic bar model in Table 1(a) demonstrates that an accurate depiction of problem structure can greatly reduce calculation demands: simply by depicting the given quantities in the correct relation to each other, it can instantly be seen that each pot of jam must weigh 100g + 100g.

This stands in contrast to more commonly known visual representations, namely pictorial and arithmetical representations, both of which frequently accompany word problem solving in mathematical text books. *Pictorial* representations contain a detailed image of some element of the problem text (e.g., an object or a person) without identifying relations between problem elements or the required calculations (Table 1[e]). These visual representations have been found to negatively influence the word problem solving process (Boonen et al., 2014; Hegarty & Kozhevnikov, 1999; Krawec, 2010, 2012; Van Garderen, 2006; Van Garderen & Montague, 2003). *Arithmetical* representations (e.g., proportion tables, Table 1[f]) are intended to support the calculation processes required to compute answers. This type of representation is generally introduced in the problem solution phase but does not contribute to problem comprehension; thus, when the problem is not well understood, arithmetical representations frequently contain erroneous information and lead to incorrect answers (Boonen et al., 2013; Cummins et al., 1988; Krawec, 2010).

Table 1. *Examples of types of visual representation for the weighing scale problem*

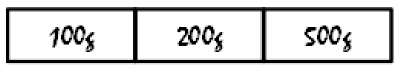
Representation type — **Example**

(a) Accurate visual-schematic: bar model



(b) Accurate visual-schematic: own construction

(c) Inaccurate visual-schematic: bar chart



(d) Inaccurate visual-schematic: own construction

(e) Pictorial

(f) Arithmetical: Proportion table

No. pots	3	1
Weight (g)	600	200

Unfortunately, very little is known about how teachers could teach students to construct accurate visual-schematic representations to support word problem comprehension. It cannot be assumed that teachers are able to do this. For a start, teachers may not know what kind of visual representations should be made or in which phase of the problem solving process to use them. Teachers may also have difficulty in constructing these types of visual representations accurately (i.e., correctly and completely). Incorrect and/or incomplete visual-schematic representations are referred to as *inaccurate* visual-schematic representations (e.g., Table 1[c] and [d]).

Furthermore, research shows that it is more effective to teach students to construct their own visual representations than to provide them ready-made, as this contributes to skill adaptivity (Van Dijk et al., 2003; Van Dijk et al., 2003a). Thus, teaching needs to focus on the construction process (i.e., how to make the representation), rather offering a representation as a given entity. Furthermore, teachers should encourage students to use visual representations in a diverse, flexible and functional way. This refers to being able to use different kinds of visual representations and to switch between them such that the representation fits the structural characteristics of the problem and is useful for helping to solve it.

In short, although using accurate visual-schematic representations to support the problem comprehension phase of word problem solving has considerable potential for improving problem solving performance, research is needed that examines how teachers implement an approach centered on the use of these visual representations. It is important to establish this point, as it is critical to the viability of this approach for supporting word problem solving in schools.

The present study

The goal of the present study is to examine *teachers' use of visual representations* when implementing a teaching intervention for supporting non-routine word problem solving that focuses on constructing accurate visual-schematic representations. This teaching intervention is embedded within a sequence of problem solving steps (comparable to the programs mentioned above) that reflect the problem solving phases described earlier. It is important to note that, just as both problem solving phases are essential for effective problem solving, so all steps are intended to be carried out fully in the prescribed order for each problem treated.

The general goal of the teaching intervention is to teach students cognitive strategies for solving non-routine word problems. The specific goal is to teach students to construct *accurate visual-schematic representations* of problem structure and to encourage them to select and use visual representations in a functional (i.e., useful for helping to solve the problem) and adaptive way. This refers to diverse and flexible use of visual representations, by which the kinds of visual representations used are varied to suit problem characteristics.
An overview of the steps and the relation of each step to problem solving phase is provided in Box 1 (based on Montague et al., 2000). The key interest of the present study, that is the construction of accurate visual-schematic representations (implemented in the third, i.e., visualization step), is indicated in bold print.

The teaching intervention comprises eight lessons that make use of teacher modelling (i.e., thinking aloud while demonstrating a cognitive activity), student modelling and independent student practice (see Methods section). We focus on *teacher modelling* of visual representations, where teachers' use of visual representations is expected to be most visible. An important aspect of the teaching intervention is that teachers are encouraged to implement it in a way that is compatible with their own manner of teaching (Rogers, 2003). This makes it possible to study natural diversity in teachers' behaviours.

Research questions

To meet the goals of the present study, the following research questions are posed:

1. *What attention do teachers give to visualization when implementing the teaching intervention, when do they use visual representations in the word problem solving process and to what purpose?* Given the focus of the teaching intervention on the use of visual representations, teachers are expected to pay most attention to the visualization step of the problem solving process, to use visual representations to structure problem elements and the relations between them, and to embed this within the full sequence of prescribed steps. However, it is possible that teachers focus on other parts of the problem solving process, that they use visual representations at other points of the process and/or for other purposes (e.g., illustrating unfamiliar words in the text, calculating answers), or that they do not use visual representations at all.

Box 1: Word problem solving steps		
Problem solving phase	***Step***	***Content***
	Step 1.	
PROBLEM COMPREHENSION: understanding text	READ the problem carefully all the way through	Each sentence of the text is studied for comprehension and not just to search for numbers and key words (such as more than, times, as much as, etc.).
	Step 2.	
PROBLEM COMPREHENSION: understanding text meaning, identifying problem components, identifying relations	UNDERSTAND the text: put it in your own words imagine the situation underline important information. what is being asked?	Deep understanding of the text is stimulated by a sequence of four substeps. The text is paraphrased (i.e.,put into own words), the described situation is imagined mentally, solution-relevant information needed for solving the problem is underlined, and the problem solver asks him/herself what question is to be answered.
	Step 3.	
PROBLEM COMPREHENSION: representing problem structure	**VISUALIZE the problem structure: make a drawing of the problem situation**	**An accurate visual-schematic representation of the text is made. This contains correct and complete depictions of spatial, linguistic and numeric relations between all solution-relevant elements of the text.**
	Step 4.	
PROBLEM SOLUTION: determining operations	HYPOTHESISE a plan to solve the problem: what kind of problem is it? (+, -, x, :) what do you need to calculate?	The number of solution steps and type(s) of operation (addition, subtraction, multiplication and/or division) are derived from the schematic representation. The required calculation is written down in standard symbolic notation.
	Step 5.	
PROBLEM SOLUTION: executing computations	COMPUTE the required operation	The required calculations are performed.
	Step 6.	
	CHECK your answer	The computation is checked and it is considered whether the result is a plausible answer to the question asked.

7

It is also possible that steps are omitted, combined or performed out of order, which may result in the visualization step not being embedded in the prescribed sequence as intended.

2. *What kinds of visual representations do teachers use and how adaptive (i.e., diverse and flexible) is this use of visual representations?* Teachers are expected to demonstrate a varied use of visual representations (i.e., diversity) and to offer different kinds of representations for solving a problem (i.e., flexibility). However, it is possible that teachers use representations in a limited and fixed way and/or that they do not consider different ways of representing problems.
3. *What is the quality of the representation process and of the visual representations used?* Teachers are expected to model the representation process transparently, correctly and completely. However, it is possible that teachers do not make their reasoning transparent (e.g., offering a visual representation without explaining which elements of the problem should be represented or without explaining how the representation can be used to solve the problem), and/or that their reasoning is incorrect (e.g., naming and/or using visual representations wrongly) and/or incomplete (e.g., naming a visual representation that can be used but not indicating why). Teachers are also expected to construct visual representations that correctly and completely depict the relations between all solution-relevant problem components. However, the visual representations made could be incorrect (i.e., containing erroneous problem components or relations) and/or incomplete (i.e., missing components and/or relations). The visual representations used are also expected to be suitable and useful (i.e., functional) for solving the problem. However, it is possible that visual representations do not fit problem characteristics well (e.g., a number line for solving a problem about percentages) and/or that they contain excess information that is not relevant for and could interfere with solving the problem.

Study approach and relevance

The present research is performed within the context of a teaching intervention for supporting non-routine word problem solving that involved eight mainstream teachers recruited through purposive sampling. These teachers had a positive attitude towards mathematics, were confident about teaching mathematics and using visual representations, and were motivated to participate in and contribute to research in this area. It is well established that a lack of motivation and interest can severely impact the way in which teachers implement educational innovations in regular classrooms. The personal willingness of teachers to adopt and integrate innovations into their classroom practice is of crucial importance (Evers, Brouwers, & Tomic, 2002; Ghaith & Yaghi, 1997; Hermans, Tondeur, Van Braak, & Valcke, 2008; Rogers, 2003). Thus, by limiting participation to individuals with these qualities, results are obtained under favorable conditions in which teacher behavior is not negatively influenced by motivational factors. This allows behaviors to be analyzed on the basis of the specified criteria, with known affective confounders excluded.

The study makes an unique contribution to research in the important and problematic area of word problem solving in regular classrooms. As far as we know, it is the first study to focus on mainstream teachers' use of visual-schematic representations in wholeclass word problem solving instruction. By focusing on how teachers perform the intended behaviours, the study provides indications for improving the design and development of instructional support and teacher training in this area.

METHODS

Participants

Directors of elementary schools in the central provinces of the Netherlands were directly approached with information about the teaching intervention and a request to participate in the present research. Eight mainstream sixth grade teachers from four elementary schools subsequently volunteered to participate. These teachers indicated that they were motivated to implement the teaching intervention and contribute to this research, that they had a positive attitude towards mathematics, that they were confident about teaching mathematics and that they believed themselves to be competent in using visual representations in the math lesson. Table 2 presents the background characteristics of the participating teachers.

Parents of students in the classes of the participating teachers were informed that their child would participate in the study and that they could withhold permission for their child to participate. No parents took up this option; thus, all participating classes were intact.

Table 2. *Background characteristics of participating teachers*

	Sex	Highest qualification level[a]	Number of years teaching	Days per week teaching	School type[b]
Teacher 1	F	1	20	5	2
Teacher 2	F	1	6	5	2
Teacher 3	M	1	13	4	2
Teacher 4	M	1	27	4	2
Teacher 5	F	2	14	5	1
Teacher 6	M	1	13	4	2
Teacher 7	M	1	39	5	2
Teacher 8	F	1	6	3	2

Note. [a] *1 = Bachelor of Applied Sciences (teacher education for primary schools) 2 = Master of Applied Sciences (teacher education for primary schools); b 1 = Urban 2 = Provincial*

Context of the research: the teaching intervention

Three weeks prior to commencement of the teaching intervention, teachers attended one afternoon session that presented the rationale for and the aims of the study and introduced the teaching intervention. A second afternoon session two weeks later focused on how to execute the problem solving steps. Professional development materials supplied to teachers contained: (a) an overview of the most important national and international literature regarding word problem solving and word problem solving instruction; (b) a presentation and explanation of each of the problem solving steps; (c) a description of how to perform each step on the basis of several examples.

Furthermore, teachers received an elaborated lesson protocol that contained a fully scripted model of how to apply the steps for each word problem to be treated. These problems were obtained from regular math textbooks and were therefore fully authentic. Rather than use the scripts verbatim, however, teachers were encouraged to use their own explanations and elaborations during the teaching intervention: as stated earlier, teachers were encouraged to implement the teaching intervention in a way that was compatible with their own teaching approach.

The eight lessons of the teaching intervention were subsequently delivered by all participating teachers over the course of four weeks, with two lessons of 40 minutes duration per week. These lessons replaced regular lessons within the standard math curriculum. On-going assistance from the research team was available throughout the study duration. Each student in the class of a participating teacher received a textbook with the problems to be treated, an exercise book and a prompt card that depicted the problem solving steps in a visually attractive way.

The first four lessons made use of teacher modelling followed by independent student practice. In these lessons, three to six word problems were intended to be modelled by the teacher, with 22 problems to be modelled in total. The last four lessons made use of student modelling followed by independent student practice. Instructional support was gradually faded out within and across lessons so that students could ultimately take control of their own problem solving work. Twenty of the teachermodelled problems were selected for analysis[5]. The two problems that were excluded were not representative of the types of non-routine problems with which students have difficulties as they required only one solution step and could be solved using a fixed algorithmic method.

The present study: Measures and analyses

The complete teaching intervention (i.e., 8 lessons) was recorded on video for each of the eight participating teachers. All recordings of the first four lessons (i.e., in which teacher modelling took place) were subsequently viewed by two members of the research team, and the way in which the teacher concerned modelled the problem solving steps (see Box 1) for each of the selected problems was recorded, described and coded according to the measurement and analysis scheme presented in the following paragraphs. Note that pictorial representations were recorded but not analysed as they could not be judged on criteria relevant to this study. Inter-rater reliability of the coding was computed as Krippendorff's alpha coefficient for nominal (here, dichotomous) and ordinal data. The obtained coefficients for nominal variables ranged from .86 to 1.00 and for ordinal variables from .91 to .98, indicating high reliability (Krippendorff, 2004).

[5] Note that not all teachers modelled all of these problems due to time constraints.

First research question (attention to visual representations, usage and purpose)

For each teacher-modelled problem and for each problem solving step, it was recorded: (1) whether and when the teacher performs the step; (2) step duration in minutes; (3) whether the teacher uses a visual representation during the step; (4) for what purpose the visual representation is used; (5) any other factors (e.g., lesson or class management) impacting teachers' attention to or use of visualization. Where steps were combined (i.e., performed simultaneously as opposed to sequentially), the time spent on the combination was allocated in equal proportions to each of the constituent steps. The total time spent on each step by each teacher across all teacher-modelled problems was then calculated, giving an indication of the attention given to each step and consequently the relative attention given to visualization compared to the other steps. Next, the data from (1), (3), (4) and (5) above were reviewed and discussed by two members of the research team. This resulted in the identification of patterns specifying when visual representations are used during the problem solving process and to what purpose.

Second research question (kinds of visual representations used and adaptivity)

Teachers' behaviors relevant to their use of different types and forms of representations were recorded. For each visual representation used, the *type* of representation (i.e., pictorial, arithmetical, visual-schematic) and *form* of representation (i.e., bar model, pie chart, number line, proportion table, own construction) was recorded; more than one representation could be recorded per problem, if applicable. For each teacher, the following measures were then derived from these data:

- Diversity: To examine the extent to which teachers demonstrated a varied use of visual representations Simpson's D was calculated (Simpson, 1949). Simpson's D is an estimation of the effective number of species, i.e., the effective number of representation types used by a teacher. In this case, D ranges from 1 (only one representational form used) to 5 (all representation types used evenly). Note that pictorial representations were not included in calculating D.
- Flexibility: The extent to which teachers demonstrated flexible use of visual representations by offering different representations to solve a problem was calculated as the number and percentage

of problems modelled for which the teacher used more than one form of arithmetical and/or visual-schematic representation (e.g., bar model plus proportion table).

Third research question (quality of representation processes and representations)

Teachers' behaviors relevant to the quality of their representation processes and representations were recorded. For each arithmetical and visual-schematic representation used, the quality of the representation process was rated in terms of:

- Transparency, i.e., the extent to which the teacher explicitly explains (or asks students to explain) step-by-step reasoning about what information is essential to represent for solving the given problem and how to represent this (1 = reasoning implicit and unexplained, 2 = reasoning partially explained, 3 = reasoning fully explained);
- Correctness, i.e., whether or not the reasoning underlying the representation process is correct;
- Completeness, i.e., whether or not the reasoning underlying the representation process is complete.

For each arithmetical and visual-schematic representation used, the quality of the representation itself was rated in terms of:

- Functionality, i.e., the extent to which the representation is suitable and useful for solving the given problem (1 = not suitable/ useful, 2 = suitable/useful but includes solution-irrelevant information, 3 = suitable/useful and includes only solution-relevant information);
- Correctness, i.e., whether or not all solution-relevant elements are correctly represented;
- Completeness, i.e., whether or not all solution-relevant elements are represented and not missing.

On the basis of these data, the average transparency rating, percent correct and complete representation processes, average functionality rating and percent correct and complete representations were then calculated for each teacher across all arithmetical and visual-schematic representations used by the teacher in question.

RESULTS

What attention do teachers give to visualization, when do they use visual representations in the word problem solving process and to what purpose?

It should first be noted that, although teachers were aware that they should perform each of the problem solving steps in the prescribed order for each problem, none of them consistently followed the full sequence of steps. All teachers occasionally omitted steps completely (often the HYPOTHESISE and CHECK steps) or partially (particularly paraphrasing the text and imagining the specific situation described therein). Furthermore, teachers frequently concatenated steps, often combining VISUALIZE with the UNDERSTAND and/or COMPUTE step. The consequences of this for teachers' use of visual representations are examined later.

Table 3 presents the time allocation for each teacher and each step across all teacher-modelled problems. Note that teachers did not spend equivalent amounts of time on modelling (range = 41-111 minutes) and did not model the same number of problems (range 4-15[6]). All but two teachers frequently solicited student input during the modelling process and engaged in extensive interaction with students, particularly during the UNDERSTAND, VISUALIZE and COMPUTE steps. Consequently, most time was spent on these steps.

It can be seen that four of the eight teachers spent the most time on understanding the problem text (UNDERSTAND), two on both visualizing the problem structure (VISUALIZE) and calculating answers (COMPUTE), and two on calculating answers (COMPUTE). Thus, although all teachers were aware of the key role of visualization, most spent more time on other parts of the problem solving process. For example, two teachers extensively discussed the general context of each problem (e.g., what kinds of things one can buy in a supermarket) and six teachers placed considerable emphasis on calculating answers.

[6] Excluding the two problems that were not selected for analysis.

Table 3. *Teachers' time allocation (minutes) per step over all teacher-modelled problems*

Step	Teacher 1	Teacher 2	Teacher 3	Teacher 4	Teacher 5	Teacher 6	Teacher 7	Teacher 8	Total
READ	5.5	8.1	6.1	6.7	5.8	1.4	2.9	6.3	42.8
UNDERSTAND	**20.6**	**48.0**	24.8	8.4	**24.5**	6.0	**17.8**	13.1	163.2
VISUALIZE	8.6	15.7	15.8	**16.4**	21.0	11.3	13.9	**20.0**	122.7
HYPOTHESISE	6.7	11.0	7.2	1.9	12.7	5.2	3.2	1.7	49.6
COMPUTE	10.6	23.4	**27.3**	**16.5**	22.0	**14.7**	4.7	**20.0**	139.2
CHECK	1.7	4.8	4.7	3.9	5.9	2.5	-	5.0	28.5
Total	*53.7*	*111.0*	*85.9*	*53.8*	*91.9*	*41.1*	*42.5*	*66.1*	*546*
Number of problems modelled	11	15	14	10	9	7	4	9	

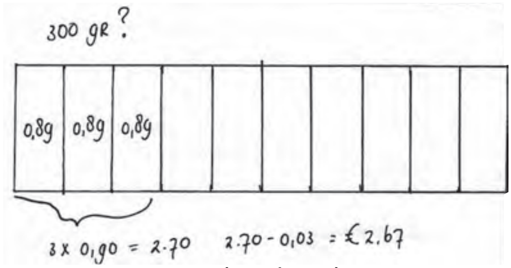
Regarding when visual representations are used in the problem solving process and to what purpose, four patterns of usage were identified: (1) UNDERSTAND & VISUALIZE; (2) VISUALIZE; (3) VISUALIZE & COMPUTE; (4) COMPUTE. The upper part of Table 4 shows the amount of time spent on visualization following each of these patterns. The lower part of the table shows the corresponding types of visual representations (i.e., pictorial, arithmetical, visual-schematic) used. Individual differences between teachers are apparent.

The first pattern was observed when the problem text contained a lot of information or when teachers did not fully comprehend what was being asked. Visual representations (specifically, visual-schematic) were then occasionally used to help clarify the situation described in the text; the steps UNDERSTAND and VISUALIZE were then combined. Half of the teachers showed this pattern of behavior.

The second pattern observed (i.e., VISUALIZE) occurred once the problem text was fully comprehended. Visual representations (pictorial, arithmetical and/or visual-schematic) were then sometimes used to separately depict the problem structure; these representations were not further used in calculating answers. All but one teacher demonstrated this pattern of usage.

With the third pattern, visual representations were used to both structure the problem and help calculate answers; the steps VISUALIZE and COMPUTE were then combined. All but one teacher used visual representations in this way. For four teachers, this was the main way in which they used visual representations and one teacher used them only in this way. Two distinct approaches were identified. With the first approach, a visual-schematic representation was used for structuring the problem and calculating answers. An example is shown in Figure 1. All teachers who showed the VISUALIZE & COMPUTE usage pattern worked in this way on at least one problem. The second approach - used by two teachers - used arithmetical representations (specifically, proportion tables) for this purpose.

This week one kilogram of cheese costs € 8.90. How much does a piece weighing 300 grams cost?



(Teacher 4)

Figure 1. Example visual-schematic representation used to calculate answers

Finally, with the fourth pattern (i.e., COMPUTE), visual representations (specifically arithmetical) were used only to calculate answers. This often followed the situation in which a pictorial or visual-schematic representation was used in the VISUALIZE step, or followed directly from the text comprehension step (UNDERSTAND) in cases where the VISUALIZE step was omitted. Six teachers used visual representations in this way.

Table 4. *Patterns of visual representation usage over all teacher-modelled problems*

Step	**Rep type**	**Teacher 1**	**Teacher 2**	**Teacher 3**	**Teacher 4**	**Teacher 5**	**Teacher 6**	**Teacher 7**	**Teacher 8**
Time allocation per usage pattern (minutes)									
UNDERSTAND & VISUALIZE		-	3.4	-	4.8	5.9	1.7	-	-
VISUALIZE		6.1	7.3	4.4	3.1	12.6	6.6	13.9	-
VISUALIZE & COMPUTE		5.1	13.4	22.7	21.7	11.0	13.0	-	40.0
COMPUTE		8.1	16.7	16.0	5.7	16.5	8.2	4.7	-
Number of visual representations per usage pattern									
UNDERSTAND & VISUALIZE	Visual-schematic	-	1	-	1	1	1	-	-
VISUALIZE	Pictorial	-	5	-	-	1	-	1	-
	Arithmetical	-	-	-	-	1	-	-	-
	Visual-schematic	4	8	5	4	8	2	3	-
VISUALIZE & COMPUTE	Arithmetical	-	-	2	-	-	-	-	7
	Visual-schematic	3	3	8	13		2	-	3
COMPUTE	Arithmetical	2	2	5	1	1	2	-	-

Note. Visual representations were never used when reading the problem text (READ), hypothesising the required calculations (HYPOTHESISE) or checking the answer (CHECK)

What kinds of visual representations do teachers use and how adaptive (i.e., diverse and flexible) is this use of visual representations?

There were considerable differences between teachers with respect to the *type* and *form* of visual representations used. Regarding *types*, (i.e., pictorial, arithmetical, visual-schematic) all but one teacher used mainly visual-schematic representations. Though teachers were aware of the importance of these visual representations, pictorial and arithmetical representations were also used, however (see Table 5).

Table 5. *Teachers' use of representation types over all teacher-modelled problems*

Rep type	Teacher 1	Teacher 2	Teacher 3	Teacher 4	Teacher 5	Teacher 6	Teacher 7	Teacher 8
Number of problems	11	15	14	10	9	7	4	9
Number of reps	9	19	20	19	14	9	4	10
Pictorial[a]	0 (0%)	5 (26%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	1 (25%)	0 (0%)
Arithmetical[a]	2 (22%)	2 (11%)	7 (35%)	1 (5%)	2 (14%)	2 (22%)	0 (0%)	7 (70%)
Visual-schematic[a]	7 (78%)	12 (63%)	13 (65%)	18 (95%)	11 (79%)	7 (78%)	3 (75%)	3 (30%)
Diversity	1.59	2.80	3.17	1.24	4.12	3.52	1.80	1.85
Flexibility[b]	1 (9%)	1 (7%)	2 (14%)	1 (10%)	3 (33%)	1 (14%)	1 (25%)	1 (11%)

Note. [a]*Percentages are of the number of visual representations used per teacher;* [b] *Percentages are of the number of problems modelled by the teacher*

Figure 2 shows teachers' relative use of different representational *forms* (i.e., bar model, pie chart, number line, proportion table, own construction, pictorial). Note that though pictorial representations are considered a *type*, they are included here to give a complete picture of the visual representations used. The most frequent form of visual-schematic representation was the bar model, which was used by all but one teacher. Number lines were used by four teachers, while pie charts were very infrequently used, by four teachers. The only form of arithmetical representation used was proportion tables. This was used by all but one teacher.

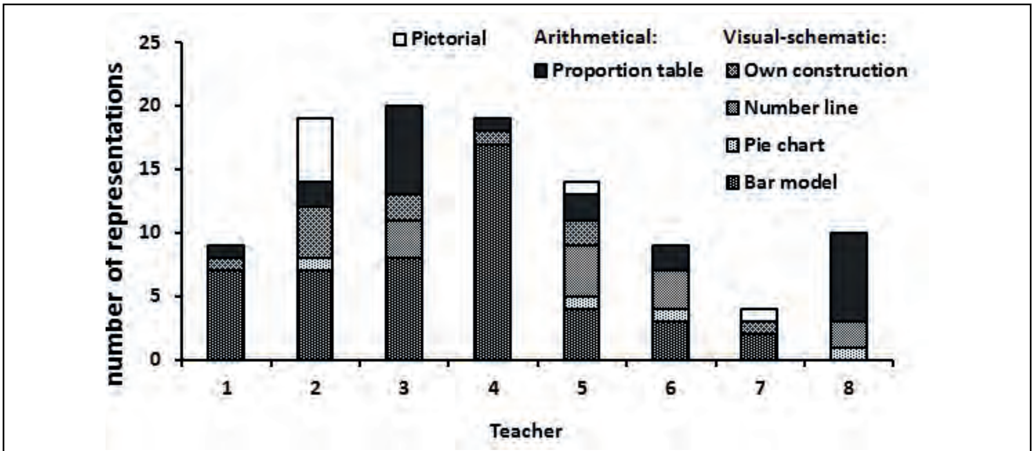


Figure 2. Teachers' use of representational forms over all teacher-modelled problems

The extent to which teachers exhibited a varied use of representational forms (i.e., *diversity*) differed between teachers, varying from 1.24 to 4.12 on a scale of 1 (only one representational form used) to 5 (all representational forms used equally) (see Table 5). Excluding pictorial representations, one teacher used five representational forms including own constructions, two teachers used four forms, four teachers used three forms and one teacher used two forms. Some teachers verbally expressed and demonstrated a preference for a certain representational form. Two teachers showed a strong preference for bar models (although one of these systematically used them incorrectly, as will be discussed later) and two teachers had a strong preference for proportion tables. Quotations that reflect teachers' representation preferences are reported in Box 2.

Box 2: Teachers' representation preferences (quotations)

Teacher 3:	"You all know that I am a big fan of proportion tables"; "I always use a proportion table. You start somewhere and end up with a proportion table"
Teacher 4:	"I can draw another bar model here"; "I feel another bar model coming on!"
Teacher 8:	"A proportion table is great to use for nearly all sorts of problems"

Teachers demonstrated a low to medium degree of *flexibility* in representation use. All offered multiple representational forms (arithmetical and visual-schematic) on the same problem at least once (see Table 5). Five teachers used different forms of visual-schematic representation in this way. Four combined a bar model with their own schematic drawing of the problem structure; two offered different visual-schematic mathematical models (number line, bar model and/or pie chart). Five teachers used both visual-schematic and arithmetical representations on the same problem at least once. Remarkably, three teachers modelled at least one problem without using any visual representation at all. Teachers rarely compared the use of different representational forms or reflected critically on what different kinds of representation contribute to word problem solving. This was particularly the case for teachers who had an expressed preference for using a specific form, who showed a tendency to use their preferred form irrespective of problem characteristics.

Quality of representation processes and visual representations

Teachers were expected to model the representation process transparently, correctly and completely, whereby they should emphasize reasoning processes and understanding the actions undertaken. Thus, teachers were expected to explicitly explain which problem elements should be modelled and how, which representational forms are suitable to do this and why, and which actions they take in constructing the visual representations step-by-step. All eight teachers showed these kinds of behaviors to a considerable degree, with average *transparency* ratings of 2.11-3.00 on a three-point scale (see Table 6). The extent to which teachers provided this reasoning themselves or stimulated students to provide it varied between teachers, however. Two teachers also encouraged students to suggest which visual representations to use.

Table 6. *Quality of representation processes and visual representations*

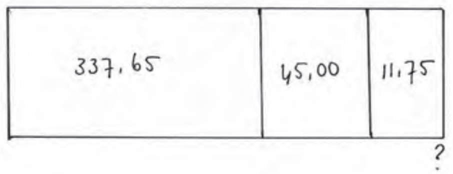
	Teacher 1	Teacher 2	Teacher 3	Teacher 4	Teacher 5	Teacher 6	Teacher 7	Teacher 8
Number of reps	9	19	20	19	14	9	4	10
				Representation processes (reasoning)				
Transparency[a]	2.11	2.43	2.70	2.84	2.62	2.89	3.00	2.90
Correctness[b]	33%	100%	100%	95%	100%	100%	100%	100%
Completeness[b]	100%	86%	90%	95%	85%	100%	100%	100%
				Visual representations				
Functionality[a]	1.56	2.64	2.85	2.63	2.69	3.00	3.00	3.00
Correctness[b]	33%	100%	100%	89%	92%	89%	100%	100%
Completeness[b]	100%	79%	85%	89%	77%	100%	100%	100%
				Accuracy of visual representation types				
Arithmetical[c]	0%	0%	100%	100%	50%	100%	-	100%
Visual-schematic[c]	29%	92%	77%	83%	73%	86%	100%	100%

Notes. [a]*Averaged over all arithmetical and visual-schematic representations;*
[b]*Percentage of all arithmetical and visual-schematic representations;*
[c]*Percentage correct and complete visual representations of this type.*

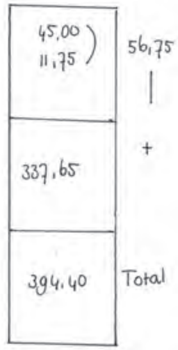
Correctness of reasoning in the representation process was high (95% and above) for seven of the eight teachers. Remarkably, one teacher (Teacher 1) reasoned correctly on only 33% of visual representations. The main source of error appeared to be an incorrect understanding of what a bar model is: she drew boxes that she called bar models within which she wrote the sum in symbolic notation. This teacher along with three others also used a bar model form as a proportion table, with the numbers written in the bars bearing no relation to the spatial-numerical relations. Another teacher (Teacher 7) expressed the belief that it is not possible to make a visualization for every type of problem, though his reasoning was correct for the representations that he did make. *Completeness of reasoning* was also high (95% and above) for five teachers, while the remaining three teachers demonstrated complete reasoning on 85-90% of the visual representations used.

There were also differences between teachers concerning the quality of the visual representations used. All teachers produced at least one accurate (i.e., *correct* and *complete*) visual-schematic representation, though accuracy varied from 29% to 100% for this representation type (see last row of Table 6). Only two teachers were fully accurate on all visual-schematic representations, while the teacher who demonstrated low correctness of reasoning produced accurate visual-schematic representations only 29% of the time. Four teachers constructed visual-schematic representations that were incorrect (i.e., depicting erroneous problem elements or relations) and four teachers constructed visual-schematic representation that were incomplete (i.e., missing solution-relevant elements and/or relations). Examples of accurate and inaccurate visual-schematic representations are given in Figure 3. The arithmetical representations of four teachers were all accurate (i.e., correct and complete), but three teachers produced incorrect or incomplete arithmetical representations (one teacher never used these visual representations, see Table 5).

Yesterday, Fenna had € 337.65 in her bank account. Her new statement shows that her grandmother deposited € 45 for her school report and she earned € 11.75 for babysitting. How much money does she have in her account now?



(Teacher 5)



(Teacher 1)

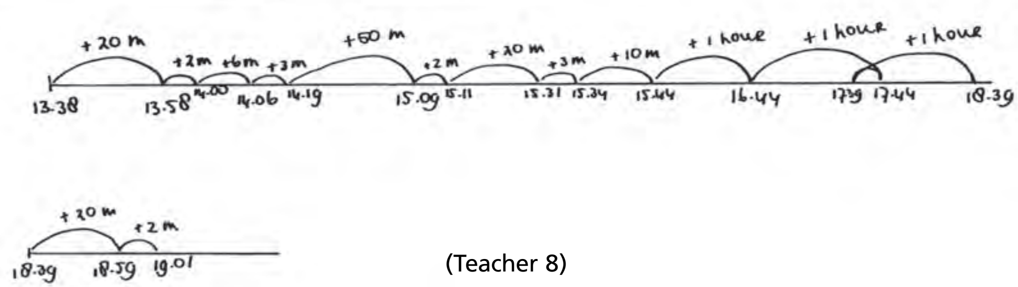
(a) Accurate (correct and complete)

(b) Inaccurate (incorrect)

Figure 3. Examples of (a) accurate and (b) inaccurate visual-schematic representations.

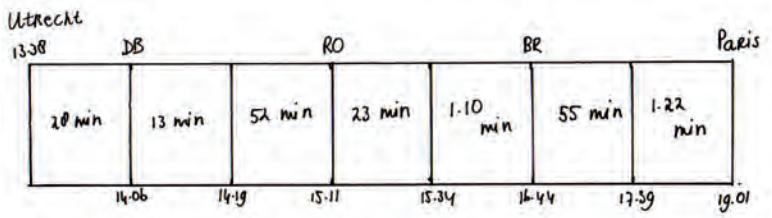
The extent to which teachers invested in understanding what was being asked in the problem text - particularly when the text contained much information - directly impacted the *functionality* (i.e., suitability and usefulness) of the visual representations used. With one exception, teachers' visual representations generally had good functionality, with average ratings of 2.63-3.00 on a three-point scale (see Table 6). However, the visual representations of the teacher who showed low correctness of reasoning had lower functionality (average rating 1.56), which could be expected. Functionality was also lower when teachers did not correctly identify the solution-relevant elements; in these cases, they tended to produce visual representations that contained excess, irrelevant information. Figure 4(a) shows an example of such a visual-schematic representation, with unnecessary calculation of intermediate arrival and departure times. Furthermore, teachers used bar models more frequently than other visual-schematic representations, even on problems involving percentages that were suitable for a pie chart, or addition and subtraction problems that were suitable for using a number line. Figure 4(b) shows an example of a bar model used (inaccurately) when a number line would have been more suitable.

Antoine and Bertrand travel with the train from Utrecht to Paris via Den Bosch, Roosendaal, and Brussels. They depart at 13.38 from Utrecht Central Station. After 28 minutes they arrive in Den Bosch. They have to wait 13 minutes before the train to Roosendaal leaves. The journey to Roosendaal takes 52 minutes. In Roosendaal they have 23 minutes to change to the train to Brussels. The journey to Brussels takes 1 hour and 10 minutes. After 55 minutes, the train departs from Brussels to Paris: a journey of 1 hour and 22 minutes. At what time do Antoine and Bertrand arrive in Paris?



(Teacher 8)

(a)



(Teacher 4)

(b)

Figure 4. Examples of (a) visual-schematic representation containing excess information and (b) bar model used to solve addition problem

DISCUSSION

The goal of this study was to examine teachers' use of visual representations when implementing a teaching intervention for supporting non-routine word problem solving. The teaching intervention focused on the construction of accurate visual-schematic representations, embedded within a sequence of six problem solving steps. It differs from previous research into supporting word problem solving (e.g., Jitendra, 2002; Montague, 2003; Montague et al., 2000) in several respects relating to both the educational setting and the way in which it is implemented. First, the teaching intervention is not implemented with low-performing students with special educational needs, but in regular, mainstream classrooms. Second, the intervention is not designed for individual or small group instruction, but for use in whole-class teaching. Third, the intervention is not carried out by researchers but by mainstream teachers.

Furthermore, while the intervention is based on existing stepwise strategy instruction programs to support non-routine word problem solving, it incorporates an important innovation. Where existing programs define visual representation in heuristic terms, the present teaching intervention specifically defines the criteria that should be satisfied: a visual representation should clarify the problem structure by making the numerical, linguistic and spatial relations between solution-relevant elements visible. The study therefore makes an unique contribution to research in the important and problematic area of word problem solving in regular classrooms.

To address the study objective, three research questions were posed. We first examined teachers' attention to visualization when implementing the teaching intervention and determined when they use visual representations in the word problem solving process and to what purpose. Answering this question provides insights into the integrity with which teachers address the key ingredient of the intervention, namely the use of visual representations. We next investigated what kinds of visual representation teachers use and the extent to which teachers show an adaptive (i.e., diverse and flexible) use of representations. This information could inform teacher training and professionalization with respect to improving teachers' expertise in visual representation use. From a didactical point of view, it is also valuable to examine teachers' competence in using visual representations, as this may give further pointers for professional development. Thus, we also investigated the quality of the representation process and of the visual representations used.

Teachers' attention to visualization and its role in word problem solving

Teachers were trained in the use of the teaching intervention and were aware of its focus on visualization (i.e., the construction of a visual representation of problem structure). It could therefore be expected that teachers would pay the most attention to the VISUALIZE step, in comparison to the other problem solving steps. This was not the case, however: most teachers spent most time on other parts of the problem solving process, namely understanding problem text (UNDERSTAND) and computing answers (COMPUTE).

Extensive time spent on the UNDERSTAND step could be explained by the fact that teachers regularly did not apply the intended strategies for clarifying what was asked (e.g., paraphrasing the text and imagining the described situation) but rather spent a lot of time discussing irrelevant contextual information (e.g., what you can buy in a supermarket). This indicates that teachers are unfamiliar with and therefore need to master strategies for supporting text comprehension if they are to effectively support word problem solving of students.

Teachers' focus on correctly performing the required arithmetical computations (i.e., the COMPUTE step) is not surprising, given that instructional methods commonly used in mainstream classrooms and teacher training award much more attention to calculating correct answers than to understanding the problem text. Consequently, teachers are generally used to spending most time on the solution phase of the word problem solving process, and this practice is likely to have been perpetuated in the current setting.

It is also striking that teachers routinely combined the VISUALIZE step with the UNDERSTAND or COMPUTE step. This stands in sharp contrast to existing research-based programs developed to support word problem solving of low-performing students, which assume that word problem solving is a sequential (i.e., step-by-step) process (Krawec, 2010; Montague et al., 2000). The results of the present study showed that, in authentic classroom settings, word problem solving need not always occur sequentially - steps can be combined such that problem comprehension and solution interact and emerge together. Theories of word problem solving therefore need to recognize the potentially iterative nature of the process, particularly for non-routine problems.

Teachers' representational use and adaptivity therein

Teachers' use of visual-schematic representations (i.e., visual representations that represent the problem structure, the solution-relevant elements and the relations between them) played a central role in this research. Most teachers made ample use of these representations, though some seemed unclear about what these representations comprise and what function they serve within the word problem solving context. Moreover, some teachers also made frequent use of arithmetical representations (specifically, proportion tables). It is important to note that, in contrast to visual-schematic representations, arithmetical representations support only calculation processes rather than problem comprehension. Thus, when the problem text is not well understood, using only this kind of representation bears the risk that it does not contain the solution-relevant elements of the problem to be solved. Even after training, teachers appear not always to be aware of this difference in the use of these types of representations.

Teachers appear to have strong preferences for using one particular form of visual representation, namely a bar model or proportion table. Given these preferences, it is not surprising that teachers showed limited diversity and flexibility (i.e., adaptivity) in representation use. This could be explained by the fact that these representational forms are frequently offered in math textbooks in elementary schools and teacher education. Teachers consequently may feel more comfortable using them, as they encounter them more often and have more knowledge about them. This may also explain the finding that teachers rarely considered or compared the suitability of the representations used. Nonetheless, an imbalanced and/or inflexible use of representations can be problematic when teachers are unable to respond appropriately to students' needs (Jitendra et al., 2007) (e.g., students may find a number line more helpful than a bar model for certain sorts of problems) or to problem characteristics (e.g., a pie chart is more suitable than a bar model for solving a problem involving percentages). Thus, the adaptive use of visual representations is clearly an issue to be addressed in teacher training and professionalization.

The quality of the representation process and the visual representations produced

The finding that teachers demonstrated medium to high transparency in the way in which they provided explicit, step-by-step reasoning about what information should be represented and how to represent it is compatible with the nature of contemporary math education, which emphasizes reasoning processes and understanding (Barnes, 2005; Van den Heuvel-Panhuizen, 2003; Webb, Van der Kooij, & Geist, 2011). Several teachers also extensively involved students in this process; the explicit interaction between teacher and students is also one of the underlying principles of contemporary math education (Van den Heuvel-Panhuizen, 2003; Webb et al., 2011).

It is worrying, though, that while the majority of teachers generally demonstrated high correctness of reasoning, one teacher appeared to hold a fundamental misconception of what a bar model is. Furthermore, the reasoning of half of the teachers was not always complete: visual representations were introduced but not further or fully explicated. Such incomplete reasoning is risky if students do not know how to use the representation in question: misconceptions (such as held by the teacher mentioned above) can then arise that can be difficult to correct (Hill, Blunk, Charalambous, Lewis, Phelps, Sleep, & Ball, 2008). A correct and complete representation process is therefore essential but is not always exhibited in teachers' natural behaviors.

With respect to the quality of the representations themselves, it is remarkable that only two teachers were consistently able to construct representations that both correctly and completely contained all solution-relevant elements, in spite of the training that they had received. It is highly likely that mainstream teachers (who have not been explicitly trained in this area) would also have difficulty in producing correct and complete visual representations. Furthermore, some teachers were unable to come up with a suitable visual representation for some problems and seemed to think that they were making a visual representation when in fact they were sometimes merely structuring the words in the text. Clearly, teachers' expertise in this area needs to be improved. Also, some teachers produced representations containing excess, irrelevant information, largely as a consequence of insufficient understanding of what was being asked. Such representations make calculating the answer more difficult and error-sensitive than necessary, and make it unclear what information is solution-relevant and what is not. This suggests that teachers, as well as students, need to develop effective strategies for understanding what is being asked in non-routine word problems.

Implications for teacher professionalization in using visual representations to support word problem solving

Based on these findings, it can be concluded that the use of visual representations to support word problem solving should be given much more attention in teacher education and teacher professionalization programs. While the quality of the representation processes and the representations produced by most of the teachers in this study was reasonable, some misconceptions and inappropriate use of certain representational forms were observed. It could be argued that anything less than teachers' full mastery in this area is undesirable.

Furthermore, the limited diversity and flexibility are matters of concern. While routine, algorithmic problems can be solved using particular types of representations, non-routine problems - such as the authentic problems used in this study - require a problem-specific approach. In these cases, a limited and inflexible use of representations can result in an ineffective and inefficient problem solving process. Teachers therefore need to possess a broad repertoire of visual representations and understand their conditions of use, so that they are able to offer representations that both match problem characteristics and support students' needs. It is important, therefore, to develop teachers' knowledge of the characteristics and purpose of different types and forms of visual representation, as well as understanding when and how to use them to support word problem solving. This requires more than just the ability to produce accurate representations and needs to be given a prominent place in teacher education and teacher professionalization programs.

In addition, attention needs to be paid to mastering strategies for supporting text comprehension. Teachers need to learn how to identify the solution-relevant information in the word problem and, on the basis of that information, how to derive the specific questions that have to be answered. This will prevent teachers from spending too much time on irrelevant details which do not facilitate - and may even hinder - problem comprehension and solution.

Directions for future research

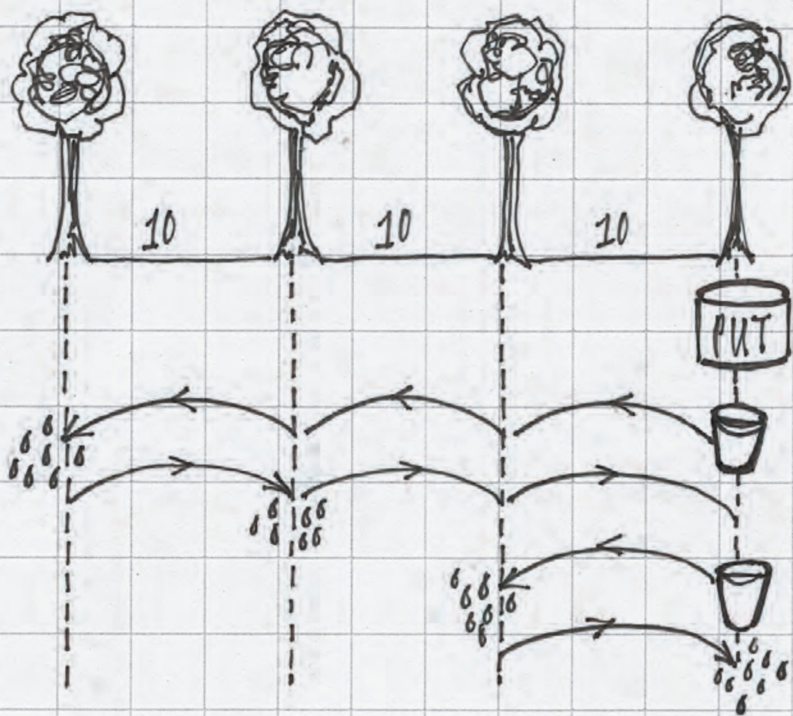
The participants in this study were recruited through purposive sampling of teachers who were motivated to participate in and considered themselves competent to contribute to research in this area. This ensured that results were obtained under favorable conditions in which teacher behaviour is not negatively influenced by motivational factors that are known to undermine the way in which teachers implement educational innovations in regular classrooms (e.g., Evers et al., 2002; Ghaith & Yaghi, 1997; Hermans et al., 2008; Rogers, 2003). Consequently, the findings cannot be directly generalised to the many mainstream teachers. Who have a negative attitude towards mathematics and are not confident about teaching mathematics (e.g., Bursal & Paznokas, 2006; Isiksal, Curran, Koc, & Askun, 2009; Swars, Daane, & Giesen, 2006). Nonetheless, it is important to realise that when even motivated and confident teachers experience difficulties in using visual representations to support word problem solving (such as limited diversity, flexibility and representational quality), difficulties experienced by other mainstream teachers might be more prominent. Thus, it is important to address the competences and needs of these teachers in future research.

Another deliberate choice is that teachers were encouraged to use their own explanations and elaborations, rather than a fully scripted lesson protocol. In this way, teachers could implement the teaching intervention in a way that is compatible with their own teaching approach and beliefs about teaching. This is important for the successful implementation and feeling of ownership of educational innovations (Ketelaar, Beijaard, Boshuizen, & Den Brok, 2012). Nevertheless, it makes the instruction vulnerable to potential shortcomings in teachers' skills. Thus, even when teachers believe they master the skills necessary to implement an instruction correctly, they should be provided with explicit training in its key ingredients and be given the opportunity to adopt and consolidate new skills before a teaching intervention such as this can be implemented in the classroom (Bitan-Friedlander, Dreyfus, & Milgrom, 2004).

To this end, training should be lengthy enough to provide teachers with enough time to internalize change, that is, accept the innovation, acquire the necessary skills and be prepared to implement it (Bitan-Friedlander et al., 2004). It is possible that the present training was not of sufficient duration so that the desired level of competence was not attained. Future research should investigate the duration and intensity of training required to achieve this level of competence.

At the same time, it is intriguing that some teachers believed themselves to be competent while in fact they were not, with one teacher even holding a fundamental misconception. As realistic beliefs about one's personal competence can positively influence individuals' willingness to invest in training and education (Beets, Flay, Vuchinich, Acock, Li, & Allred, 2008; Han & Weiss, 2005), it would be of interest to investigate how teachers can be helped to develop realistic beliefs about their pedagogical and didactical proficiency in the use of visual representations.

8 CONCLUDING REMARKS



Four young trees were set out in a row 10 meters apart. A well was situated beside the last tree. A bucket of water is needed to water two trees. How far would a gardener have to walk altogether if he had to water the four trees using only one bucket?

INTRODUCTION

The research presented in this thesis focused on the component processes and skills that underlie the successful comprehension of word problems. The studies conducted as part of this research investigated both students' use of visual representations and the quality of these representations. Specifically, we were interested in the extent to which different types of visual representation increase or decrease the chance of solving a word problem correctly.

The first objective of this thesis was to examine students' performances, notably the extent to which students use different types of visual representations. Furthermore, we examined the role that students' spatial and semantic-linguistic skills played in the solving of routine and non-routine word problems in early (second) and later (sixth) grades of elementary school.

The second objective of this thesis was to investigate how teachers implemented an innovative instructional approach – based on the didactical use of visual-schematic representations – in their own classroom teaching practice. This instructional approach required teachers to use visual-schematic representations that visualized the problem structure in a diverse and flexible way. Moreover, they were required to vary the kinds of visual representations in a way that suited the problem characteristics and students' individual needs. This approach made it necessary for teachers to model the representation process transparently, correctly and completely, as well as to construct visual representations that correctly and completely depicted the relations between all the components relevant to the solution of the problem.

Both of the objectives of this thesis, which were stated in Chapter 1 (General Introduction), were met by the research. The concluding remarks in this final chapter are focused on reflecting on the findings. In addition, I will examine possible implications of these findings for educational practice. On the basis of the findings described in the six studies that were included in this thesis, several observations and their consequences merit particular discussion and each will be addressed in due course.

First, I will discuss the findings with regard to the following: (a) the underlying processes of word problem solving; (b) the role of different types of visual representations; (c) the importance of semantic-linguistic and visual-spatial skills; and, (d) the didactical use of visual representations by elementary school teachers.

Secondly, I will draw implications for teacher training and teacher professionalization. Finally, I will conclude by reflecting on the reasons why the implementation of instructional innovations in mainstream educational practice is a challenging matter. This will also include an examination of the best course that is open to us if we are to bridge the existing gap between educational research and educational practice.

The underlying processes of word problem solving

In line with previous studies (see e.g., Carpenter, Corbitt, Kepner, Lindquist, & Reys, 1981; Cummins, Kintsch, Reusser, & Weimer, 1988; Krawec, 2010; Van der Schoot, Bakker-Arkema, Horsley, & Van Lieshout, 2009), this thesis revealed that comprehension of a word problem text is a determining factor with regard to students' performance on non-routine and routine word problems. In addition, we have gained insight into the component skills and abilities that underlie the comprehension of word problems.

On the basis of the results from Chapter 2, we can conclude that comprehension of non-routine word problems consists of the following elements: 1) the identification of relevant numerical and linguistic components, and of the relations between these components; 2) the visual representation of these components and relations in a complete and coherent way. While the identification of linguistic components and the relations between these components is semantic-linguistic in nature, the visual representation of these components and relations lies in the visual-spatial processing domain. Furthermore, the key basic ability in the semantic-linguistic domain is reading comprehension; whereas in the visual-spatial domain spatial ability is key.

In many previous studies these two processing domains were investigated separately. The importance of semantic-linguistic factors was mostly investigated in relation to routine word problem solving (Pape, 2003; Van der Schoot et al., 2009), whereas the use of visual representations was often examined in relation to non-routine word problem solving (Hegarty & Kozhevnikov, 1999; Van Garderen, 2006). However, the findings of our research reveal that both processing domains play a prominent role in non-routine word problem solving in the later grades of elementary school. Moreover, our findings make it clear that the identification of numerical and linguistic components and the visual representation of these relations are not separate processes, but seem to develop parallel to each other.

The production of different types of visual representations

When we looked more closely at the importance of the visual-spatial elements in word problem solving, by specifically examining the production of visual representations, our findings revealed that in most cases (sixth) grade students did not make use of a visual representation to solve a word problem. Furthermore, when a student did decide to use a visual representation, he/she did not always have the ability to construct a visual-schematic representation that could be called accurate. That is, a visual-schematic representation that contained the correct relations between all the elements relevant to the solution, and gave a coherent and complete view of the problem structure. The study described in Chapter 4 showed, for example, that students did construct pictorial representations and provide images of a specific element (i.e., object or person) of the word problem text when solving non-routine word problems. Furthermore, several students who made use of visual-schematic representations included incorrect relations in these types of representations, or produced incomplete visual-schematic representations (i.e., solution-relevant relations were missing). This finding showed that only making a distinction between pictorial and visual-schematic representations, as was the case in previous research (Hegarty & Kozhevnikov, 1999; Van Garderen, 2006; Van Garderen & Montague, 2003), is too limited. Our findings suggested that introducing a third type of visual representations, namely inaccurate visual-schematic representations, is justified.

Furthermore, because we used item-level analysis, we were better able to examine the extent to which these different types of visual representations contributed to students' word problem solving performance. We can conclude that only accurate visual-schematic representations increase the chance of solving a word problem correctly, and that our approach offers a more comprehensive view with respect to the importance of different types of visual representations than previous studies, which were based on test-level analysis.

The importance of visual-spatial and semantic-linguistic skills in routine word problems

Besides underscoring the importance of both the visual-spatial and semantic-linguistic processing domains in non-routine word problems (as shown in Chapter 2 and Chapter 4), our research also provided more insight into the role played by visual-spatial and semantic-linguistic skills in solving a specific type of routine word problems, namely compare word problems.

Based on the findings described in Chapter 3, we can conclude that sixth grade students who built a high quality mental representation of the problem structure experienced less difficulties with solving inconsistent compare problems that contained an unmarked relational term (i.e., 'more than'). The relational term in an inconsistent compare problem primed an inappropriate mathematical operation (e.g., 'more than' when the required operation is subtraction). However, translating a marked relational term like 'less than' into an addition operation was found to be closely associated with a student's performance on a general measure of semantic-linguistic skills, namely reading comprehension. In other words, reading comprehension skills together with sophisticated visual representation skills were found to be essential in dealing with semantically complex word problems.

The difficulties with solving compare word problems were, however, already visible in the early grades of elementary school. In line with Cummins et al., (1988), the findings of the study described in Chapter 6 showed that second grade students experienced more difficulties solving this type of routine word problem compared to combine and change word problems. Thus, semantic-linguistic skills do not only play a role in the solution of non-routine word problems, but are also important in routine word problems that have a complex semantic structure or, as in the case of students in higher grades of elementary school, in routine problems that contain semantically complex text elements (like relational terms).

In short, the findings of this research showed that both younger and older elementary school students experience difficulties with solving word problems. These difficulties were clearly caused by difficulties with the comprehension of the text of a word problem. Moreover, our findings also indicate that students often lacked a sophisticated visual representation strategy and/or were not able to deal with the semantic complexities of certain types of word problems.

Teachers' didactical use of visual representations

As teachers' behavior in the classroom (i.e., what teachers do in the classroom and how they interact with their students) plays a prominent role in the development of students (Kyriakides, Christoforou, & Charalambous, 2013; Kyriakides & Creemers, 2008), it can be assumed that teachers might perpetuate or even cause the difficulties that students experience.

On the basis of the findings of Chapter 3 for example, we can question whether enough attention is paid to the training of semantic-linguistic skills in educational practice. These findings showed that students who performed well on a standardized math test experienced a lot of difficulties with the solution of compare problems that asked for semantic-linguistic skills (i.e., inconsistent marked compare problems). If indeed these semantic-linguistic skills are not, or inadequately, trained in elementary school, this may lead to problems in secondary education where word problems contain more verbal information and become more complex.

When we examined the attention paid to the visual-spatial domain of word problem solving specifically, the findings of Chapter 7 showed that several mainstream teachers, like their students, experienced difficulties using accurate visual-schematic representations during word problem solving instruction in their own lessons. Significantly, even teachers who had followed a training on visual-schematic representation and who indicated that they felt competent, were not able to construct these visual-schematic representations in an accurate way in all the word problems they modeled. Moreover, we saw that teachers use mathematical representations instead of visual-schematic representations in a lot of situations. This is understandable when we take into account that math text books in contemporary math education and teacher training predominantly offer and teach student teachers to use mathematical representations. However, these specific types of visual representations play an important role only during the solution phase (i.e., they support the calculation process of mathematical operations and not the comprehension process).

These findings highlight the fact that the importance of accurate visual-schematic representations in the comprehension phase of word problem solving is still largely unrecognized both by those who set the curriculum and by teacher educators. Changing this situation is of the utmost importance, since the research presented in the present thesis indicates that comprehension of word problems causes the most difficulties in students and that these difficulties can be overcome by teaching the use of accurate visual-schematic representations.

Finally, the findings of the study in Chapter 7 revealed that when teachers did use accurate visual-schematic representations, in most cases a *bar model* was used. Other forms of visual-schematic representations that could have been used to elucidate the structure of the word problem, like *number lines*, *pie charts* or *own constructions,* were used only to a limited extent. Therefore, in general the level of diversity of forms of visual representations used by teachers was low. Moreover, the visual representations that teachers used did not always suit the problem characteristics and/or meet the individual needs of the students.

These findings also suggest that mainstream teachers who do not receive any training in word problem solving, and specifically in the use of visual-schematic representations as part of the solution process, experience even more difficulties. Evidently, if teachers are not able to construct accurate visual-schematic representations themselves, and, in addition, are not able to construct these representations in a diverse and flexible way, there is little reason to expect that their students will learn to use these types of visual representations. Therefore, in the next paragraph I will focus specifically on the implications of our findings for teacher professionalization and teacher training.

Implications for teacher professionalization and teacher training

The results of the studies described in this thesis, and of the study presented in Chapter 7 in particular, have clear implications for teacher professionalization and teacher training. These are summarized in the following list of recommendations:

School teachers and their students should be competent at constructing accurate visual-schematic representations as an aid in word problem solving, and the development of this competence should have a prominent place in the math curriculum of regular classrooms. (Student) teachers should have knowledge about the purpose of accurate visual-schematic representations and be trained in the construction and proper use of these types of visual representations.

Teacher professionalization and training should focus on the correct use of different forms of visual-schematic representations while solving a word problem. (Student) teachers should learn how to construct *number lines, pie charts, own constructions* and other appropriate forms of visual-schematic representations.

There should be a particular emphasis on the construction process of these types of visual representation. Teachers should be able to use the construction process in a transparent, correct, and complete manner. (Student) teachers should learn to make their reasoning transparent by explaining which elements of the problem should be represented, and how the representation can be used to solve the problem. This reasoning process should also be correct (e.g., naming and using visual representations correctly), as well as complete (e.g., indicating why and how a visual representation can be used).

Finally, teacher training should pay particular attention to teaching how to identify the characteristics of word problems. (Student) teachers should know the distinction between routine and non-routine word problems, and the role that accurate visual-schematic representations play in these word problem types. Namely, in routine word problems (like combine, change and compare problems) the use of only one type of visual representation can suffice, because the problem structure of each of these types of word problems is identical. However, the problem structure of non-routine word problems varies, which makes it inappropriate to offer only one kind of accurate visual-schematic representation. Hence, (student) teachers should learn to use visual-schematic representations in a way that is both diverse (i.e., demonstrating a varied use of visual representations) and flexible (i.e., offering different visual representations to solve one word problem). Moreover, these visual representations should be functional and suit the specific characteristics of the word problem (e.g., the use of a pie chart while solving word problems involving percentages).

Besides their importance in teacher professionalization and teacher training, the recommendations listed above provide interesting aspects for further research about the importance of visual representations in word problem solving. Based on our findings the focus of future studies should initially be on teachers' own competence and didactical use of visual representations during word problem solving instruction. Once teachers have more knowledge about the importance of visualization in the word problem solving process, they can use this knowledge to help their students successfully overcome the difficulties that they are experiencing (Antoniou & Kyriakides, 2013; Penuel, Fishman, Yamaguchi, & Gallagher, 2007).

Bridging the gap between educational research and educational practice

The research presented in this thesis adds valuable new insights into the processes, factors and skills that influence performance in word problem solving to a well-established discussion in the field of educational research. Moreover, throughout the years several innovative word problem solving instructions have been developed which were meant to implement findings of educational research into practice (e.g., Jitendra & Star 2012; Jitendra et al., 2009; Montague, 2003; Montague, Warger, & Morgan, 2000). Nevertheless, based on the observations of teaching practice reported in Chapter 7, and on the difficulties experienced by students in early and later grades of elementary school reported in Chapters 3 and 6, we can conclude that it is hard for these innovative instructions to find their way into actual classroom practice.

These findings contribute to the discussion of an important issue that has been frequently debated in the last decades, namely the gap between educational research and educational practice (e.g., Broekkamp & Van Hout-Wolters, 2007; McIntyre, 2005; VanderLinde & Van Braak, 2010). Doubts have been sometimes raised about the quality and relevance of educational research, because educational research often does not provide educational professionals with clear, practical answers (Broekkamp & Van Hout-Wolters, 2007; Biesta, 2007). The gap between educational research and educational practice is in this view a result of the existence of two mutually exclusive types of knowledge: on the one hand, research-based knowledge that is published in scientific journals; and on the other hand, pedagogical knowledge that is used by classroom teachers in their day-to-day teaching practice (McIntyre, 2005).

However, to maintain that educational research should be left to academics, and that educational practice should be the sole domain of practicing teachers, could be detrimental to both fields of education (Biesta, 2007). Educational research should be or should become evidence-based, and teaching should be or should become an evidence-based profession, as is already the case in several countries in the world (Biesta, 2007). This approach considers evidence-based teaching practice to be of great value. The role of research in education should to be to tell us 'what works' and a preferred way to discover 'what works' is through experimental studies (Biesta, 2007; Slavin, 2002). Hence, besides the systematic observation, recoding, and analysis of data and the publication of its findings, educational research also has an important role to play in the improvement of educational processes and the evidence-based evaluation of outcomes.

However, in a lot of cases the improvements and innovations proposed by educational researchers only make it to the stage of publication in journals, and never reach teachers and students in the classrooms (Vanderlinde & Van Braak, 2010). Important findings are consequently rarely brought to the attention of teachers. Once research has been published in an academic journal, researchers move on the next study, rather than attempting to relate their findings to the teaching practice (Stevens, 2004). A good line of communication between researchers and practitioners is absent, and practitioners are not encouraged to get actively involved in the research process (Könings, Brand-Gruwel, & Van Merriënboer, 2007). It is therefore necessary that more opportunities are made available to practitioners and researchers to collaborate, disseminate findings, co-construct ideas and set research agendas. Involving practitioners in the design and implementation of research makes it possible to link research and practice. More cooperation between researchers and practitioners goes hand in hand with a change in thought concerning the way that research is disseminated. There still is a traditional top-down model when it comes to the dissemination of educational innovations, in which innovations are developed by the researcher and then transferred to others in oral or written form. This linear way of dissemination should be replaced by a circular model, which emphasizes a two-way flow of information between researchers and practitioners and encourages practitioners to adapt and negotiate research findings within the context of their use (Nutley, Walter, & Davies, 2007).

When educational practitioners are actively involved in the research process it often becomes clear that what works in practice is not always as straightforward as research findings suggest and that success depends on several factors. Also, sometimes it becomes clear that innovations that work are not desirable in educational practice. Biesta (2007) gives a good example of this situation. He concludes that, although we have conclusive empirical evidence that in all cases physical punishment is the most effective way of deterring or controlling disruptive behavior, most societies would find it undesirable to choose an option that involves such a violation of human rights.

Another reason why it is difficult to know whether an innovation works, is that research can tell us what works in a particular situation, but not what will work in any future situation. Furthermore, teachers often perceive innovation as involving more work, time and energy than the traditional well-known methods, to which they may therefore tend to revert (Könings et al., 2007). As the quest for an evidence-based educational practice might not be viable, it might be more suitable to strive for a less stringent approach, such as evidence-informed, evidence-influenced or evidence-aware practice (Biesta, 2007).

Nevertheless, in order to realize an evidence-informed practice and successfully implement educational innovations, close collaboration and transparent communication between researchers and practitioners is vital. It is essential that the content of the innovation fits with practitioners' beliefs and desires about teaching. Research into teachers' adoption of educational innovations showed that their valuation of what they are required to perform (i.e., perceived importance and value of the innovation, and perceived amount of effort required to implement it) could determine the extent to which they act as intended (Bitan-Friedlander, Dreyfus, & Milgrom, 2004). Teachers' 'ownership' of an innovation, and therefore its success in practice, develops only once teachers are prepared to invest mentally or physically in its implementation at classroom level (Ketelaar, Beijaard, Boshuizen, & Den Brok, 2012).

These points have been taken into consideration when conducting the research described in this thesis. We tried to establish a close collaboration between educational research and practice in several ways: by making an inventory of the needs and desires of teachers with regard to word problem solving; by using authentic word problems that arose in the regular classroom practice of teachers and students; by involving teachers in the design of an innovative word problem instruction; and by letting teachers implement the innovation in their own classroom practice.

Our primary goal was to give teachers ownership of the innovation by giving them the opportunity to adapt the innovation to their wishes. We also took into consideration how important it is that teachers get the feeling that they are in control of their own actions when implementing the innovation (i.e., agency; Bijeaard, 2009; Ketelaar et al., 2012). Therefore, teachers were encouraged to use their own explanations and elaborations while implementing the innovative instruction, and to implement the teaching intervention in a way that was compatible with their own teaching approach.

An innovation may be regarded to have been successfully introduced once teachers have adopted it, are able to and willing to implement it in their classes, and are confident in their ability to adapt the innovation to the needs and abilities of their students (Bitan-Friedlander et al., 2004). However, in spite of the close collaboration with the educational practice the implementation of the innovative word problem instruction, which played a central role in Chapter 7, was not completely successful. This was not the consequence of a top-down approach or a lack of collaboration, but because teachers did not completely master the skills necessary to implement the innovation. This was particularly remarkable because the teachers who implemented the innovative instruction had earlier reported that, after they had received training in the use of the innovation, they felt confident and competent in using it. This indicates that teachers might not always be able to critically assess their competence and behavior in their own teaching practice.

This highlights another task that might be considered part of educational researchers' domain, namely teaching (student) teachers and other educational practitioners to evaluate and observe their own skills and teaching behavior in order to become 'self-reflecting' teachers. In the past four years this has been one of my main tasks as a teacher educator; educating student teachers to become teachers who are curious, open-minded and self-critical. Educational researchers who work in teacher training should also make (student) teachers aware of developments in research and help them to evaluate the significance of these developments. In order to do this successfully, educational researchers need to have a thorough understanding of educational practice, and to keep themselves informed about the wishes and desires of teachers and the circumstances in which they do their work.

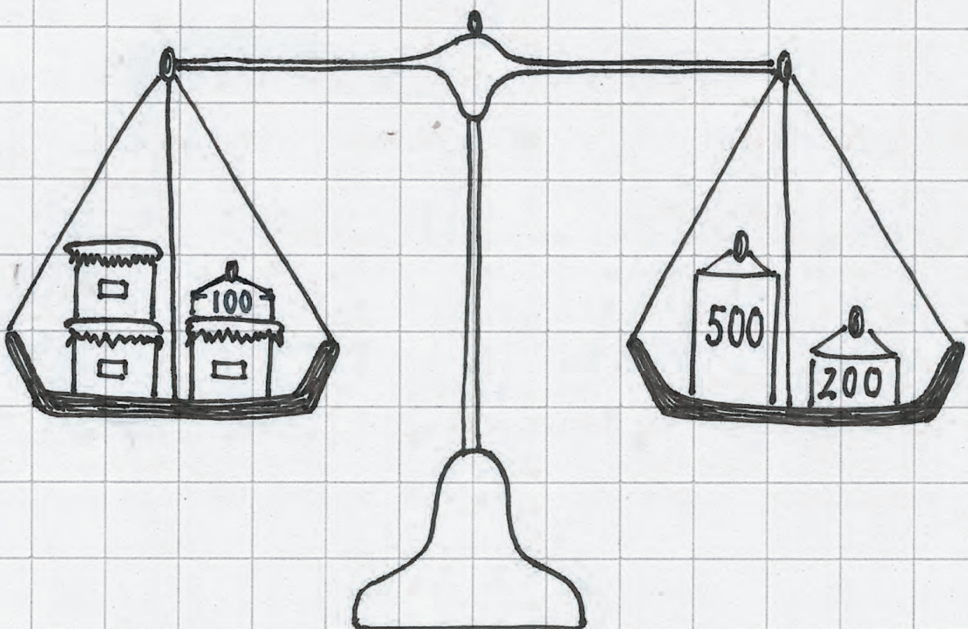
More cooperation between researchers and practitioners can, for example, be realized by promoting 'design-based research' or, maybe even more promising, by establishing 'professional learning communities' (PLC; Vescio, Ross, & Adams, 2008). In these PLCs a group of educators and researchers would meet regularly, share expertise, and work collaboratively to improve teaching skills and the academic performance of students. Initiatives such as the establishment of PLCs might be able to bridge the gap between educational research and educational practice. This could be an important step toward successfully fulfilling the main task of both educational researchers and teachers: the improvement of educational practice.

I APPENDIX

Teaching low performing second grade students to solve combine, change and compare mathematical word problems:

A feasibility study in four subjects

Anton J. H. Boonen, & Jelle Jolles
(under review)



On one side of a scale there are three pots of jam and a 100 g weight. On the other side there are a 200 g and a 500 g weight. The scale is balanced. What is the weight of a pot of jam?

ABSTRACT

The solution of combine, change and compare mathematical word problems causes many difficulties in especially young elementary school students. The principles underlying instructional programs like *Solve It!* and *schema-based instruction* could prove helpful in mastering these types of word problems.

This feasibility study examined four second-grade students who were less successful word problem solvers. These students received protocolled instruction during a five-week intervention period. The effectiveness of the word problem solving instruction was reported by comparing students' performances on the combine, change and compare problems before and after the intervention period, as well as by examining whether they executed the solution steps of the instruction correctly.

The results of the pre- and post-test comparison showed that the total word problem solving performance of all four students had improved. However, this improvement was not always visible in all three types of word problems. The study showed that the extent to which the solution steps had been executed correctly was a determining factor for the correct solution of the word problems.

While our findings do not imply that every student will benefit from a word problem instruction like the one we investigated, this feasibility study does provide important insights with regard to varying ways in which a word problem solving instruction can influence the solution strategies and performances of students who perform poorly on mathematical word problems.

INTRODUCTION

> ***[Word problem example]***
> *Mary has 9 marbles. She has 4 marbles more than John. How many marbles does John have?*

Tim, a seven-year-old boy who is in the second grade of elementary school, has difficulties with solving word problems like the one that is given in the example above. While solving these word problems, Tim often uses an impulsive, superficial solution strategy. Notably, he only focuses on selecting the presented numbers (9 and 4) and identifying the relational keywords (more than), which subsequently form the basis for his mathematical calculations. Tim's strategy often leads to an incorrect answer to the word problem. In this situation, Tim performed an addition operation where a subtraction operation was required, that is $9 + 4 = 13$ instead of $9 - 4 = 5$. The incorrect answer is not the result of a lack of calculation ability, but a result of a problem with deeply and correctly understanding the word problem text.

Mathematical word problem solving plays a prominent role in the curriculum of contemporary approaches to teaching mathematics (see Barnes, 2005; Elia, Van den Heuvel-Panhuizen, & Kovolou, 2009; Gravemeijer & Doorman, 1999; Van den Heuvel-Panhuizen, 2003). The solution of a word problem generally depends on two major phases: (1) *problem representation,* and (2) *problem solution.* The problem representation phase involves the identification and representation of the problem structure of the word problem. The identification and representation of the problem structure facilitate the correct understanding of the word problem text and help distill the mathematical operation(s) that should be performed. In the problem solution phase, on the other hand, the mathematical operations to be used are identified and the planned computations are executed to solve the problem (Krawec, 2010; Lewis & Mayer, 1987). Hence, errors in word problem solutions frequently occur in the problem representation phase, rather than in the problem solution phase. Improving students' problem representation skills is therefore of pivotal importance in order to help them master these word problems.

This article describes a feasibility study in which a word problem instruction was used to help students overcome their difficulties with understanding and representing the word problem text. It was set up as a multiple single-case study involving four subjects who participated in a five-week educational intervention course. The effectiveness of the instruction was reported by comparing students' performances before and after the intervention period, as well as by examining whether they executed the solution steps of the instruction correctly.

All of the subjects were second grade students who performed poorly as word problem solvers. Students from the second grade of elementary school were used as subjects in this study because difficulties with solving word problems already arise at an early age (see the example of Tim mentioned above, Cummins, Kintsch, Reusser, & Weimer, 1988). In our view, the current article presents convincing evidence that early intervention in the first grades of elementary school is imperative in order to address the difficulties experienced by young children. Our findings show that dedicated word problem solving instruction by means of a direct instruction method can play a pivotal role in this respect.

In the remainder of this introduction we provide some background information concerning the two instructional programs which were used as the basis of in our study: the *Solve it!* method and *schema-based instruction.*

Solve it! and schema-based instruction

Many researchers in the domain of mathematical word problem solving have documented the fact that students of all ages experience difficulties with solving a variety of types of word problems (e.g., Cummins et al., 1988; Hegarty, Mayer, & Green, 1992; Hegarty, Mayer, & Monk, 1995; Van der Schoot, Bakker-Arkema, Horsley & Van Lieshout, 2009). Yet, there is a scarcity of effective evidence-based instructional programs that address both the identification and representation of the problem structure (phase 1, see above), as well as the execution of the planned mathematical operations (phase 2, see above).

Instructional programs focusing on the explicit instruction of cognitive strategies to help students identify and represent the problem structure of a word problem seem to be effective (Jitendra et al., 2013; Jitendra & Star, 2012; Jitendra et al., 2009; Krawec, 2012). An example of this type of instruction is the *Solve It!* method (Montague, 2003). This is a heuristic instructional approach that teaches both elementary and middle school students how to: (a) read the problem for understanding; (b) paraphrase the problem by putting it into their own words; (c) visualize the problem (i.e., by constructing an external or internal visual-schematic representation); (d) set up a plan for solving the problem; (e) compute; and (f) verify the solution of the problem (Montague, Warger & Morgan, 2000).

Another commonly investigated example of a word problem solving instruction is *schema-based instruction* (SBI, developed by Jitendra et al.,). Schema-based instruction is generally provided in the early grades of elementary school and uses schema training to help students see the underlying (mathematical) structure of the word problem. Students are taught to recognize the similarities and differences between types of word problems and to identify and represent their problem structures (Jitendra, George, Sood, & Price, 2010). Schema-based instruction is particularly prescriptive in nature, because visual representations of the problem structure of word problems are provided to the children. However, several studies have recently shown that it is more effective to teach students to construct their own visual-schematic representations, instead of providing them with the representations (Van Dijk, Van Oers, & Terwel, 2003; Van Dijk, Van Oers, Terwel, & Van den Eeden, 2003a). Moreover, SBI has been shown to be difficult to master for low achieving students (Jitendra et al., 2002, 2013).

The aim of the present study is to describe and evaluate an instructional approach primarily based on the principles of the Solve it! method and of SBI. An important 'technical' element of the approach taken by us is that the students should be trained to identify and construct the visual representations themselves, instead of being provided with visual representations by teachers. Specifically, this approach addresses visual representations of the problem structures of the three types of word problems that are frequently offered in the first grades of elementary school, namely combine, change and compare word problems.

METHODS

Participants and instruments

Four second-grade students (three boys *Hugo*, *Peter*, *Tim* and a girl *Lisa*) took part in the study. They attended a mainstream sub-urban elementary school in the Netherlands and were native Dutch speakers. All students were healthy and their intelligence was in the normal range. Table 1 shows students' age and their performances on the nationwide standardized norm-referenced CITO (Institute for Educational Measurement; www.cito.nl) Mathematics and Technical Reading test. According to these norms, level A corresponds to the highest 25% of the norm-referenced population, level B to the above-average 25%, level C to the below-average 25%, and level D and E together to the lowest 25%. The four students were selected on the basis of their performance on a nine-item *word problem pre-test* (Cronbach's alpha = .82). This test included three combine, three change and three compare problems (see Table 2). The word problem items were presented on a different page and administered by the teacher in a session – attended by the four subjects - of approximately 30 minutes. Each word problem was read out loud twice by the teacher to control for differences in decoding skills. After reading the word problem, students had to solve the word problem within three minutes and during this time the teacher did not speak to the student. An examination of the number of word problems completed (see Table 1) showed that all four students experienced significant difficulties solving them. In contrast to their classmates (who solved the most of the nine word problems correctly; $M = 8.53$, $SD = 1.23$), the four research subjects solved less than half of the nine word problems items correctly.

After the intervention period a post-test with nine similar word problems was administered. Although the problem structure of these word problems was identical to the word problems that were included in the pre-test, we adjusted the figures in order to prevent the occurrence of a learning effect. The female teacher who executed all testing and instruction was 28 years old and had three years of experience in teaching second grade students. She had also obtained her master degree in Educational Sciences.

Table 1. *Age, Mathematical and Technical Reading achievement and pre-test score of the four research subjects*

Variable	Student			
	Hugo	**Peter**	**Tim**	**Lisa**
Age (months)	84	90	94	92
Achievement				
Mathematics[a]	B	B	C	B
Technical Reading[a]	E	D	B	C
Pre-test score	4/9	0/9	2/9	4/9

Note: a Norm scores on the nationwide standardized CITO (Institute for Educational Measurement) Mathematics & Technical reading test (2012). Between brackets: A= high 25%, B = above average 25%, C= average 25%, D = below average 25%, E = low 25%.

Table 2. *The nine items of the word problem solving pre-test (taken from Cummins et al., 1988)*

Combine word problems

1. Mary has 2 marbles. John has 5 marbles. How many marbles do they have altogether?
2. Mary has 4 marbles. John has some marbles. They have 7 marbles altogether. How many marbles does John have?
3. Mary and John have 8 marbles altogether. Mary has 7 marbles. How many marbles does John have?

Change word problems

1. Mary had 3 marbles. Then John gave her 5 marbles. How many marbles does Mary have now?
2. Mary had 2 marbles. Then John gave her some marbles. Now Mary has 9 marbles. How many marbles did John give to her?
3. Mary had some marbles. Then John gave her 3 marbles. Now Mary has 5 marbles. How many marbles did Mary have in the beginning?

Compare word problems

1. Mary has 5 marbles. John has 8 marbles. How many marbles does John have more than Mary?
2. Mary has 3 marbles. John has 4 marbles more than Mary. How many marbles does John have?
3. Mary has 4 marbles. She has 3 marbles less than John. How many marbles does John have?

Intervention materials

The intervention was a word problem solving instruction primarily based on the Solve It! instruction program (Montague, 2003) and on SBI (Jitendra, DiPipi, & Perron-Jones, 2002, Jitendra et al., 2009; Jitendra & Star, 2012). Specifically, the solution steps that were part of both instructional methods were merged and adjusted in such a way that they were understandable for second grade students and followed the solution process of three different types of word problems. This resulted in five solution steps that taught students to:

Step 1: READ the word problem for understanding: Students were taught to read each sentence of the word problem text critically and not only look for numbers and keywords, like more than, times, just as much, etcetera;

Step 2: VISUALIZE the word problem: Students were taught to identify and externally represent the problem structures of the combine, change and compare word problem type (i.e., the production of a visual-schematic representation);

Step 3: Add a QUESTION MARK to the visual-schematic representation to indicate the variable that has to be calculated;

Step 4: COMPUTE the required operation; and

Step 5: DRAW A CIRCLE around the variable that had to be calculated in order to check if the required solution was reported.

Procedure: The execution of the solution steps of the instruction

Over the course of five weeks, the four students were taught to use these five solution steps to improve their solution strategies and performances. Each of the ten sessions that were offered to the students (in two group sessions of 30 minutes per week) included teacher-mediated instruction that addressed the use of the solution steps. The amount and intensity of the instructional support of the teacher was gradually faded within and across the sessions. The way in which the instruction for each of the three types of word problems was offered is elaborated below.

<u>The solution of combine word problems (week 1 of the instruction)</u>

In the first week the five solution steps of the instructional approach were introduced by teaching the students to solve the type of problem known as a combine word problem (see word problem example 1).

[Word problem example 1]
Mary has 3 marbles. John has 5 marbles. How many marbles do they have altogether?

In a combine word problem a subset or superset must be computed given the information about two other sets. This type of problem involves understanding part-whole relationships and knowing that the whole is equal to the sum of its parts (Cummins et al., 1988; Jitendra, 2002, Jitendra et al., 2002). The five solution steps of the instruction to solve word problem example 1 are offered to the students in the following way:

Step 1: The combine problem is read aloud twice by one of the four students.
Step 2: Students are taught to visualize the problem structure of the combine problem by making an external visual representation.

The three marbles that Mary has are drawn first (see Figure 1):

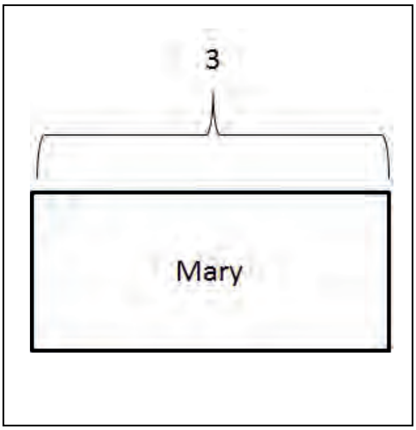


Figure 1. Step 2: VISUALIZE the marbles that Mary has

Next, the five marbles that John has are drawn (see Figure 2):

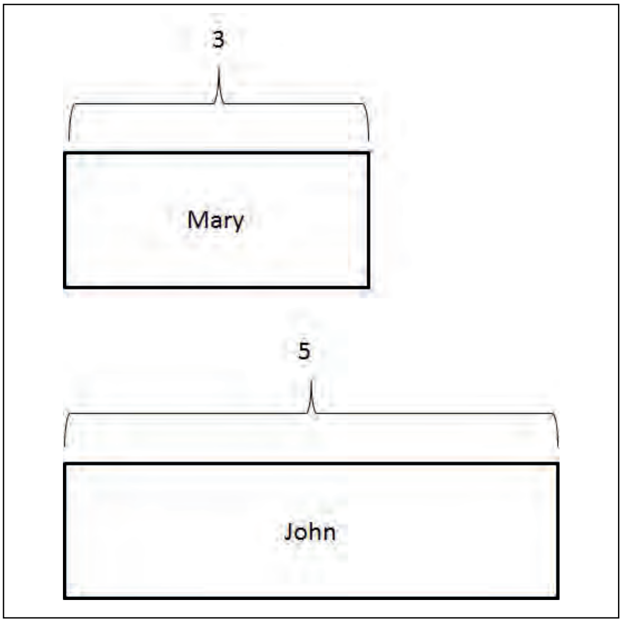


Figure 2. Step 2: VISUALIZE the marbles that John has

Step 3: A question mark is added to the visual-schematic representation to indicate the 'unknown' value that should be calculated (see Figure 3):

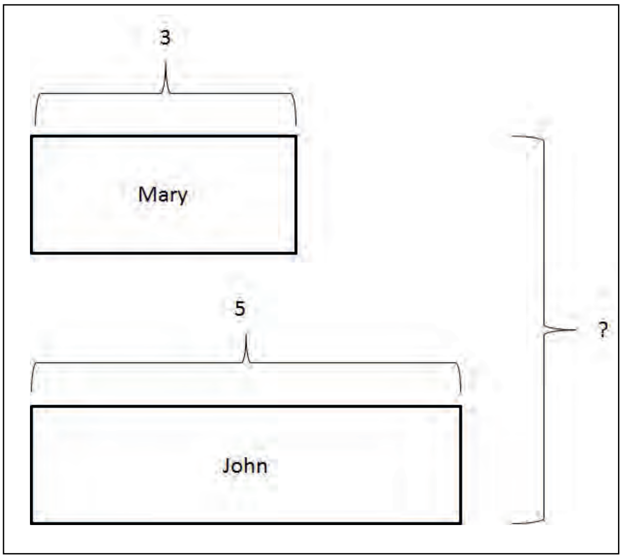


Figure 3. Step 3: Add a QUESTION MARK to the visual-schematic representation to indicate the variable that has to be calculated

Step 4: The required mathematical operation is written down (3 + 5 =) and solved (3 + 5 = 8).
Step 5: A circle is drawn around the 'unknown' value (i.e., 8), to be sure that the required answer is reported (see Figure 4).

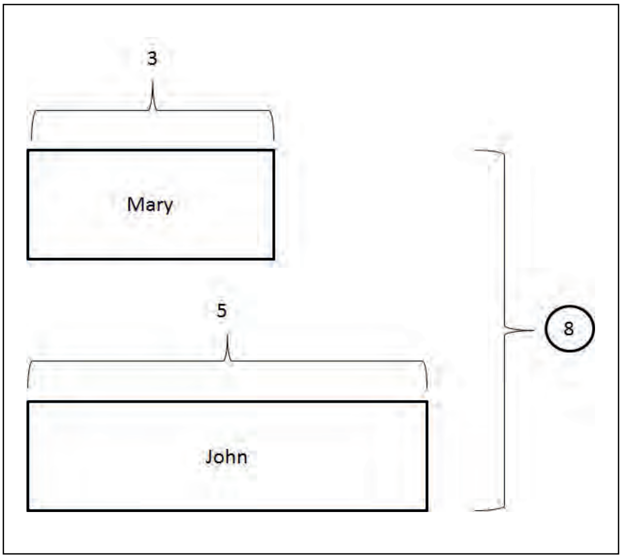


Figure 4. Step 5: DRAW A CIRCLE around the variable that had to be calculated

The solution of change word problems (week 2 of the instruction)
In the second week of the intervention period the four students were trained to solve change word problems (see word problem example 2).

> ***[Word problem example 2]***
> *Mary had 2 marbles. Then John gave her some marbles. Now Mary has 9 marbles. Hoe many marbles did John give to her?*

A change problem starts with a beginning set in which the object identity and the amount of the object are defined. Then a change occurs to the beginning set that results in an 'ending set' in which the new amount is defined (Jitendra, 2002). The solution of a change problem is instructed in the following way:

Step 1: The change problem was read aloud twice by the one of the students.

Step 2: Students are taught to visualize the problem structure of the change problem by making an external visual representation.

The two marbles that Mary has are drawn first (see Figure 5):

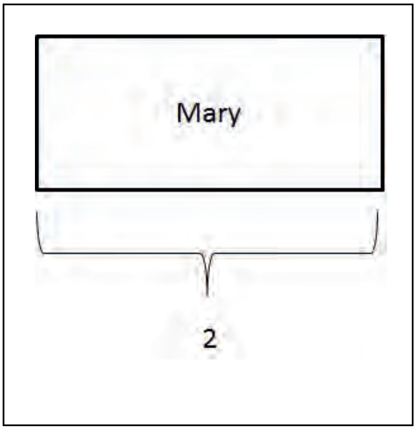


Figure 5. Step 2: VISUALIZE the marbles that Mary has

Next, the marbles that John gave to Mary are drawn (see Figure 6):

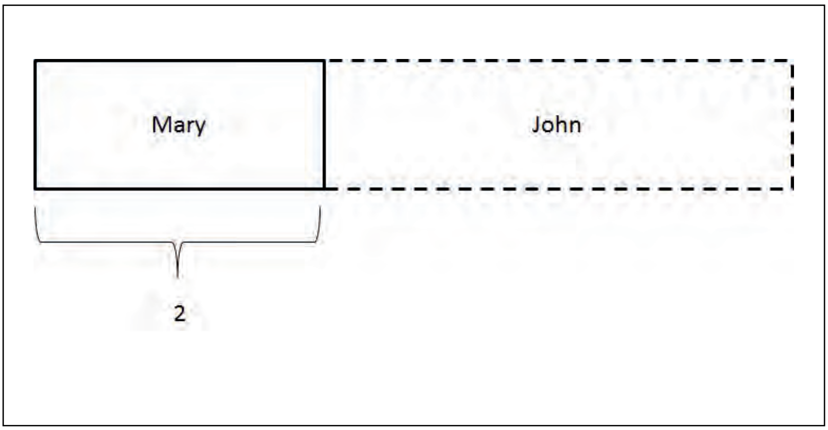


Figure 6. Step 2: VISUALIZE the marbles that John gave to Mary

Finally, the total amount of marbles is added to the visual-schematic representation (see Figure 7).

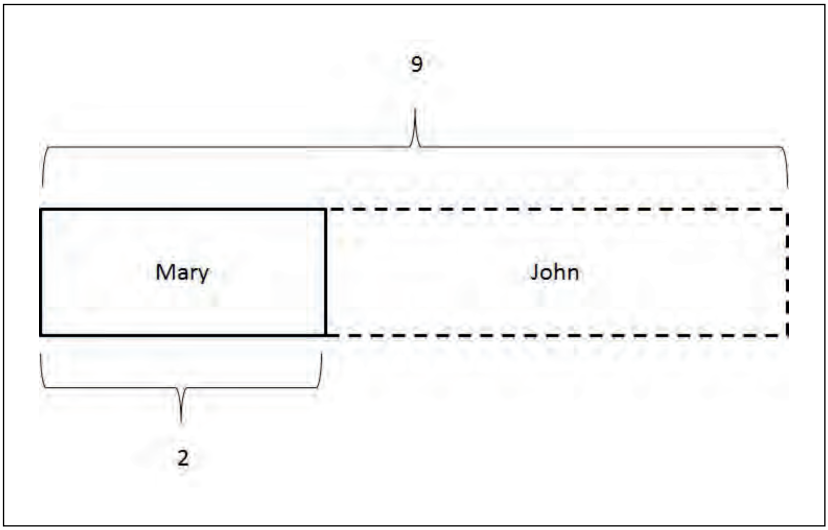


Figure 7. Step 2: VISUALIZE the amount of marbles that Mary and John have

Step 3: A question mark is added to the visual-schematic representation to indicate the 'unknown' value that should be calculated (see Figure 8):

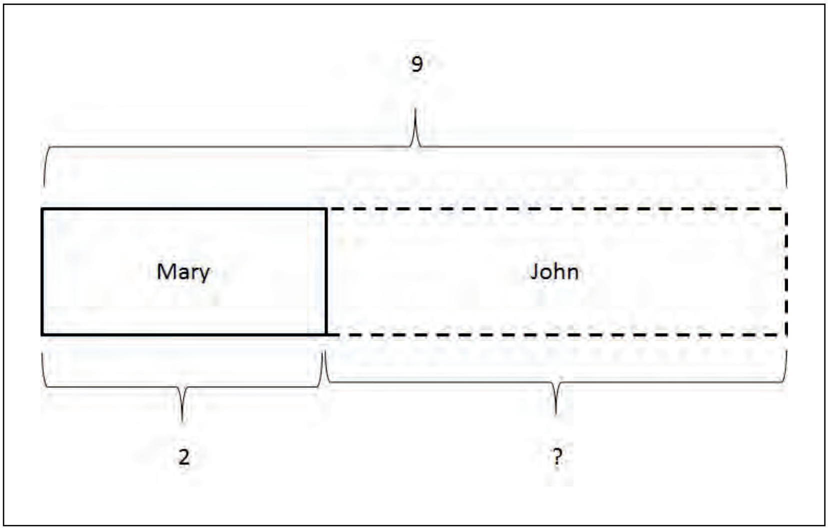


Figure 8. Step 3: Add a QUESTION MARK to the visual-schematic representation to indicate the variable that has to be calculated

Step 4: The mathematical operation is distilled from the visual representation, written down (2 + ? = 9) and solved (2 + 7 = 9).

Step 5: A circle is drawn around the 'unknown' value (i.e., 7), to be sure that the required answer is reported (see Figure 9).

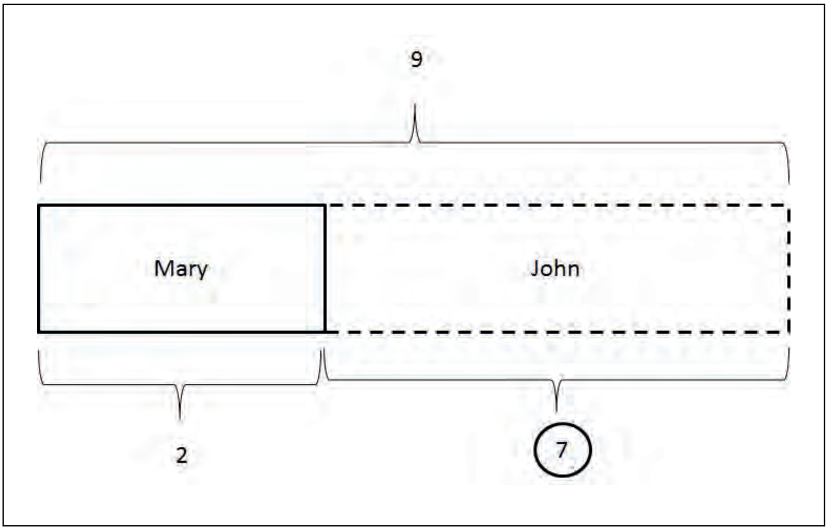


Figure 9. Step 5: DRAW A CIRCLE around the variable that had to be calculated

The solution of compare word problems (week 3 of the instruction)

In the third week of the intervention period the four students were trained to solve compare word problems (see word problem example 3).

> ***[Word problem example 3]***
> *Mary has 5 marbles. John has 8 marbles. How many marbles does John have more than Mary?*

In compare problems the cardinality of one set must be computed by comparing the information given about relative sizes of the other set sizes; one set serves as the comparison set and the other as the referent set. The solution of a compare problem is instructed in the following way:

Step 1: The compare problem was read aloud twice by the one of the students.

Step 2: Students are taught to visualize the problem structure of the compare problem by making an external visual representation.

The five marbles that Mary has are drawn first (see Figure 10):

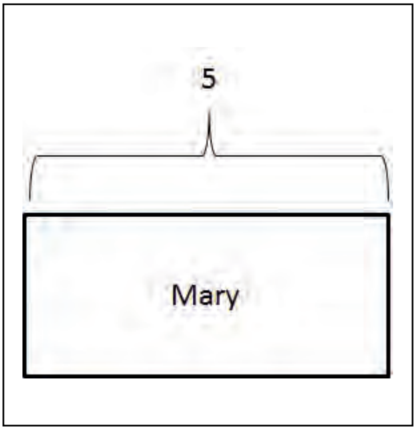


Figure 10. Step 2: VISUALIZE the marbles that Mary has

Next, the 8 marbles that John has are drawn (see Figure 11):

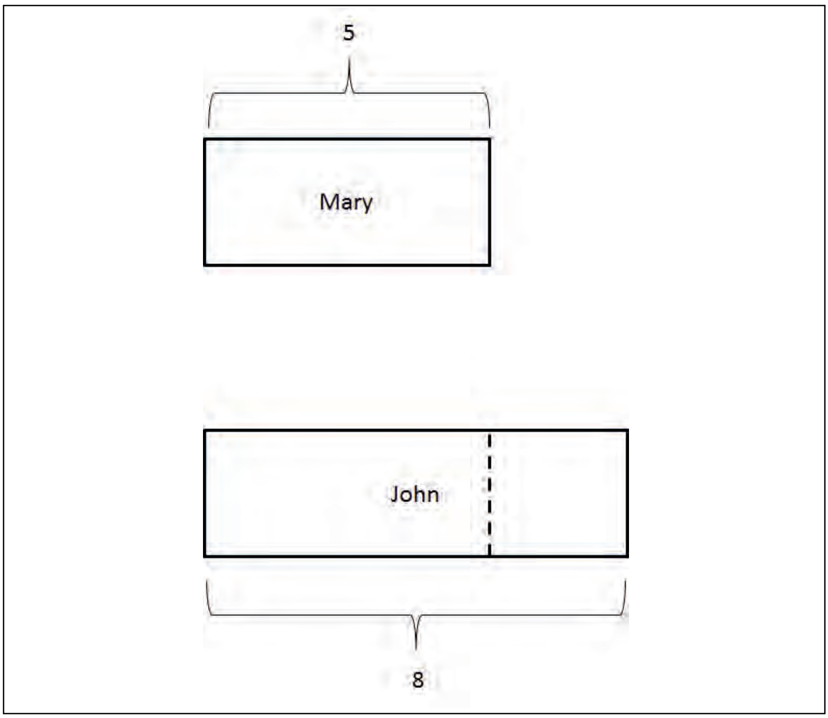


Figure 11. Step 2: VISUALIZE the marbles that John has

Step 3: A question mark is added to the visual-schematic representation to indicate the 'unknown' value that should be calculated (see Figure 12):

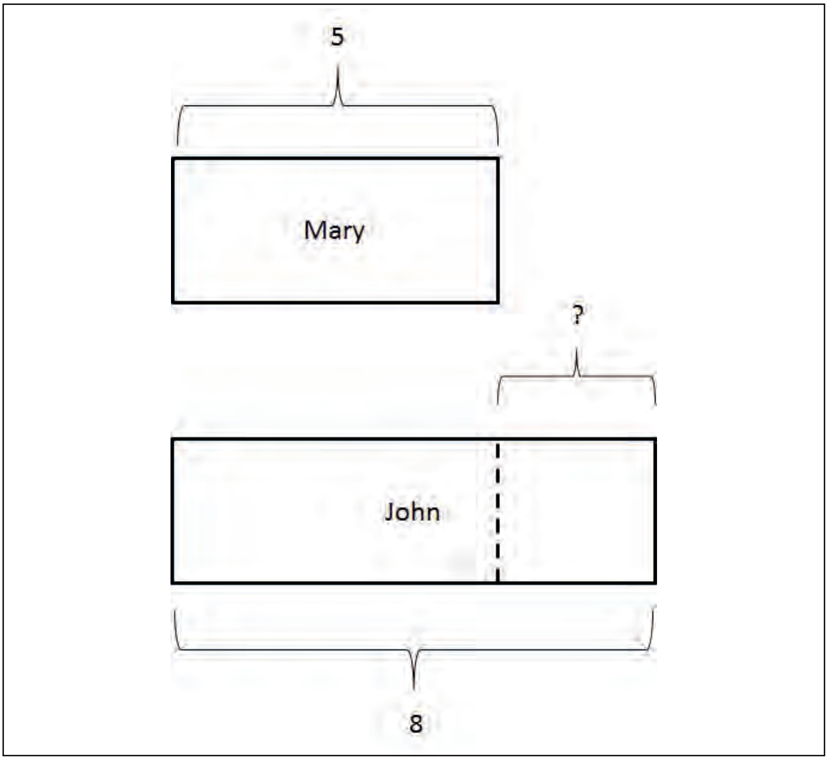


Figure 12. Step 3: Add a QUESTION MARK to the visual-schematic representation to indicate the variable that has to be calculated

Step 4: The required mathematical operation is written down (5 + ? = 8) and solved (5 + 3 = 8).
Step 5: A circle is drawn around the 'unknown' value (i.e., 3), to be sure that the required answer is reported (see Figure 13).

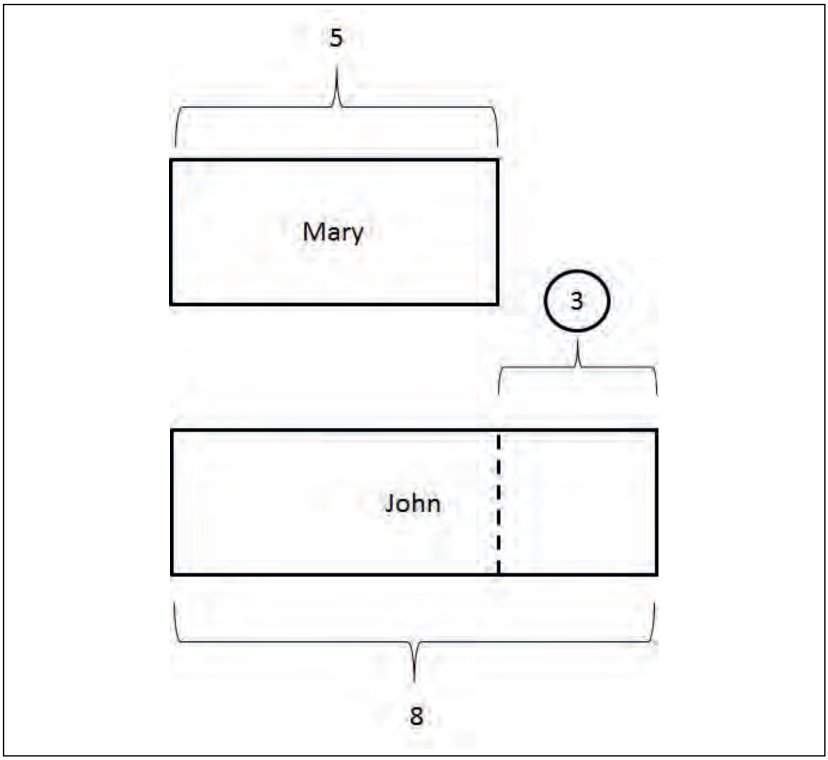


Figure 13. Step 5: DRAW A CIRCLE around the variable that had to be calculated

In the fourth and fifth week the three types of word problems were repeated and randomly presented to the students to check whether they were able to make the transfer between the problem structures of each type of word problem.

RESULTS

Students' word problem solving performances and use of strategy in the pre- and post-test

Table 3 gives an overview with respect to the performances on combine, change and compare problems before and after the intervention period.

Table 3. *Performance pre- and post-test for combine, change and compare problems*

Subject	Type of word problem	Pre-test score	Post-test score
Hugo	Combine	1	3
	Change	1	2
	Compare	2	1
	Total	4	6
Peter	Combine	0	3
	Change	0	3
	Compare	0	3
	Total	0	9
Tim	Combine	1	3
	Change	1	3
	Compare	0	3
	Total	2	9
Lisa	Combine	1	2
	Change	2	0
	Compare	1	3
	Total	4	5

For each of the four subjects the performances and solution strategies used are reported for each type of word problem separately.

Combine word problems

The following three word-problem items were included in the pre-test and post-test (between brackets the adjusted figures of the post-test):

Combine 1. Mary has 2 (3) marbles. John has 5 (6) marbles. How many marbles do they have altogether?
Combine 2. Mary has 4 (4) marbles. John has some marbles. They have 7 (6) marbles altogether. How many marbles does John have?
Combine 3. Mary and John have 8 (10) marbles altogether. Mary has 7 (9) marbles. How many marbles does John have?

Hugo

Pre-test. The results of the pre-test showed that Combine 1 was solved correctly by Hugo (i.e., 2 + 5 = ? [7]). Combine 2 was, however, solved incorrectly (answer Hugo = 7, required answer = 3). The mathematical operation that he reported showed that Hugo had difficulties finding the required answer (Hugo: 4 + 3 = ? [7]; required: 4 + ? = 7 [3]). Also Combine 3 was solved incorrectly (Hugo: 7 + 8 = ? [15]; required: 7 + ? = 8 [1]). Hugo's decision to add the two known figures of Combine 3 reflected a difficulty with comprehending the text of the word problem. In the pre-test Hugo only wrote down the mathematical operations that he performed and reported no other solution strategies.

Post-test. The results of the post-test showed that Combine 1 still did not cause any difficulties (i.e., 3 + 6 = ? [9]). Also Combine 2 and Combine 3 were solved correctly after the intervention period (Combine 2: 4 + ? = 6 [2]; Combine 3: 9 + ? = 10 [1]). With regard to the solution strategies that were used in the post-test, Hugo correctly visualized the problem structure, added a question mark in the right place in the visual-schematic representation, and correctly drew a circle around the 'unknown' variable in all three word-problem items.

I

Peter

Pre-test. The results of the pre-test showed that Peter solved Combine 1 incorrectly (Peter: 5 x 2 = ? [10]; required: 2 + 5 = ? [7]). Instead of performing an addition operation, Peter used a multiplication operation, reflecting a difficulty with distilling the correct mathematical operation from the word problem text. Combine 2 was also solved incorrectly (answer Peter = 7; required answer = 3). The mathematical operation that he reported showed that Peter had difficulties finding the required answer (Peter: 4 + 3 = ? [7]; required: 4 + ? = 7 [3]. The same situation applied to Combine 3, where Peter reported the incorrect mathematical operation (Peter: 7 + 1 = ? [8]; required: 7 + ? = 8 [1]). In the pre-test Peter only wrote down the mathematical operations that he performed, and reported no other solution strategies.

Post-test. The results of the post-test showed that Combine 1 was solved correctly (i.e., 3 + 6 = ? [9]). Also Combine 2 and Combine 3 were solved correctly after the intervention period (Combine 2: 4 + ? = 6 [2]; Combine 3: 9 + ? = 10 [1]). With regard to the solution strategies that were used in the post-test, Peter correctly visualized the problem structure, added a question mark in the right place in the visual-schematic representation, and correctly drew a circle around the 'unknown' variable in Combine 1 and 2. In Combine 3 Peter only used one specific solution step: he correctly drew a circle around the unknown variable.

Tim

Pre-test. The results of the pre-test showed that Combine 1 did not cause any difficulties (i.e., 2 + 5 = ? [7]). Combine 2 and Combine 3 were, however, solved incorrectly by Tim (answer Tim Combine 2 = 10; required answer = 3; answer Tim Combine 3 = 8; required answer = 1). These errors were a result of the fact that Tim had difficulties understanding the text of the word problem. This was reflected by the mathematical operations that he reported (Combine 2: 3 + 7 = ? [10]; required: 4 + ? = 7 [3]; Combine 3: 4 + 4 = ? [8]; required 7 + ? = 8 [1]). In the pre-test Tim only wrote down the mathematical operations that he performed and reported no other solution strategies.
Post-test. The results of the post-test showed that Combine 1 still did not cause any difficulties (i.e., 3 + 6 = ? [9]). Also Combine 2 and Combine 3 were solved correctly after the intervention period (Combine 2: 4 + ? = 6 [2]; Combine 3: 9 + ? = 10 [1]). With respect to the solution strategy that was used in the post-test, Tim only used the step in which a circle had to be drawn around the 'unknown' variable. He executed this step in all three Combine word problems correctly.

Lisa

Pre-test. The results of the pre-test showed that Combine 1 did not cause any difficulties (i.e., 2 + 5 = ? [7]). Combine 2 was, however, solved incorrectly (answer Lisa = 10; required answer = 3). Lisa made two errors with respect to this specific word problem: (1) she reported the incorrect mathematical operation (Lisa: 4 + 7 = ?; required: 4 + ? = 7); and (2) she made a calculation error (Lisa: 4 + 7 = 10; required: 4 + 7 = 11). Also Combine 3 was solved incorrectly (Lisa: 7 + 8 = ? [15]; required: 7 + ? = 8 [1]). Lisa's decision to add the two known figures in Combine 3 reflected a difficulty with comprehending the text of a word problem. In the pre-test Lisa only wrote down the mathematical operations that she performed, and reported no other solution strategies.
Post-test. The results of the post-test showed that Combine 1 still did not cause any difficulties (i.e., 3 + 6 = ? [9]. Looking at the solution strategies that were used, the results showed that Lisa correctly visualized the problem structure, added a question mark in the right

place in the visual-schematic representation, and correctly drew a circle around the unknown variable. Although these three steps were also correctly executed in Combine 2, Lisa's answer on this word problem was incorrect (Lisa: 4 + 2 = ? [6]; required: 4 + ? = 6 [2]). This incorrect answer was a result of her difficulty with finding the required answer. Combine 3 was solved correctly by Lisa (i.e., 9 + ? = 10 [1]). In this word problem Lisa executed only the last solution step.

Change word problems

The following three word-problem items are included in the pre-test and post-test (between brackets the adjusted figures of the post-test):

Change 1. Mary had 3 (2) marbles. Then John gave her 5 (4) marbles. How many marbles does Mary have now?
Change 2. Mary had 2 (3) marbles. Then John gave her some marbles. Now Mary has 9 (7) marbles. Hoe many marbles did John give to her?
Change 3. Mary had some marbles. Then John gave her 3 (4) marbles. Now Mary has 5 (9) marbles. How many marbles did Mary have in the beginning?

I

Hugo
Pre-test. The results of the pre-test showed that Change 1 was solved correctly by Hugo (i.e., 3 + 5 = ? [8]). Change 2 was, however, solved incorrectly (Hugo: 2 + 9 = ? [11]; required: 2 + ? = 9 [7]). Hugo's decision to add the two known figures in Change 2 reflected a difficulty with comprehending the text of a word problem. Also Change 3 was solved incorrectly (Hugo: 2 + 5 = ? [7]; required: 3 + ? = 5 [2]). Hugo's solution of this word problem showed that he had difficulties comprehending the word problem text and distilling the correct mathematical operation. In the pre-test Hugo only wrote down the mathematical operations that he performed and reported no other solution strategies.
Post-test. The results of the post-test showed that Change 1 still did not cause any difficulties (i.e., 2 + 4 = ? [6]). Also Change 2 was solved correctly after the intervention period (Hugo: 3 + ? = 7 [4]). With regard to the solution strategies that were used in Change 1 and Change 2, Hugo correctly visualized the problem structure, added a question mark in the right place in the visual-schematic representation, and correctly drew a circle around the 'unknown'

variable in these two word problem items. Change 3 was, however, solved incorrectly (no answer was given). Looking at the execution of the solution steps, Hugo seemed to have difficulties visualizing the problem structure and adding a question mark in the right place in the visual-schematic representation (see Figure 14).

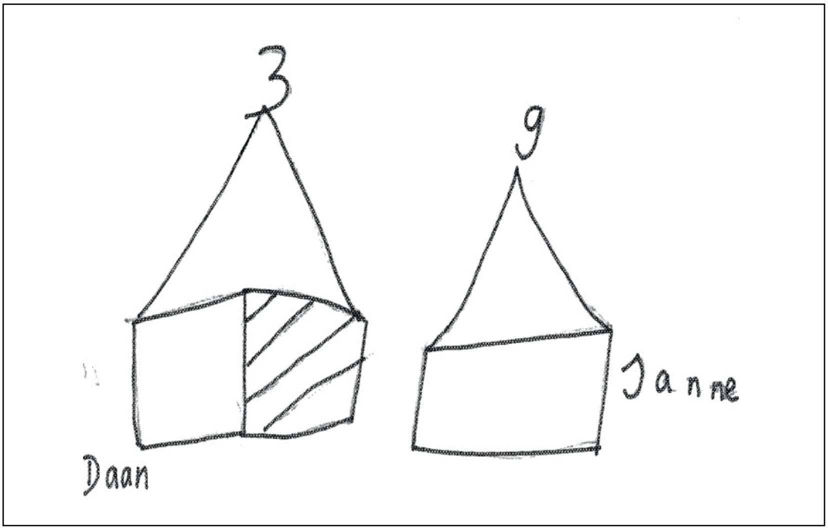


Figure 14. Hugo's incorrect visualization of Change 3 (Note: in the Dutch translation of the word problem test Mary and John were replaced by Sanne and Daan)

Peter

Pre-test. The results of the pre-test showed that Peter solved Change 1 incorrectly (Peter: 5 x 3 = ? [15]; required 3 + 5 = ? [8]). Instead of performing an addition operation, Peter used a multiplication operation, reflecting a difficulty with distilling the correct mathematical operation from the word problem text. Change 2 was also solved incorrectly (answer Peter = 9; required answer = 7). The mathematical operation that was reported showed that Hugo had difficulties finding the required answer (Peter: 2 + 7 = ? [9]; required: 2 + ? = 9 [3]). The same applies to Change 3 where Peter reported the incorrect mathematical operation (Peter: 2 + 3 = ? [5]; required: 3 + ? = 5 [2]). In the pre-test Peter only wrote down the mathematical operations that he performed and reported no other solution strategies.

Post-test. The results of the post-test showed that Change 1 was solved correctly (i.e., 2 + 4 = ? [6]). Also Change 2 and Change 3 were solved correctly after the intervention period (Change 2: 3 + ? = 7 [4]; Change 3: 4 + ? = 9 [5]). With regard to the solution strategies that were used in the post-test, Peter correctly visualized the problem structure, added a question mark in the right place in the visual-schematic representation, and correctly drew a circle around the 'unknown' variable in all Change word problems.

Tim

Pre-test. The results of the pre-test showed that Change 1 did not cause any difficulties (i.e., 5 + 3 = ? [8]). Change 2 and Change 3 were, however, solved incorrectly (answer Tim Change 2 = 9; required answer = 7; answer Tim Change 3 = 5; required answer = 2). The mathematical operations that were reported showed that Tim had difficulties finding the required answer (Change 2: 2 + 7 = ? [9]; required: 2 + ? = 9 [7]; Change 3: 3 + 2 = ? [5]; required: 3 + ? = 5 [2]). In the pre-test Tim only wrote down the mathematical operations that he performed, and reported no other solution strategies.
Post-test. The results of the post-test showed that Change 1 still did not cause any difficulties (i.e., 3 + 4 = ? [6]). Also Change 2 and Change 3 were solved correctly after the intervention period (Change 2: 3 + ? = 7 [4]; Change 3: 4 + ? = 9 [5]). With regard to the solution strategies that were used in the post-test, Tim only used the step in which a circle should be drawn around the 'unknown' variable. This step was executed correctly in all three Change word problems.

Lisa

Pre-test. The results of the pre-test showed that Change 1 did not cause any difficulties (i.e., 3+ 5 = ? [8]). Change 2 was, however, solved incorrectly (answer Lisa = 11; required answer = 7). Lisa reported the incorrect mathematical operation (Lisa: 2 + 9 = ?; required: 2 + ? = 9). Her decision to add the two known figures in Change 2 reflected a difficulty with comprehending the text of a word problem. Change 3 was solved correctly (i.e., 3 + ? = 5 [2]). In the pre-test Lisa only wrote down the mathematical operations that she performed and reported no other solution strategies.
Post-test. The results of the post-test showed that Change 1 was not answered. Looking at the execution of the solution steps, Lisa seemed to have difficulties visualizing the problem structure (see Figure 15).

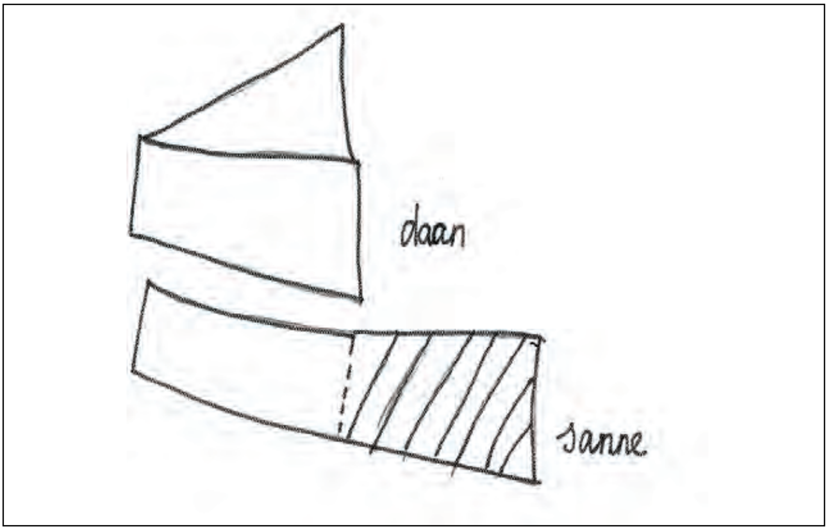


Figure 15. Lisa's incorrect visualization of Change 1 (Note: in the Dutch translation of the word problem test Mary and John were replaced by Sanne and Daan)

Also Change 2 was solved incorrectly by Lisa (Lisa: 4 + ? = 7 [3]; required: 3 + ? = 7 [4]). In this word problem only the last solution step was used, but incorrectly executed. Lisa seemed to have problems finding the 'unknown' variable. The same applies to Change 3 (Lisa: 5 + ? = 9 [4]; required: 4 + ? = 9 [5]).

Compare word problems

The following three word-problem items are included in the pre-test and post-test (between brackets the adjusted figures of the post-test):

Compare 1. Mary has 5 (4) marbles. John has 8 (9) marbles. How many marbles does John have more than Mary?
Compare 2. Mary has 3 (2) marbles. John has 4 (5) marbles more than Mary. How many marbles does John have?
Compare 3. Mary has 4 (5) marbles. She has 3 (2) marbles less than John. How many marbles does John have?

Hugo
Pre-test. The results of the pre-test showed that Hugo answered Compare 1 incorrectly (answer Hugo = 13; required answer = 3) His decision to add the two known figures, indicated that Hugo had difficulties comprehending the word problem text (Hugo: 5 + 8 = ? [13]; required: 5 + ? = 8 [3]). Compare 2 was, however, solved cor-

rectly (i.e., 3 + 4 = ? [7]). Also Compare 3 was correctly answered, despite of complex problem structure of this word problem (i.e., 4 + 3 = ? [3]). Although this word problem had almost the same problem structure as Compare 1, no comprehension or calculation errors were made by Hugo while solving Compare 3.

Post-test. The results of the post-test showed that Compare 1 was answered correctly (i.e., 4 + ? = 9 [5]). With regard to the solution strategies that were used in Compare 1, Hugo correctly visualized the problem structure, added a question mark in the right place in the visual-schematic representation, and correctly drew a circle around the 'unknown' variable. However, Compare 2 was solved incorrectly in the post-test (Hugo: 5 + ? = 7 [2]; required: 2 + 5 = ? [7]). Hugo seemed to have difficulties finding the 'unknown' variable. This is also reflected in the incorrect visual-schematic representation that he made (see Figure 16).

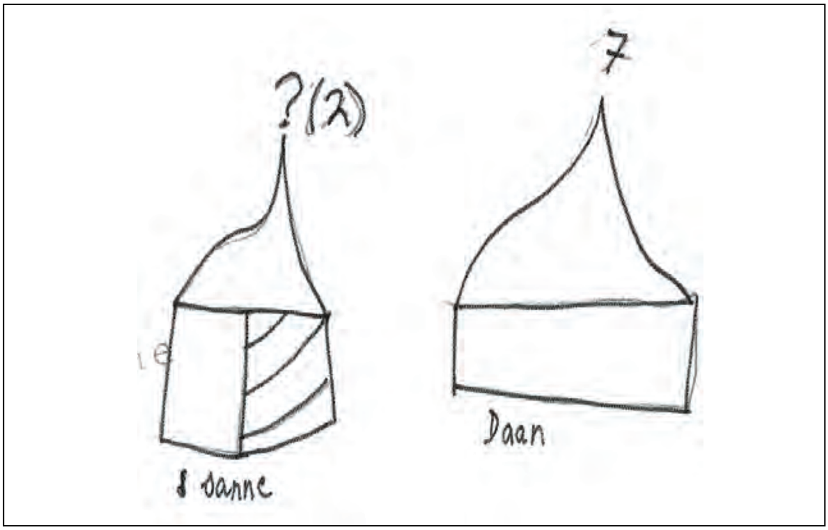


Figure 16. Hugo's incorrect visualization of Compare 2 (Note: in the Dutch translation of the word problem test Mary and John were replaced by Sanne and Daan)
word problem test Mary and John were replaced by Sanne and Daan)

Also Compare 3 was solved incorrectly (Hugo: 5 -2 = ? [3]; required:5 + 2 = ? [7]). Hugo apparently was distracted by the relational keyword 'less than', and performed a subtraction operation instead of the required addition operation. The step in which the problem structure was visualized was executed correctly. However, Hugo drew a circle around the wrong variable.

Peter

Pre-test. The results of the pre-test showed that Peter solved Compare 1 incorrectly (Peter: 5 – 3 = ? [2]; required: 5 + ? = 8 [3]). Peter had difficulties distilling the correct mathematical operation from the word problem text. Also Compare 2 was solved incorrectly (Peter: 2 x 2 = ? [4]). It seemed that Peter randomly performed a mathematical operation for this word problem. Compare 3 was solved incorrectly because Peter performed a subtraction operation (Peter: 4 -3 = ? [1]), instead of an addition operation (required: 4 + 3 = ? [7]). Peter was probably distracted by the relational keyword 'less than'.
Post-test. The results of the post-test showed that Compare 1 was solved correctly (i.e., 4 + ?= 9 [5]). Also Compare 2 and Compare 3 were solved correctly after the intervention period (Compare 2: 2 + 5= ? [7]; Compare 3: 5 + 2 = ? [7]). With regard to the solution strategies that were used in the post-test, Peter correctly visualized the problem structure, added a question mark in the right place in the visual-schematic representation, and correctly drew a circle around the 'unknown' variable in all Compare word problems.

Tim

Pre-test. The results of the pre-test showed that Tim gave no answer to Compare 1, because he apparently did not know the mathematical operation that had to be performed. The answer on Compare 2 (answer Tim = 8; required answer = 7) suggested that Tim recognized the structure of this word problem, but made a calculation error (Tim: 3 + 4 = ? [8]; required: 3 + 4 = ? [7]). The incorrect answer on Compare 3, however, did not reflect a calculation error, but rather a comprehension error (Tim: 4 -3 = ? [1]; required 4 + 3 = ? [7]). Tim apparently was distracted by the relational keyword 'less than', and performed a subtraction operation instead of the required addition operation. In Compare 2 and Compare 3 Tim only wrote down the mathematical operations that he performed and reported no other solution strategies.
Post-test. The results of the post-test showed that Compare 1 was answered correctly (i.e., 4 + ? = 9 [5]). Also Compare 2 and Compare 3 were solved correctly after the intervention period (Compare 2: 2 + 5 = ? [7]; Compare 3: 5 + 2 = ? [7]). With regard to the solution strategies that were used in the post-test, Tim only used the step in which a circle should be drawn around the 'unknown' variable. This step was executed correctly in all three Compare word problems.

Lisa

Pre-test. The results of the pre-test showed that Lisa answered Compare 1 incorrectly (answer Lisa = 13; required answer = 3) Her decision to add the two known figures, indicated that Lisa did not comprehend the word problem text (Lisa: 5 + 8 = ? [13]; required: 5 + ? = 8 [3]). Compare 2 was, however, solved correctly (i.e., 3 + 4 = ? [7]). Compare 3 was solved incorrectly; this was not caused by a comprehension error, but by a calculation error (Lisa: 4 + 3 = ? [5]; required: 4 + 3 = ? [7]).

Post-test. The results of the post-test showed that Compare 1 was solved correctly (i.e., 4 + ?= 9 [5]). With regard to the solution steps that were used in Compare 1, Lisa correctly visualized the problem structure, added a question mark in the right place in the visual-schematic representation, and correctly drew a circle around the 'unknown' variable. Also Compare 2 was solved correctly (i.e., 2 + 5 = ? [7]). Remarkably, Lisa did not use any solution steps while solving this word problem. The same applies to Compare 3 (i.e., 5 + 2 = ? [7]).

Conclusions

I

The results of the pre-test showed that Peter and Tim experienced difficulties solving all three types of word problems. As Hugo solved two out of three compare problems correctly, his difficulties lay mainly in solving combine and change problems. Lisa, on the other hand, had relatively less difficulty with solving change problems compared to combine and compare problems (see Table 3). The types of errors that were made in the pre-test can be distinguished as: (1) calculation errors, and (2) comprehension errors. The four research subjects made relatively less calculation errors. Incorrect answers to the word problems were mainly the result of comprehension errors. In several situations, the students reported the incorrect mathematical operation, had difficulties finding the required answer, or just performed an addition operation with the known figures in the word problem without carefully understanding and identifying the problem structure. While solving Compare 3 for example, students got distracted by the relational term 'less than', which falsely referred to a subtraction operation. With respect to the solution strategies used during the pre-test, all four students only wrote down the mathematical operations that they performed and reported no other solution strategies.

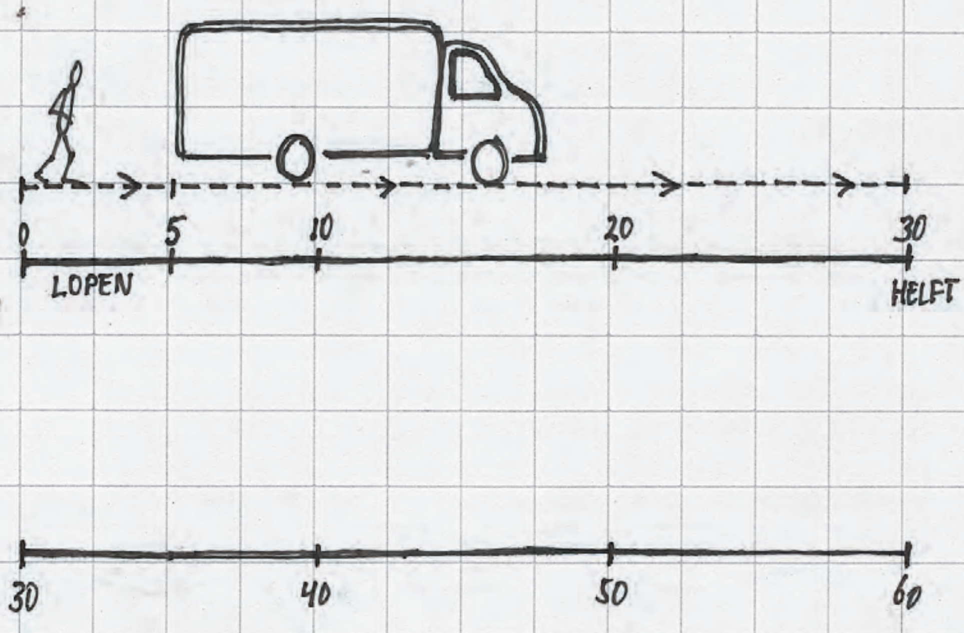
After the intervention period, Peter and Tim made no errors on the combine, change and compare problems. Although Hugo showed a better total performance in the post-test, he performed poorer on compare problems compared to the pre-test. The same applied to Lisa's performance on change problems. Therefore, we can only tentatively conclude that the difference between pre-test and post-test results can be ascribed to the word problem solving instruction. When we looked at the correct use of the solution steps of this instruction, we saw that Peter mastered all steps and executed them correctly. Apparently, this had a positive influence on his word problem solving performance. Tim, on the other hand, performed only one solution step (i.e., draw a circle on the required answer) in all the word problems during the post-test. The correct execution of this solution step seemed to help him solve all the word problems in the post-test correctly. That it is important to execute the solution steps correctly is also reflected by the performance of Hugo on the change and compare problems. Although Hugo executed all the solution steps, in some situations he had difficulties constructing the correct visual-schematic representation and drawing a circle on the 'unknown' variable. As a result he answered these word problems incorrectly. The findings with regard to the solution strategies used by Lisa showed that she did not master the solution steps for all types of word problems. In two types of word problems (Combine 1 and Compare 1), Lisa executed all the steps correctly. However, in the other type of word problems not all solution steps were executed, or some of the solution steps were executed incorrectly.

It should be noted that our findings do not imply that every student will benefit from a word problem solving instruction like the one that we investigated. Teachers should take the individual differences between students' educational needs into account in all circumstances. Moreover, the absence of a control group deters us from drawing strong conclusions concerning the effectiveness of the word problem solving instruction. Nevertheless, this feasibility study provides insights with regard to the varying ways in which a word problem solving instruction can influence the solution strategies and performances of students who perform poorly. It provides the basis for a more formal study in the future, conducted using a controlled design. The present study gives us clues about which methodological elements should be taken into account.

FINAL REMARKS

This article describes a feasibility study in which the performances of four second-grade students were investigated as individual cases. Therefore, its findings regarding the effectiveness of the word problem solving instruction used should be interpreted with caution. Nevertheless, the word problem solving instruction described in this study shows promise as a tool for teaching students to solve word problems. Future research is required to provide more insights concerning the effectiveness of the instruction on students' performances and use of strategy in solving word problems. This feasibility study marks an important starting point in the search for instructional programs that could be implemented in the educational practice of contemporary math approaches where word problem solving plays a prominent role.

REFERENCES



A hitchhiker set out on a journey of 60 miles. He walked the first 5 miles and then got a lift from a lorry driver. When the driver dropped him he still had half of his journey to travel. How far had he traveled in the lorry?

A —

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: the same of different constructs? *Psychological Bulletin, 131*, 30-60. doi: http://dx.doi.org/10.1037/0033-2909.131.1.30

Ahmad, A., Tarmizi, R. A., & Nawawi, M. (2010). Visual representations in mathematical word problem solving among form four students in Malacca. *Procedia Social and Behavioral Sciences, 8*, 356-361. doi: http://dx.doi.org/10.1016/j.sbspro.2010.12.050

Antoniou, P., & Kyriakides, L. (2013). A Dynamic Integrated Approach to teacher professional development: Impact and sustainability of the effects on improving teacher behaviour and student outcomes. *Teaching and Teacher Education*, *29*, 1-12. doi: 10.1016/j.tate.2012.08.001

B —

Barnes, H. (2005). The theory of Realistic Math Education as a theoretical framework for teaching low attainers mathematics. *Pythagoras, 61*, 42-57.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182. doi: http://dx.doi.org/10.1037/0022- 3514.51.6.1173

Battista, M. T. (1990) Spatial visualization and gender differences in high school geometry. *Journal for Research in Mathematics Education, 21*, 47-60. doi: http://www.jstor.org/stable/749456

Beentjes, I. (2008). *Ruimtelijke exploratie en de ontwikkeling van number sense* [Spatial exploration in the development of number sense] (Master thesis Orthopedagogiek, Utrecht University). Utrecht.

Beets, M. W., Flay, B. R., Vuchinich, S., Acock, A. C., Li, K.-K., & Allred, C. (2008). School climate and teachers' beliefs and attitudes associated with implementation of the positive action program: A diffusion of innovations model. *Prevention Science, 9*, 264-275. doi: 10.1007/s11121-008-0100-2

Beijaard, D. (2009). Leraar worden en leraar blijven: over de rol van identiteit in professioneel leren van beginnende docenten [Becoming a teacher and staying a teacher: about the role of

identity in professional development of beginning teachers]. Inaugural lecture. Eindhoven University of Technology.
Berends, I. E., & Van Lieshout, E. C. D. M. (2009). The effect of illustrations in arithmetic problem-solving: Effects of increased cognitive load. *Learning and Instruction, 19*, 345-353. doi: 10.1016/j.learninstruc.2009.06.012
Bernardo, A. B. I. (1999). Overcoming obstacles to understanding and solving word problems in mathematics. *Educational Psychology, 19,* 149-163. doi: 10.1080/0144341990190203
Biesta, G. (2007). Why "what works" won't work: Evidence-based practice and the democratic deficit in educational research. *Educational theory, 57*, 1-22. doi: 10.1111/j.1741-5446.2006.00241.x
Bitan-Friedlander, N., Dreyfus, A., & Milgrom, Z. (2004). Types of "teachers in training": the reactions of primary school science teachers when confronted with the task of implementing an innovation. *Teaching and Teacher Education*, *20*, 607-619. doi: 10.1016/j.tate.2004.06.007
Bjorklund, D. F., & Brown, R. D. (1998). Physical play and cognitive development: Integrating activity, cognition, and education. *Child Development*, *69*, 604-606. doi: 10.1111/j.1467-8624.1998.tb06229.x
Blatto-Vallee, G., Kelly, R. R., Gaustad, M. G., Porter, J., & Fonzi, J. (2007). Visual-spatial representation in mathematical problem solving by deaf and hearing students. *Journal of Deaf Studies and Deaf Education, 12*, 432-448. doi: 10.1093/deafed/enm022
Boonen, A. J. H., Van der Schoot, M., Van Wesel, F., De Vries, M. H., & Jolles, J. (2013). What underlies successful word problem solving? A path analysis in sixth grade students. *Contemporary Educational Psychology, 38*, 271-279. doi: http://dx.doi.org/10.1016/j.cedpsych.2013.05.001
Boonen, A. J. H., Van Wesel, F., Jolles, J., & Van der Schoot, M. (2014). The role of visual representation type, spatial ability, and reading comprehension in word problem solving: An item-level analysis in elementary school children. *International Journal of Educational Research, 68*, 15-26. doi: http://dx.doi.org/10.1016/j.ijer.2014.08.001
Booth, R. D. L., & Thomas, M. O. J. (1999). Visualization in

mathematics learning: Arithmetic problem solving and student difficulties. *Journal of Mathematical Behavior, 18*, 169-190. http://dx.doi.org/10.1016/S0732-3123(99)00027-9

Broekkamp, H., & Van Hout-Wolters, B. (2007). The gap between educational research and practice: A literature review, symposium, and questionnaire. *Educational Research and Evaluation*, *13*, 203-220. doi: 10.1080/13803610701626127

Brosnan, M. J. (1998). Spatial ability in children's play with Lego blocks. *Perceptual andMotor Skills, 87*, 19-28. doi: 10.2466/pms.1998.87.1.19

Bryant, D. J., & Tversky, B. (1999). Mental representations of perspective and spatial relations from diagrams and models. *Journal of Experimental Psychology, 25*, 137-156. doi: 10.1037/0278-7393.25.1.137

Bursal, M., & Paznokas, L. (2006). Mathematics anxiety and preservice elementary teachers' confidence to teach mathematics and science. *School Science & Mathematics, 106*, 173-180. doi: 10.1111/j.1949-8594.2006.tb18073.x

C —

Caldera, Y. M., Culp, A. M., O'Brian, M., Truglio, R. T., Alvarez, M., & Huston, A. (1999). Children's play preferences, construction play with blocks and visual spatial skills: Are they related? *International Journal of Behavioral Development, 23*, 855-872. doi: 10.1080/016502599383577

Campbell, J. I. D. (Ed.). (1992). *The nature and origins of mathematical skills.* Amsterdam: Elsevier Science Publishers.

Carpenter, T. P., Corbitt, M. K., Kepner, H. S., Lindquist, M. M., & Reys, R. E. (1981). National assessment. In E. Fennema (Ed.), *Mathematics education research; Implications for the 80's* (pp. 22- 38). Reston, VA: National Council of Teachers of Mathematics.

Caruso, D. A. (1993). Dimensions of quality in infants' exploratory behavior: relationships to problem solving ability. *Infant Behavior and Development, 16*, 441-454. doi: 10.1016/0163-6383(93)80003-Q

Casey, M. B., Andrews, N., Schindler, H., Kersh, J. E., Samper, A., & Copley, J. (2008). The development of spatial skills through interventions involving block building activities. *Cognition and Instruction, 26,* 269-309.

doi: 10.1080/07370000802177177

Casey, M. B., Nuttall, R. L., Benbow, C. P., & Pezaris, E. (1995). The influence of spatial ability on gender differences in mathematics college entrance test scores across diverse samples. *Developmental Psychology, 31*, 697-705. doi: 10.1037/0012-1649.31.4.697

Casey, M. B., Nuttall, R. L., & Pezaris, E. (1997). Mediators of gender differences in mathematics college entrance test scores: A comparison of spatial skills with internalized beliefs and anxieties. *Developmental Psychology, 33*, 669-680. doi: 10.1037/0012-1649.33.4.669

Casey, M. B., Nuttall, R. L., & Pezaris, E. (2001). Spatial-mechanical reasoning skills versus mathematics self-confidence as mediators of gender differences on mathematics subtests using cross-national gender-based items. *Journal for Research in Mathematics Education, 32*, 28-57. http://www.jstor.org/stable/749620

Casey, M. B., Pezaris, E., & Nuttall, R. L. (1992). Spatial ability as a predictor of math achievement: The importance of sex and handedness patterns. *Neuropsychologia, 30,* 35–45. doi: 10.1016/0028-3932(92)90012-B

CITO (2010). *Leerling- en onderwijsvolgsysteem, Begrijpend Lezengroep 7* [Reading comprehension test grade 5]. Arnhem: Cito.

CITO (2008). *Leerling- en onderwijsvolgsysteem, Rekenen-Wiskunde groep 7* [Mathematics test grade 5]. Arnhem: Cito.

Clark, H. H. (1969). Linguistic processes in deductive reasoning. *Psychological Review, 76*, 387–404. doi: 10.1037/h0027578

Clark, H. H., & Card, S. K. (1969). Role of semantics in remembering comparative sentences. *Journal of Experimental Psychology, 82*, 545–553. doi: http://dx.doi.org/10.1037/h0028370

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155–159. doi: 10.1037/0033-2909.112.1.155

Cummins, D. D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in solving word problems. *Cognitive Psychology, 20*, 405-438. doi: http://dx.doi.org/10.1016/0010-0285(88)90011-4

D —

d'Ailley, H. H., Simpson, J., & MacKinnon, G. E. (1997). Where should 'you' go in a math compare problem. *Journal of Educational Psychology, 89*, 562-567. doi: http://dx.doi.org/10.1037/0022-0663.89.3.562

De Corte, E., & Verschaffel, L. (1987). The effect of semantic structure on 1st-graders strategies for solving addition and subtraction word problems. *Journal of Research in Mathematics Education, 18*, 363-381. doi: http://www.jstor.org/stable/749085

De Corte, E., Verschaffel, L., & De Win, L. (1985). Influence of rewording verbal problems on children's problem representations and solutions. *Journal of Educational Psychology, 77*, 460-470. doi: 10.1037/0022-0663.77.4.460

De Corte, E., Verschaffel, L., & Pauwels, A. (1990). Influence of the semantic structureof word problems on second graders' eye movements. *Journal of Educational Psychology, 82*, 359–365. doi: 10.1037/0022-0663.82.2.359

Depaepe, F., De Corte, E., & Verschaffel, L. (2010a). Teachers' approaches towards word problem solving: Elaborating or restricting the problem context. *Teaching and Teacher Education, 26*, 152-160. doi: 10.1016/j.tate.2009.03.016

Dewolf, T., Van Dooren, W., Cimen, E. E., & Verschaffel, L. (2014). The impact of illustrations and warnings on solving mathematical word problems realistically. *Journal of Experimental Education, 82*, 103-120. doi: 10.1080/00220973.2012.745468

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 Mathematics Education, 39*, 405-418. doi: 10.1007/s11858-007-0043-2

E —

Edens, K., & Potter, E., (2008). How students unpack the structure of a word problem: Graphic representations and problem solving. *School Sciences and Mathematics, 108*, 184-196. doi: 10.1111/j.1949-8594.2008.tb17827.x

Ekstrom, R. B., French, J. W., Harman, H., & Derman, D. (1976). *Kit of factor-referenced cognitive tests.* Princeton, NJ: Educational Testing Service.

Elia, I., Van den Heuvel-Panhuizen, M., & Kovolou, A. (2009). Exploring strategy use and strategy flexibility in non-routine

problem solving in primary school high achievers in mathematics. *ZDM The International Journal of Mathematics Education, 41*, 605-618.
doi: 10.1007/s11858-009-0184-6

Evers, W. J. G., Brouwers, A., & Tomic, W. (2002). Burnout and self-efficacy: A study on teachers' beliefs when implementing an innovative educational system in the Netherlands. *British Journal of Educational Psychology, 72*, 227-243.
doi: 10.1348/000709902158865

F —

Fairchild, A. J. Mackinnon, D. P., Taborga, M. P., & Taylor, A. B. (2009). R2 effect-size measures for mediation analysis. *Behavior Research Methods, 41*, 486-498.
doi: 10.3758/BRM.41.2.486

French, P. L. (1979). Linguistic marking, strategy, and affect in syllogistic reasoning. *Journal of Psycholinguistic Research, 8*, 425–449.
doi: 10.1007/BF01067329

G —

Geary, D. C., Saults, S. J., Liu, F., & Hoard, M. K. (2000). Sex differences in spatial cognition, computational fluency, and arithmetic reasoning. *Journal of Experimental Child Psychology*, 337-353. doi:10.1006/jecp.2000.2594

Ghaith, G., & Yaghi, H. (1997). Relationships among experience, teacher efficacy, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education, 13*, 451-458.
doi: 10.1016/S0742-051X(96)00045-5

Giroux, J., & Ste-Marie, A. (2001). The solution of compare problems among first grade students. *European Journal of Psychology and Education, 16*, 141-161.
doi: 10.1007/BF03173022

Gravemeijer, K., & Doorman, M. (1999). Context problems in realistic mathematics education: A calculus course as an example. *Educational Studies in Mathematics, 39*, 111-129. doi: 10.1023/A:1003749919816

Green, S. B., & Salkind, N. J. (2008). *Using SPSS for Windows and Macintosh. Analyzing and Understanding Data (5th edition).* London: Pearson Education.

Grimshaw, G. M., Sitarenios, G., & Finegan, J. A. K. (1995). Mental rotation At 7 years-relations with prenatal testosterone levels and spatial play experiences. *Brain and cognition*, *29*, 85-100. doi: 10.1006/brcg.1995.1269

Guay, R. B., & McDaniel, E. D. (1977). The relationship between mathematics achievement and spatial abilities among elementary school children. *Journal for Research in Mathematics Education, 8*, 211-215. doi: http://www.jstor.org/stable/748522

Guoliang, Y., & Pangpang, Z. (2003). Visual-spatial representations and mathematical problem solving among mathematical learning disabilities. *Acta Psychologica Sinica*, *35*, 643-648.

H —

Haenggi, D., Kintsch, W., & Gernsbacher, M. A. (1995). Spatial situation models and text comprehension. *Discourse Processes, 19*, 173-199. doi: 10.1080/01638539509544913

Hajer, M. (1996). *Leren in een tweede taal. Interactie in een meertalige mavo-klas* [Learning a second language. Interaction in a multilingual classroom]. The Netherlands, Groningen: Wolters Noordhoff.

Han, S. S., & Weiss, B. (2005). Sustainability of teacher implementation of school-based mental health programs. *Journal of Abnormal Child Psychology, 33*, 665–679. doi:10.1007/s10802-005-7646-2

Hannon, B., & Daneman, M. (2001). A new tool for measuring and understanding individual differences in the component processes of reading comprehension. *Journal of Educational Psychology*, *93*, 103-128. doi: 10.1037/0022-0663.93.1.103

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in new millennium. *Communication Monographs, 76*, 408-420. doi: 10.1080/03637750903310360

Hegarty, M., & Kozhevnikov, M. (1999). Types of visual–spatial representations and mathematical problem solving. *Journal of Educational Psychology, 91*, 684–689. doi: http://dx.doi.org/10.1037/0022-0663.91.4.68

Hegarty, M., Mayer, R., E., & Green, C. E. (1992). Comprehension of arithmetic word problems: Evidence from students' eye fixations. *Journal of Educational Psychology, 84*, 76-84. doi: 10.1037/0022-0663.84.1.76

Hegarty, M., Mayer, R. E., & Monk, C. A. (1995). Comprehension of arithmetic word problems: A comparison of successful and unsuccessful problem solvers. *Journal of Educational Psychology, 87*, 18-32. doi: 10.1037/0022-0663.87.1.18

Hegarty, M., & Waller, D. A. (2005). Individual differences in spatial abilities (Eds.), *The Cambridge handbook of visuospatial thinking* (pp. 121-169). Cambridge: University Press.

Helwig, R., Rozek-Tedesco, M. A., Tindal, G., Heath, B., & Almond, P. J. (1999). Reading as an access to mathematics problem solving on multiple-choice tests for sixth-grade students. *The Journal of Educational Research, 93*, 113-125. doi:10.1080/00220679909597635

Hermans, R., Tondeur, J., Van Braak, J., & Valcke, M. (2008). The impact of primary school teachers' educational beliefs on the classroom use of computers. *Computers & Education, 51*, 1499-1509. doi: 10.1016/j.compedu.2008.02.001

Hickendorff, M. (2011). *Explanatory latent variable modeling of mathematical ability in primary school: Crossing the border between psychometrics and psychology.* (doctoral dissertation, Leiden University). The Netherlands, Leiden.

Hickendorff, M. (2013). The effects of presenting multidigit mathematics problems in a realistic context on sixth graders' problem solving. *Cognition and Instruction, 31*, 314-344. doi: 10.1080/07370008.2013.799167

Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction*, *26*, 430-511. doi: 10.1080/07370000802177235

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariancestructure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55. doi:10.1080/10705519909540118

I —

Isiksal, M., Curran, J. M., Koc, Y., & Askun, C. S. (2009). Mathematics anxiety and mathematical self-concept: Considerations in preparing elementary-school teachers. *Social Behavior and Personality, 37*, 631-644. doi: 10.2224/sbp.2009.37.5.631

J —

Janssen, J., Verhelst, N., Engelen, R., & Scheltens, F. (2010). *Wetenschappelijke Verantwoording papieren toetsen Rekenen-Wiskunde groep 3 tot en met 8* [Scientific justification of the mathematics test]. The Netherlands, Arnhem: Cito.

Jiminez, L., & Verschaffel, L. (2014). Development of children's solutions of non-standard arithmetic word problem solving. *Revista de Psicodidáctica, 2014, 19,* 93-123. doi: 10.1387/RevPsicodidact.7865

Jitendra, A.K. (2002). Teaching students math problem-solving through graphic representations. *Teaching Exceptional Children, 34*, 34-38.

Jitendra, A.K., DiPipi, C. M., & Perron-Jones, N. (2002). An exploratory study of schema-based word problem solving instruction for middle school students with learning disabilities: An emphasis on conceptual and procedural understanding. *The Journal of Special Education, 36*, 23–38. http://dx.doi.org/10.1177/00224669020360010301

Jitendra, A. K., George, M. P., Sood, S., & Price, K. (2010). Schema-based instruction: facilitating mathematical word problem solving for students with emotional and behavioral disorders. *Preventing School Failure, 54*, 145-151. doi: 10.1080/10459880903493104

Jitendra, A. K., Griffin, C. C., Haria, P., Leh, J., Adams, A., & Kaduvettoor, A. (2007). A comparison of single and multiple strategy instruction on third-grade students' mathematical problem solving. *Journal of Educational Psychology*, *99*, 115-127. doi: 10.1037/0022-0663.99.1.115

Jitendra, A., & Hoff, K. (1996). The effects of schema-based instruction on mathematical word-problem-solving performance of students with learning disabilities. *Journal of Learning Disabilities, 29*, 422-431. doi: 10.1177/002221949602900410

Jitendra, A. K., Petersen-Brown, S., Lein, A. E., Zaslofsky, A. F., Kunkel, A. K., Jung, P.-G., & Egan, A. M. (2013). Teaching mathematical word problem solving: The quality of evidence for strategy instruction priming the problem structure. *Journal of Learning Disabilities, XX*, 1-22. doi: 10.1177/0022219413487408

Jitendra, A. K., & Star, J. R. (2012). An exploratory study contrasting high- and low achieving students' percent word problem solving. *Learning and Individual Differences, 22*, 151-158. doi: 10.1016/j.lindif.2011.11.003

Jitendra, A. K., Star, J. R., Rodriguez, M., Lindell, M., & Someki, F. (2011). Improving students' proportional thinking using schema-based instruction. *Learning and Instruction, 21*, 731-745. doi: 10.1016/j.learninstruc.2011.04.002

Jitendra, A. K., Star, J. R., Starosta, K., Leh, J. M., Sood, S., Caskie, G., … Mack, T. R. (2009). Improving seventh grade students' learning of ratio and proportion: The role of schema-based instruction. *Contemporary Educational Psychology, 34*, 250-264. doi:10.1016/j.cedpsych.2009.06.001

K —

Kaufmann, S. B., (2007). Sex differences in mental rotation and spatial visualization ability: Can they be accounted for by differences in working memory capacity? *Intelligence, 35*, 211-223. doi: http://dx.doi.org/10.1016/j.intell.2006.07.009

Keith, T. Z., Reynolds, M. R., Patel, P. G., & Ridley, K. R. (2008). Sex differences in latent cognitive abilities ages 6 to 59: Evidence from the Woodcock Johnson III Tests of Cognitive Abilities. *Intelligence, 36*, 502–525. doi: http://dx.doi.org/10.1016/j.intell.2007.11.001

Kendeou, P., Papadopoulos, T. C., & Spanoudis, G. (2012). Processing demands of reading comprehension tests in young readers. *Learning and Instruction, 22*, 354-367. doi: http://dx.doi.org/10.1016/j.learninstruc.2012.02.001

Ketelaar, E., Beijaard, D., Boshuizen, H., & Den Brok, P. J. (2012). Teachers' positioning towards an educational innovation in the light of ownership, sense-making and agency. *Teaching and Teacher Education*, *28*, 273-282. doi: 10.1016/j.tate.2011.10.004

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction–integration model. *Psychological Review, 95*, 163–183. doi: 10.1037/0033-295X.95.2.163

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* Cambridge: Cambridge University Press.

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review, 92*, 109-129. doi: 10.1037/0033-295X.92.1.109

Kline, R. B. (2005). *Principles and practice of structural equation* *modeling (second edition)*. New York: The Guilford Press.

Koedinger, K. R., & Nathan, M. J. (2004). The real story about story problems: Effects of representations on quantitative reasoning. *The Journal of the Learning Sciences, 13*, 129-164. doi: 10.1207/s15327809jls1302_1

Könings, K. D., Brand-Gruwel, S., & van Merriënboer, J. J. (2007). Teachers' perspectives on innovations: Implications for educational design. *Teaching and Teacher Education*, *23*, 985-997. doi: 10.1016/j.tate.2006.06.004

Kozhevnikov, M., Hegarty, M., & Mayer, R. E. (2002). Revising the visualizer-verbalizer dimension: Evidence for two types of visualizers. *Cognition and Instruction, 20*, 47-77. doi: 10.1207/S1532690XCI2001_3

Krawec, J. L. (2010). *Problem representation and mathematical problem solving of students with varying abilities* (doctoral dissertation, University of Miami). Miami.

Krawec, J. L. (2012). Problem representation and mathematical problem solving of students of varying math ability. *Journal of Learning Disabilities, X*, 1-13. doi: 10.1177/0022219412436976

Krawec, J. L., Huang, J., Montague, M., Kressler, B., & Melia de Alba, A. (2013). The effects of cognitive strategy instruction on knowledge of math problem-solving processes of middle school students with learning disabilities. *Learning Disability Quarterly, 36*, 80-92. doi: 10.1177/0731948712463368

Krippendorff, K. (2004). Reliability in content analysis: Some common misconceptions and recommendations. *Human Communication Research, 30*, 411-433. doi: http://dx.doi.org/10.1111/j.1468-2958.2004.tb00738.x

Kyriakides, L., Christoforou, C., & Charalambous, C. Y. (2013). What matters for student learning outcomes: A meta-analysis of studies exploring factors of effective teaching. *Teaching and Teacher Education*, *36*, 143-152. doi: 10.1016/j.tate.2013.07.010

Kyriakides, L., & Creemers, B. P. (2008). A longitudinal study on the stability over time of school and teacher effects on student outcomes. *Oxford Review of Education*, *34*, 521-545. doi: 10.1080/03054980701782064

L —

Lean, G., & Clements, M. A. K. (1981). Spatial ability, visual imagery,

and mathematical performance. *Educational Studies in Mathematics, 12*, 267-299.
doi: 10.1007/BF00311060

Lee, K., Ng, S.-W., Ng, E.-L., & Lim, Z.-Y. (2004). Working memory and literacy as predictors of performance on algebraic word problems. *Journal of Experimental Child Psychology, 89*, 140-158.
doi: http://dx.doi.org/10.1016/j.jecp.2004.07.001

Lee, K., Ng, E. L., & Ng, S. F. (2009). The contributions of working memory and executive functioning to problem representation and solution generation in algebraic word problems. *Journal of Educational Psychology, 101*, 373-387.
doi: 10.1037/a0013843

Leutner, D., Leopold, C., & Sumfleth, E. (2009). Cognitive load and science text comprehension: Effects of drawing and mentally imagining text content. *Computers in Human Behavior, 25*, 284-289.
doi:10.1016/j.chb.2008.12.010

Levine, S. C., Ratliff, K. R., Huttenlocher, J., & Cannon, J. (2012). Early puzzle play: a predictor of preschoolers' spatial transformation skill. *Developmental Psychology, 48*, 530-542.
doi: 10.1037/a0025913

Lewis, A. B., & Mayer, R. E. (1987). Students' miscomprehension of relational statements in arithmetic word problems. *Journal of Educational Psychology, 79*, 363-371.
doi: 10.1037/0022-0663.79.4.363

Lichtman, A. J. (1974). Correlation, regression and the ecological fallacy: A critique. *Journal of Interdisciplinary History, IV*, 417-433.
doi: http://www.jstor.org/stable/202485

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of gender differences in spatial abilities: a meta-analysis. *Child Development, 56,* 1479-1498.
doi: http://www.jstor.org/stable/1130467

M —

Marzocchi, G. M., Lucangeli, D., De Meo, T., Fini, F., & Cornoldi, C. (2002). The disturbing effect of irrelevant information on arithmetic problem solving in inattentive children. *Developmental Neuropsychology, 21*, 73-92.
doi: 10.1207/S15326942DN2101_4

Mayer, R. E. (1985). Mathematical ability. In R. J. Sternberg (Ed.), *Human abilities: An information processing approach* (pp. 127-150). San Francisco: Freeman.

Mayer, R.E. (1998). Cognitive, metacognitive, and motivational aspects of problem solving. *Instructional Science 26*, 49–63. doi: 10.1023/A:1003088013286

McGee, M. G. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological Bulletin, 86*, 889-918. doi: http://dx.doi.org/10.1037/0033-2909.86.5.889

McIntyre, D. (2005). Bridging the gap between research and practice. *Cambridge Journal of Education*, *35*, 357-382. doi: 10.1080/03057640500319065

Mitchell, E. (1973). The learning of sex roles through toys and books: a woman's view. *Young Children, 118,* 226-231.

Montague, M. (2003). *Solve It! A practical approach to teaching mathematical problem solving skills.* VA: Exceptional Innovations, Inc.

Montague, M., & Applegate, B. (2000). Middle school students' perceptions, persistence, and performance in mathematical problem solving. *Learning Disability Quarterly, 23,* 215-227. doi: http://www.jstor.org/stable/1511165

Montague, M., Enders, C., & Dietz, S. (2011). Effects of cognitive strategy instruction on math problem solving of middle school students with learning disabilities. *Learning Disability Quarterly, 34*, 262-272. doi: 10.1177/0731948711421762

Montague, M., Warger, C., & Morgan, T. H. (2000). *Solve It!* strategy instruction to improve mathematical problem solving. *Learning Disabilities Research & Practice, 15*, 110-116. doi: 10.1207/SLDRP1502_7

Moreno, R., Ozogul, G. & Reisslein, M. (2011). Teaching with concrete and abstract visual representations: Effects on students‘ problem solving, problem representations, and learning perceptions. *Journal of Educational Psychology, 103*, 32–47. doi: 10.1037/a0021995

Muthén, L. K., & Muthén, B. (2006). *Mplus user's guide* (Version 4). Los Angeles, CA: Muthén & Muthén.

N —

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231-259.

doi: http://dx.doi.org/10.1037/0033-295X.84.3.231

Nutley, S. M., Walter, I., & Davies, H. T. (2007). *Using evidence: How research can inform public services*. Policy Press.

O —

Orde, B. J. (1997). Drawing as visual-perceptual and spatial ability training. *In Proceedings of* selected research and development presentations at the 1997 National Convention of the Association for Educational Communications and Technology.

P —

Pantziara, M., Gagatsis, A., & Elia, I. (2009). Using diagrams as tools for the solution of non-routine mathematical problems. *Educational Studies in Mathematics, 72*, 39-60. doi: 10.1007/s10649-009-9181-5

Pape, S. J. (2003). Compare word problems: Consistency hypothesis revisited. *Contemporary Educational Psychology, 28*, 396-421. doi: http://dx.doi.org/10.1016/S0361-476X(02)00046-2

Pape, S. J. (2004). Middle school children's problem-solving behavior: A cognitive analysis from a reading comprehension perspective. *Journal for Research in Mathematics Education, 35,* 187–219. doi: http://www.jstor.org/stable/30034912

Penuel, W. R., Fishman, B. J., Yamaguchi, R., & Gallagher, L. P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal*, *44*, 921-958. doi: 10.3102/0002831207308221

Phillips, C. E., Jarrold, C., Baddeley, A. D., Grant, J., & Karmiloff-Smith, A. (2004). Comprehension of spatial language terms in Williams Syndrome: Evidence for an interaction between strength and weakness. *Cortex, 40*, 85-101. doi: 10.1016/S0010-9452(08)70922-5

Pierce, C.A., Block, R.A., & Aguinis, H. (2004). Cautionary note on reporting eta-squared values from multifactor anova designs. *Educational and Psychological Measurement, 64*, 916-924. doi: 10.1177/0013164404264848

Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (2003). Cognitive load in reading a foreign language text with multimedia aids and the influence on verbal and spatial abilities. *Computers in Human Behavior, 19*, 221-243. doi: 10.1037/0022-0663.90.1.25

Pomerleau, A., Malcuit, G., & Seguin, R. (1990). Five-month-old girls' and boys' exploratory behaviors in the present of familiar and unfamiliar toys. *The Journal of Genetic Psychology, 153*, 47-61. doi: 10.1080/00221325.1992.10753701

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*, 879-891. doi: 10.3758/BRM.40.3.879

Prenger, J. (2005). *Taal telt! Een onderzoek naar de rol van taalvaardigheid en tekstbegrip in het realistische rekenonderwijs. [Language counts! A study into the role of linguistic skill and text comprehension in realistic mathematics education].* Doctoral dissertation, University of Groningen, The Netherlands.

Presmeg, N. C. (1997). Generalization using imagery in mathematics. In L. D. English (Ed.), *Mathematical reasoning: Analogies, metaphors, and images* (pp. 299–312). Mahwah, NJ: Erlbaum.

Q —

Quaiser-Pohl, C. (2003). The Mental Cutting test 'Schnitte" and the Picture Rotation test – Two new measures to assess spatial ability. *International Journal of Testing, 3*, 219-231. doi: 10.1207/S15327574IJT0303_2

R —

Rasmussen, C. L., & King, K. D. (2000). Locating starting points in differential equations: a realistic mathematics education approach. *International Journal of Mathematical Education in Science and Technology, 31*, 161-172. doi: 10.1080/002073900287219

Riley, M. S., Greene, J. G., & Heller, J. I. (1983). *Development of children's problem solving ability in arithmetic.* In H. P. Ginsberg (Ed.), The development of mathematical thinking. New York: Academic Press.

Robinson, W.S. (2009). Ecological correlations and the behavior of individuals. *International Journal of Epidemiology, 38*, 337-341. doi: 10.1093/ije/dyn357

Rogers, E. M. (2003). *Diffusion of innovations (5th ed.).* New York: Free Press.

Ruijssenaars, A. J. J. M., Van Luit, J. E. H., & Van Lieshout, E. C. D. M. (Eds.). (2004). *Rekenproblemen en dyscalculie* [Arithmetic problems and dyscalculia]. Rotterdam, Netherlands: Lemniscaat.

S —

Sanchez, C. A. & Wiley, J. (2006). An examination of the seductive details effect in terms of working memory capacity. *Memory & Cognition, 34*, 344-355. doi: 10.3758/BF03193412

Schnotz, W., & Kürschner, C. (2008). External and internal representations in the acquisition and use of knowledge: visualization effects on mental model construction. Instructional Science, 36, 175-190. doi: 10.1007/s11251-007-9029-2

Schoenfeld, A. H. (1992*). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics.* D. Grouws (Ed.), Handbook of research on mathematics teaching and learning, MacMillan, New York (1992), 334–370.

Scholten, F. (2008). *Geslachtsverschillen in spelgedrag tussen jongens en meisjes.* [Sex differences in play behavior] (Bachelor thesis Kinder- en Jeugdpsychologie, University of Tilburg). Tilburg.

Schoppek, W., & Tulis, M. (2010). Enhancing arithmetic and word-problem solving skills efficiently by individualized computer-assisted practice. *The Journal of Educational Research*, *103*, 239-252. doi: 10.1080/00220670903382962

Schriefers, H. (1990). Lexical and conceptual factors in the naming of relations. *Cognitive Psychology, 22*, 111–142. doi: 10.1016/0010-0285(90)90005-O

Schumacher, R. F., & Fuchs, L. S. (2012). Does understanding relational terminology mediate effects of intervention on compare word problem? *Journal of Experimental Child Psychology, 111*, 607-628. doi: 10.1016/j.jecp.2011.12.001

Serbin, L. A., & Connor, J. M. (1979). Sex typing of children: play preference and patterns of cognitive performance. *Journal of Genetic Psychology, 134,* 315-316. doi: 10.1080/00221325.1979.10534065

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General, 125*, 4-27. doi: 10.1037/0096-3445.125.1.4

Sharma, U., Loreman, T., & Forlin, C. (2012). Measuring teacher efficacy to implement inclusive practices. *Journal of Research in Special Educational Needs, 12*, 12-21. doi: 10.1111/j.1471-3802.2011.01200.x

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and non-experimental studies: New procedures and recommendations. *Psychological Methods, 7*, 422-445.

Silver, E. A., & Cai, J. (1996). An analysis of arithmetic problem posing by middle school students. *Journal for Research in Mathematics Education, 27*, 521-539. doi: http://www.jstor.org/stable/749846

Simpson, E. H. (1949). Measurement of diversity. *Nature*.

Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational researcher*, *31*, 15-21. doi: 10.3102/0013189X031007015

Søvik, N., Frostrad, P., & Heggberget, M. (1999). The relation between reading comprehension and task-specific strategies used in arithmetical word problems. *Scandinavian Journal of Educational Research, 43*, 371-398. doi:10.1080/0031383990430403

Sprafkin, C., Serbin, L. A., Denier, C., & Connor, J. M. (1983). Sex-differentiated play: Cognitive consequences and early interventions. In MB Liss (Red), *Social and cognitive skills: Sex roles and child's play*. New York: Academic Press.

Stevens, R. J. (2004). Why do educational innovations come and go? What do we know? What can we do?. *Teaching and Teacher Education*, *20*, 389-396. doi: 10.1016/j.tate.2004.02.011

Swanson, H. L., Lussier, C. M., & Orosco, M. J. (2013). Cognitive strategies, working memory, and growth in word problem solving in children with math difficulties. *Journal of Learning Disabilities, XX*, 1-20. doi: 10.1177/0022219413498771

Swars, S. L., Daane, C. J., & Giessen, J. (2006). Mathematics anxiety and mathematics teacher efficacy: What is the relationship in elementary preservice teachers? *School Science and Mathematics, 106*, 306-315. doi: 10.1111/j.1949-8594.2006.tb17921.x

T —

Tabachnick, B. G., & Fidell, L. S. (2006). *Using multivariate statistic,* *fifth edition*. Boston: Allyn & Bacon.

Thevenot, C. (2010). Arithmetic word problem solving: Evidence for the construction of a mental model. *Acta Psychologica, 133*, 90-95. doi: http://dx.doi.org/10.1016/j.actpsy.2009.10.004

Thevenot, C., & Oakhill, J. (2005). The strategic use of alternative representation in arithmetic word problem solving. *Quarterly Journal of Experimental Psychology – A, 58*, 1311-1323. doi:10.1080/02724980443000593

Thevenot, C., & Oakhill, J. (2006). Representations and strategies for solving dynamic and static arithmetic word problems: The role of working memory capacities. *European Journal of Cognitive Psychology, 18*, 756-775. doi: 10.1080/09541440500412270

Timmermans, R. E., Van Lieshout, E. C. D. M., & Verhoeven, L. (2007). Gender related effects of contemporary math instruction for low performers on problem-solving behavior. *Learning and Instruction, 17*, 42-54. doi: http://dx.doi.org/10.1016/j.learninstruc.2006.11.005

Tolar, T. D., Fuchs, L., Cirino, P. T., Fuchs, D., Hamlett, C. L., & Fletcher, J. M. (2012). Predicting development of mathematical word problem solving across the intermediate grades. *Journal of Educational Psychology, 104*, 1083-1093. doi: 10.1037/a0029020

Tracy, D. M. (1987). Toys, spatial ability and science and mathematics achievement: Are they related? *Sex Roles, 17*, 115-138. doi: 10.1007/BF00287620

V —

Van den Boer, C. (2003). *Als je begrijpt wat ik bedoel. Een zoektocht naar verklaringen voor achterblijvende prestaties van allochtone leerlingen in het wiskundeonderwijs* [If you see what I mean. A quest for an explanation of the lower achievement levels of minority students in mathematics education]. Utrecht, CD-ß Press.

Van den Heuvel-Panhuizen, M. (2003). The didactical use of models in realistic mathematics education: An example from a longitudinal trajectory on percentage. *Educational Studies in Mathematics, 54*, 9-35. doi: 10.1023/B:EDUC.0000005212.03219.dc

Van den Heuvel-Panhuizen, M. (2005). The role of contexts in assessment problems in mathematics. *For the learning of*

Mathematics, 25, 2-9.
doi: http://www.jstor.org/stable/40248489

Vanderlinde, R., & Van Braak, J. (2010). The gap between educational research and practice: views of teachers, school leaders, intermediaries and researchers. *British Educational Research Journal*, *36*, 299-316.
doi: 10.1080/01411920902919257

Van der Schoot, M., Bakker-Arkema, A. H., Horsley, T. M., & Van Lieshout, E. C. D. M. (2009). The consistency effect depends on markedness in less successful but not successful problem solvers: An eye movement study in primary school children. *Contemporary Educational Psychology, 34*, 58-66.
doi: http://dx.doi.org/10.1016/j.cedpsych.2008.07.002

Van der Veen, I., Smeets, E., & Derriks, M. (2010). Children with special educational needs in the Netherlands: Number, characteristics and school career. *Educational Research, 52, 15-43.*
doi: 10.1080/00131881003588147

Van Dijk, I. M. A. W., Van Oers, H. J. M., & Terwel, J. (2003). Providing or designing? Constructing models in primary math education. *Learning and Instruction, 13*, 53-72.
doi: http://dx.doi.org/10.1016/S0959-4752(01)00037-8

Van Dijk, I. M. A. W., Van Oers, B., Terwel, J., & Van den Eeden, P. (2003). Strategic learning in primary mathematics education: Effects of an experimental program in modeling. *Educational Research and Evaluation, 9*, 161-187.
doi: http://dx.doi.org/10.1076/edre.9.2.161.14213

Van Eerde, H. A. A. (2009). Rekenen-wiskunde en taal: Een didactisch duo [Arithmetic and language: a didactical duo]. *Panamapost. Reken-wiskunde onderwijs:onderzoek, ontwikkeling en praktijk, 28*, 19-32.

Van Garderen, D. (2006). Spatial visualization, visual imagery, and mathematical problem solving of students with varying abilities. *Journal of Learning Disabilities, 39*, 496-506.
doi: 10.1177/00222194060390060201

Van Garderen, D., & Montague, M. (2003). Visual-spatial representation, mathematical problem solving, and students of varying abilities. *Learning Disabilities Research & Practice, 18*, 246-254.
doi: 10.1111/1540-5826.00079

Van Meter, P. (2001). Drawing construction as a strategy for learning from text. *Journal of Educational Psychology, 2001*, 93, 129-140.
doi: 10.1037/0022-0663.93.1.129

Van Meter, P., & Garner, J. (2005). The promise and practice of learner-generated drawing: literature review and synthesis. *Educational Psychology Review, 17*, 285-325. doi: 10.1007/s10648-005-8136-3

Velez, M. C., Silver, D., & Tremaine, M. (2005). Understanding visualization through spatial ability differences. *Visualization,VIS 05*, 511-518. doi: 10.1109/VISUAL.2005.1532836

Verschaffel, L. (1994). Using retelling data to study elementary school children's representations and solutions of compare problems. *Journal of Research in Mathematics Education, 25*, 141-165. doi: http://www.jstor.org/stable/749506

Verschaffel, L., De Corte, E., & Pauwels, A. (1992). Solving compare problems: An eye movement test of Lewis and Mayer's consistency hypothesis. *Journal of Educational Psychology, 84*, 85-94. doi: 10.1037/0022-0663.84.1.85

Verschaffel, L., Greer, B., & De Corte, E. (2000). *Making sense of word problems.* Lisse: Swets & Zeitlinger B.V.

Vescio, V., Ross, D., & Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and teacher education*, *24*, 80-91. doi: 10.1016/j.tate.2007.01.004

Vilenius-Tuohimaa, P. M., Aunola, K., & Nurmi, J.-E. (2008). The association between mathematical word problems and reading comprehension. *Educational Psychology, 28*, 409-426. doi:10.1080/01443410701708228

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117*, 25-270. doi: http://dx.doi.org/10.1037/0033-2909.117.2.250

W —

Wassenberg, R. (2007). *Differential cognitive development: A neuropsychological approach.* (doctoral dissertation, Maastricht University). Maastricht, The Netherlands.

Wassenberg, R., Hurks, P. P. M., Hendriksen, J. G. M., Feron, F. J. M., Meijs, C. J. C., Vles, J. S. H., & Jolles, J. (2008). Age-related improvement in complex language comprehension: Results

of a cross-sectional study with 361 children aged 5 to 15. *Journal of Clinical and Experimental Neuropsychology, 30*, 435-448.
doi: 10.1080/13803390701523109

Webb, D. C., Van der Kooij, H., & Geist, M. R. (2011). Design research in the Netherlands: Introducing logarithms using Realistic Mathematics Education. *Journal of Mathematics Education at Teachers College, 2*, 47-52.

Weekers, A., Groenen, I., Kleintjes, F. & Feenstra, H. (2011). *Wetenschappelijke Verantwoording papieren toetsen Begrijpend lezen voor groep 7 en 8* [Scientific justification of the reading comprehension test]. Arnhem: Cito.

Wolfgang, C. H., Stannard, L. L., & Jones, I. (2001). Block play performance among pre-schoolers as a predictor of later school achievement in mathematics. *Journal of Research in Childhood Education, 15*, 173-180.
doi: 10.1080/02568540109594958

Wolfgang, C. H., Stannard, L. L., & Jones, I. (2003). Advanced constructional play with LEGOs among preschoolers as a predictor of later school achievement in mathematics. *Early Child Development and Care, 173*, 467-475.
doi: 10.1080/0300443032000088212

Woolderink, S. (2009). *Zicht op Ouderschap.*[Perspectives on parenthood] (Master thesis Orthopedagogiek, VU University Amsterdam). Amsterdam.

Z —

Zhu, Z. (2007). Gender differences in mathematical problem solving patterns: A review of literature. *International Educational Journal, 8*, 187-203.
doi: http://iej.com.au

SUMMARY

Two objectives were examined in the research presented in this thesis. The first objective was to examine the extent to which students use different types of visual representations, and the role that spatial and semantic-linguistic skills play in the solving of routine and non-routine word problems in early (second) and later (sixth) grades of elementary school. This objective was the focus of the studies described in chapters 2 to 6. The second objective was to investigate how teachers implemented an innovative instructional approach – based on the didactical use of visual representations – in their own classroom teaching practice. This was the focus of the study described in chapter 7.

Students' difficulties with solving mathematical word problems are widely recognized by both researchers and teachers. The difficulties experienced by many students often do not rise from their inability to execute computations, but from difficulties with understanding the problem text. Two component processes, namely the production of visual-schematic representations and relational processing (i.e., deriving the correct relations between solution-relevant elements of the word problem text) and their underlying basic abilities play an important role in the successful understanding of word problems. Previous studies generally examined these component processes separately from each other by distinguishing a visual-spatial (the production of visual-schematic representations and spatial ability) and semantic-linguistic (relational processing and reading comprehension) processing domain. The study described in ***Chapter 2*** is the first study to examine the two component processes and basic abilities in one hypothesized model. On the basis of the findings of this study we can conclude that these component processes and basic abilities explain 49% of the variance of students' word problem solving performance. Moreover, both processing domains proved important and explained unique variance; 21% of the relation between spatial ability and word problem solving performance was explained by the production of visual-schematic representations, 34% of the relation between reading comprehension and word problem solving performance was explained by relational processing. On the basis of the path analyses, it appeared that the component processes are parallel processes which aid the successful comprehension of word problems in association with each other.

In spite of the proven importance of both visual representation strategies and semantic-linguistic skills, contemporary realistic math approaches tend to pay limited attention to the training of semantic-linguistic skills during word problem solving instruction. To test this assumption, we designed a study (see ***Chapter 3***) in which we not only manipulated the extent to which a sophisticated visual representation strategy was required (by distinguishing consistent and inconsistent word problems), but also varied the semantic complexity of the word problems by using highly semantic complex marked, and low semantic complex unmarked relation terms. Research showed that semantic-linguistic skills (i.e., reading comprehension) are necessary to overcome the semantic complexity of marked relational terms. In this study we classified students as successful and unsuccessful on the basis of their performance on an independent RME-specific mathematics test. The most important finding of this study, namely that successful word problem solvers in the RME curriculum had substantial difficulties in solving marked inconsistent word problems, substantiated the assumption that RME apparently pays little attention to the training of semantic-linguistic skills. It is important to start developing such skills early in elementary school, as word problems get semantically more complex as students progress in their educational career, for example when making the transition from elementary to secondary education.

The study reported in ***Chapter 4*** examined the importance of different types of visual representations, spatial ability and reading comprehension in word problem solving from *an item-level approach* rather than from the test-level approach that was often used in previous studies. This change in statistical modeling provided a more thorough and sophisticated understanding of the role of representation, spatial and reading comprehension skills in word problem solving. This item-level approach showed that the distinction between pictorial and visual-schematic representations made by previous studies is too narrow. To be more specific, we demonstrated that only the production of accurate visual-schematic representations was more frequently associated with a correct than with an incorrect answer to a word problem. Accurate, in contrast to inaccurate, visual-schematic representations contain a complete and coherent image of the problem situation, including the correct relations between the word problem's key variables. Accurate visual-schematic representations increased the chance of solving a word problem correctly by almost six times. In contrast, the production of inaccurate visual-schematic representations and pictorial representation decreased the chance of word problem solving success, respectively 2.94 and 2.78 times. As pictorial representations merely concern images of the visual appearance of objects or persons described in the word problem, they probably took the problem solvers' attention away from constructing a coherent model of the word problem, including the appropriate relations between the solution-relevant elements contained in it. Although inaccurate visual-schematic representations do include these relations, they are either incorrectly drawn or missing. As a consequence, this type of representation may have put problem solvers on the wrong track when solving a word problem.

Furthermore, besides contributing to a better understanding of the effects of the type of visual representations on the chance of word problem solving success, in this chapter we tried to reproduce, at the item-level, the findings of previous studies using a test-level approach concerning the importance of spatial ability and reading comprehension in word problem solving. In line with these earlier findings, the current study showed that spatial ability was a significant and relevant basic ability which increased the chance of solving a word problem successfully. However, our findings showed that the extent to which reading comprehension skills increase the chance of problem solving success is very limited. The results of the logistic regression analyses showed that although reading comprehension was a significant predictor in the model (due to the large number of items involved), the relevancy of its contribution was negligible (i.e., reading comprehension increased the chance of problem solving success only 1.02 times). Our item-level finding that reading comprehension was not a relevant factor contradicted the test-level findings from this study (r = .45), as well as previous studies demonstrating that reading comprehension and word problem solving performance were related. In other words, a relation between reading comprehension and word problem solving found at the test level does not imply that reading comprehension positively affects the chance of problem solving success at the item level.

In ***Chapter 5*** the specific relation between constructive play, spatial ability and word problem solving performance was examined. In previous studies, the relation between constructive play and spatial ability, and between spatial ability and mathematical word problem solving performance was reported. The relation between constructive play and mathematical word problem solving had, however, not been established yet. The findings of our study showed that spatial ability acted as a partial mediator in the relation between these two variables. This implied that children who had frequently engaged in constructive play in their past had better spatial skills and, as a result, showed a higher performance on mathematical word problems. The variables of this study (i.e., constructive play, spatial ability and sex) explained 38.16% of the variance in students' word problem solving performance. Furthermore, 31.58% of the relation between constructive play and mathematical word problem solving performance was explained by spatial ability.

While the first chapters of this thesis focus on the visual strategies, solution processes and performances on non-routine word problems of students in higher grades of elementary school (i.e., grade 6), the study described in ***Chapter 6*** was focused on second grade students' performances on routine word problems. The findings of this study showed that second grade students made more errors on compare word problems than on combine and change word problems, which was in line with research performed more than 25 years ago. Rather than the existence of a consistency effect (second grade students performed equally on inconsistent and consistent compare problems), a more general difficulty in processing relational terms (like 'more than' and 'less than') seemed to be a more plausible explanation of the finding. Second graders apparently did no (yet) have the knowledge to comprehend and process the linguistic input of compare problems and recall the appropriate problem structure. So, these students might have difficulties understanding the fact that the quantitative difference between the same sets could be expressed in parallel ways with both the terms more and fewer.

The difficulties that students experience with the comprehension and solution of word problems can be traced back to the role of the teacher in the word problem solving process. The findings of ***Chapter 7*** showed that most teachers were able to construct mathematical representations (e.g., a proportion table) and that they offered these representations to their students mainly in the solution phase of the word problem solving process. Like their students, several teachers experienced difficulties using accurate visual-schematic representations. Some teachers often used pictorial and inaccurate visual-schematic representations during the comprehension process of word problem solving. The findings of the study in Chapter 4, however, showed that these visual representation types did not increase the chance of solving a word problem correctly. The findings of our study in Chapter 7 revealed that when teachers did use accurate visual-schematic representations, in most cases a *bar model* was used. Other forms of visual-schematic representations like *number lines*, *pie charts* or *own constructions* were used only to a limited extent. In general, the level of diversity of forms of visual representations used by teachers was low. Moreover, the visual representations that teachers used did not always suit the problem characteristics and/or meet the individual needs of the students.

On the basis of the findings of studies described in this thesis, the following recommendations for teacher professionalization and teacher training could be made:

School teachers should be competent at constructing accurate visual-schematic representations as an aid in word problem solving, and the development of this competence should have a prominent place in the math curriculum of regular classrooms. (Student)teachers should have knowledge about the purpose of accurate visual-schematic representations and be trained in the construction and proper use of these types of visual representations.

Teacher professionalization and training should focus on the correct use of different forms of visual-schematic representations while solving a word problem. (Student) teachers should learn how to construct *number lines, pie charts and own construction* and other appropriate forms of visual-schematic representations.

There should be a particular emphasis on the construction process of these types of visual representation. Teachers should be able to use the construction process in a transparent, correct, and complete manner. (Student) teachers should learn to make their reasoning transparent by explaining which elements of the problem should be represented, and how the representation can be used to solve the problem. This reasoning process should also be correct as well as complete.

Finally, teacher training should pay particular attention to teaching how to identify the characteristics of word problems. (Student) teachers should know the distinction between routine and non-routine word problems, and the role that accurate visual-schematic representations play in these word problem types. Namely, in routine word problems (like combine, change and compare problems) the use of only one type of visual representation can suffice, because the problem structure of each of these types of word problems is identical. However, the problem structure of non-routine word problems varies, which makes it inappropriate to offer only one kind of accurate visual-schematic representation. Hence, (student) teachers' should learn to use visual-schematic representations in a way that is both diverse (i.e., demonstrating a varied use of visual representations) and flexible (i.e., offering different visual representations to solve one word problem). Moreover, these visual representations should be functional and suit the specific characteristics of the word problem (e.g., the use of a pie chart while solving word problems involving percentages).

The recommendations listed above provide interesting aspects for further research about the importance of visual representations in word problem solving. Based on the findings of the research presented in this thesis the focus of future studies should initially be on teachers' own competence and didactical use of visual representations during word problem solving instruction. Once teachers have more knowledge about the importance of visualization in the word problem solving process, they can use this knowledge to help their students successfully overcome the difficulties that they are experiencing.

Finally, the feasibility study reported in **Appendix I** examined four second-grade students who were less successful word problem solvers. These students received protocolled instruction during a five-week intervention period. The effectiveness of the word problem solving instruction was reported by comparing students' performances on the combine, change and compare problems before and after the intervention period, as well as by examining whether they executed the solution steps of the instruction correctly. The results of the pre- and post-test comparison showed that the total word problem solving performance of all four students had improved. However, this improvement was not always visible in all three types of word problems. The study showed that the extent to which the solution steps had been executed correctly was a determining factor for the correct solution of the word problems. While our findings do not imply that every student will benefit from a word problem instruction like the one we investigated, this feasibility study does provide important insights with regard to varying ways in which a word problem solving instruction can influence the solution strategies and performances of students who perform poorly on mathematical word problems.

SAMENVATTING

Het onderzoek binnen deze thesis bespreekt twee onderwerpen. Het eerste onderwerp, dat beschreven wordt in de hoofdstukken 2 tot en met 6, heeft betrekking op de prestaties van leerlingen uit groep 4 en groep 8 van het reguliere basisonderwijs met betrekking tot het oplossen van routine en non-routine talige rekenopgaven. Naast deze prestaties is in het bijzonder de mate waarin leerlingen verschillende typen visuele representaties gebruiken alsmede de rol die ruimtelijke en semantisch-linguïstische vaardigheden daarin spelen een belangrijk onderdeel van deze thesis. Een tweede onderwerp dat wordt besproken, in het specifiek in hoofdstuk 7, betreft de manier waarop leerkrachten een innovatieve instructie die betrekking heeft op het gebruik van visuele representaties in hun eigen lespraktijk implementeren.

De problemen die leerlingen ervaren tijdens het oplossen van talige rekenopgaven worden door zowel onderwijsonderzoek als de onderwijspraktijk erkend. De moeilijkheden die veel leerlingen ervaren ontstaan vaak niet door problemen in pure rekenvaardigheid, maar door problemen in het begrijpen van de tekst van een talige rekenopgave. Twee component processen, namelijk (1) het produceren van visueel-schematische representaties en (2) het destilleren van de correcte relaties tussen oplossingsrelevante informatie van de tekst van een talige rekenopgave, spelen een belangrijke rol in het goed kunnen begrijpen van deze tekst. Ruimtelijk inzicht en begrijpend lezen zijn belangrijke vaardigheden die onderliggend zijn aan beide component processen. In voorafgaand onderzoek zijn beide component processen vaak apart van elkaar onderzocht. Er werd daarbij een onderscheid gemaakt tussen een visueel-ruimtelijk domein, waar het produceren van visueel-schematische representaties en ruimtelijk inzicht onder vallen, en een semantische-linguïstisch domein, dat bestaat uit het destilleren van correcte oplossingsrelevante relaties en begrijpend lezen.

Het onderzoek dat beschreven wordt in ***Hoofdstuk 2*** is één van de eerste onderzoeken die beide domeinen in één theoretisch model onderzocht. Op basis van de resultaten van dit onderzoek kan geconcludeerd worden dat beide component processen en onderliggende vaardigheden 49% van de variantie van de prestaties van leerlingen (groep 8) op talige rekenopgaven verklaarden. Zowel het visueelruimtelijke als semantisch-linguïstische domein bleken van belang en verklaarden unieke variantie; 21% van de relatie tussen ruimtelijk inzicht en de prestaties op talige rekenopgaven werd verklaard door het produceren van visueel-schematische representaties, 34% van de relatie tussen begrijpend lezen en de prestaties op talige rekenopgaven werd verklaard door het destilleren van de correcte oplossingsrelevante relaties. Op basis van de pad analyses die gedaan zijn in dit hoofdstuk kan worden geconcludeerd dat beide component processen parallel van elkaar bestaan, en samen bijdragen tot het beter begrijpen van de tekst van een talige rekenopgaven.

Ondanks het belang van visuele representaties en semantisch-linguïstische vaardigheden, lijkt het hedendaagse realistische rekenonderwijs in Nederland beperkte aandacht te hebben voor het ontwikkelen van semantische-linguïstische vaardigheden gedurende het onderwijs met betrekking tot talige rekenopgaven. Om deze aanname kracht bij te zetten, is in ***Hoofdstuk 3*** een onderzoek gerapporteerd waarin allereerst de mate waarin efficiënte visuele representaties nodig waren werd gemanipuleerd. Dit gebeurde door een onderscheid te maken tussen consistente en inconsistente vergelijkingsproblemen. Vervolgens werd ook gevarieerd met de semantische complexiteit van dit type talige rekenopgave, door gebruik te maken van semantisch complexe (i.e., minder dan), en semantisch minder complexe (i.e., meer dan) sleutelwoorden. Onderzoek toont namelijk aan dat in het bijzonder semantisch-linguïstische vaardigheden nodig zijn om semantisch complexe sleutelwoorden te verwerken. In ons onderzoek werd een onderscheid gemaakt tussen succesvolle en minder succesvolle leerlingen op basis van hun prestaties op CITO Rekenen, een gestandaardiseerde toets die een afspiegeling geeft van de manier waarop het realistisch rekenen wordt aangeboden in Nederland.
Een belangrijke bevinding van het onderzoek in Hoofdstuk 3 is dat succesvolle leerlingen op basis van dit instrument substantiële moeilijkheden hadden met het oplossen van semantisch complexe inconsistente vergelijkingsproblemen. Dit vormt in zekere zin een ondersteuning voor de aanname dat het realistisch rekenonderwijs in Nederland kennelijk beperkt aandacht heeft voor het trainen van semantisch-linguïstische vaardigheden. Omdat talige rekenopgaven langer en complexer worden wanneer leerlingen naar het middelbaar en hoger onderwijs gaan, is het van belang dat deze vaardigheden al in het basisonderwijs voldoende worden aangeleerd.

Het onderzoek dat in ***Hoofdstuk 4*** wordt gerapporteerd onderzocht het belang van verschillende type visuele representaties, ruimtelijk inzicht en begrijpend lezen in het correct oplossen van talige rekenopgaven. Door gebruik te maken van statistische analyses op itemniveau, waren we, in tegenstelling tot het voorafgaand onderzoek dat veelal op testniveau is gedaan, in staat om diepgaander te kijken naar de rol van deze variabelen. Het onderzoek in Hoofdstuk 4 toonde aan dat het onderscheid tussen picturale en visueel-schematische representaties, zoals gemaakt in voorafgaand onderzoek, te smal is. Ons onderzoek liet namelijk zien dat het produceren van accurate visueel-schematische representaties vaker geassocieerd kan worden met een correct dan een incorrect antwoord op een talige rekenopgave. Accurate, in vergelijking met inaccurate visueel-schematische representaties geven een compleet en coherent beeld van de probleemstructuur, en bevatten de correcte relaties tussen oplossingsrelevante elementen. Accurate visueel-schematische representaties vergrootten de kans op het correct oplossen van een talige rekenopgave bijna zes keer. In tegenstelling tot accurate visueel-schematische representaties verkleinden inaccurate visueel-schematische representaties en picturale representaties de kans op het succesvol oplossen van talige rekenopgaven, respectievelijk 2.94 en 2.78 keer. Omdat picturale representaties betrekking hebben op de visuele verschijning van bepaalde elementen (objecten en/of personen) van de tekst van een talige rekenopgave, leiden ze de leerling af van het construeren van een coherente weergave van de rekenopgave, die alle correcte relaties tussen oplossingsrelevante informatie bevat. Ondanks dat inaccurate visueel-schematische representaties relaties bevatten, worden deze relaties incorrect gelegd of ontbreken er oplossingsrelevante relaties. Een gevolg daarvan is dat dit type visuele representaties leerlingen op het verkeerde been zet tijdens het oplossingsproces.

Naast een beter begrip over de mate waarin verschillende type visuele representaties de kans op het correct oplossen van talige rekenopgaven vergroten of verkleinen, werd er in dit hoofdstuk ook onderzocht of het belang van ruimtelijk inzicht en begrijpend lezen (zoals aangetoond in voorafgaand testniveau onderzoek) gerepliceerd kon worden door middel van analyses op itemniveau. In lijn met de bevindingen van vorig onderzoek, toonde ons onderzoek aan dat ruimtelijk inzicht een significante en relevante vaardigheid is die de kans op het succesvol oplossen van een talige rekenopgaven vergroot. Echter, ons onderzoek toonde aan dat begrijpend lezen in veel mindere mate deze kans vergroot. De resultaten van de logistische regressie analyses toonden namelijk aan dat, ondanks dat begrijpend lezen een significante voorspeller was, de relevantie van begrijpend lezen beperkt was. Op basis van de resultaten van het onderzoek uit Hoofdstuk 4 kan dus geconcludeerd worden dat variabelen die op testniveau aan elkaar gerelateerd zijn niet per definitie aan elkaar gerelateerd hoeven te zijn op itemniveau.

In ***Hoofdstuk 5*** worden de specifieke relaties tussen constructief spelgedrag, ruimtelijk inzicht en de prestaties op talige rekenopgaven onderzocht. In voorafgaand onderzoek werd aandacht besteed aan zowel de relatie tussen constructief spelgedrag en ruimtelijk inzicht, als aan de relatie tussen ruimtelijk inzicht en de prestatie van leerlingen op talige rekenopgaven. De relatie tussen constructief spelgedrag en de prestaties op talige rekenopgaven werd echter nog niet gerapporteerd. De resultaten van ons onderzoek toonden aan dat ruimtelijk inzicht een partiële mediator is in de relatie tussen deze twee variabelen. Dit impliceert dat leerlingen die veel spelen met constructieve spellen (zoals Lego, blokken en tangram) betere ruimtelijke vaardigheden hebben, en daarom een betere prestaties laten zien in hun prestaties op talige rekenopgaven. De drie variabelen die centraal stonden in dit onderzoek verklaarden 38.16% van de variantie in de prestaties van (groep 8) leerlingen op talige rekenopgaven. Daarnaast bleek de relatie tussen constructief spelgedrag en de prestaties op talige rekenopgaven voor 31.58% verklaard te worden door ruimtelijk inzicht.

Terwijl de focus van de eerste hoofdstukken uit deze thesis op het gebruik van visuele strategie, oplossingsprocessen en prestaties op non-routine talige rekenopgaven in de bovenbouw van het basisonderwijs ligt, gaat het onderzoek dat beschreven is in ***Hoofdstuk 6*** over prestaties op routine talige rekenopgaven van leerlingen uit groep 4. De resultaten van dit onderzoek toonden aan dat deze leerlingen meer fouten maakten tijdens het oplossen van *vergelijkings*problemen dan bij het oplossen van *deel-geheel* problemen en *voor-na* problemen. Dit kwam overeen met voorafgaand onderzoek dat de discrepanties in prestaties tussen deze drie typen talige rekenopgaven heeft onderzocht. In plaats van de aanwezigheid van een consistentie-effect (i.e., de leerlingen uit groep 4 behaalden vergelijkbare prestaties op consistente en inconsistente opgaven), is een meer algemeen probleem met het verwerken van sleutelwoorden en het begrijpen van de tekst een plausibelere verklaring voor de problemen met het oplossen van vergelijkingsproblemen. Leerlingen uit groep 4 hebben blijkbaar nog niet de kennis om de linguïstische aspecten van een vergelijkingsproblemen te begrijpen en de juiste probleemstructuur uit de opgave te destilleren.

De moeilijkheden die leerlingen ervaren tijdens het begrijpen en oplossen van talige rekenopgaven kunnen mogelijk terug te voeren zijn naar de rol van de leerkracht in het representatie- en oplossingsproces. De resultaten van ***Hoofdstuk 7*** toonden aan dat de meeste leerkrachten in staat waren om rekenmodellen in te zetten (met name verhoudingstabellen). Deze modellen zijn, in tegenstelling tot visueel-schematische representaties, slechts zinvol wanneer ze worden ingezet in de oplossingsfase van het totale oplossingsproces. Net als veel van hun leerlingen, hadden verschillende leerkrachten moeite met het maken van accurate visueel-schematische representaties. Sommige leerkrachten maakten veelvuldig gebruik van picturale en inaccurate visueel-schematische representaties gedurende de begripsfase van het oplossingsproces. De bevindingen van Hoofdstuk 4 lieten echter zien dat deze typen visuele representaties het kans op het correct oplossen van een talige rekenopgaven niet vergroten. De resultaten van Hoofdstuk 7 lieten zien dat, wanneer leerkrachten accurate visueel-schematische representaties gebruikten, in de meeste gevallen een strookmodel werd geconstrueerd. Andere vormen van visueel-schematische representaties, zoals getallenlijnen, cirkeldiagrammen en eigen producties werden beperkt gebruikt. Over het algemeen was de diversiteit in gebruik van visuele representaties dus laag. Bovendien waren de visuele representaties die werden ingezet niet altijd in overeenstemming met de karakteristieken van de rekenopgave en/of de behoeften van de leerlingen.

Op basis van de bevindingen van deze PhD-thesis kunnen de volgende aanbevelingen worden gedaan voor leerkrachtprofessionalisering en leerkrachtopleidingen:

De vaardigheid om accurate visueel-schematische representaties te maken tijdens het oplossingsproces van talige rekenopgaven zou een prominente plaats moeten krijgen in het rekencurriculum op basisscholen. Leerkrachten zouden kennis moeten hebben over het doel van dit type visuele representatie en worden getraind in het gebruik van dergelijke representaties. Ook moet aandacht besteed worden aan het moment waarop deze visueel-schematische representaties zouden moeten worden ingezet in het totale oplossingsproces.

Leerkrachtprofessionalisering en leerkrachtopleidingen zouden zich moeten focussen op het correct gebruiken van verschillende vormen van visueel-schematische representaties. Men zou leerkrachten moeten leren hoe getallenlijnen, cirkeldiagrammen en andere, zelf-geconstrueerde, visueel-schematische representaties gebruikt moeten worden.

Het constructieproces zou daarbij in het specifiek aandacht moeten krijgen. Leerkrachten moeten in staat zijn om het constructieproces transparant, correct en compleet weer te geven. Leerkrachten maken het proces transparant door uit te leggen welke elementen van de opgave moeten worden gerepresenteerd. Ook moet uitgelegd worden hoe de visuele representatie gebruikt moet worden om de talige rekenopgave op te lossen. Dit proces moet bovendien correct en compleet zijn.

Tenslotte, zou er specifieke aandacht moeten zijn voor de karakteristieken van verschillende type talige rekenopgaven. Leerkrachten moeten bijvoorbeeld weten wat het onderscheid is tussen routine en non-routine rekenopgaven, en de rol die accurate visueel-schematische representaties spelen binnen deze type talige rekenopgaven. In routine rekenopgaven (zoals deel-geheel, voor-na, en vergelijkingsproblemen) kan het gebruik van één type visuele representaties volstaan; de probleemstructuur van ieder van deze routine opgaven is namelijk identiek. De probleemstructuur van een non-routine talige rekenopgave varieert waardoor het niet passend is om slechts één accurate visueel-schematische representatie aan te bieden. Het is dus zaak dat leerkrachten leren om visueel-schematische representaties te gebruiken in een diverse (i.e., een gevarieerd aanbod aan visuele representaties) en flexibele (i.e., verschillende visuele representaties aanbieden per rekenopgave) manier. Bovendien moeten deze visueel-schematische representaties functioneel zijn en passend zijn bij de specifieke karaktereigenschappen van een rekenopgave (bijvoorbeeld het gebruik van een cirkeldiagram bij het oplossen van een rekenopgaven over percentages).

De bovenstaande aanbevelingen zijn interessante onderwerpen voor toekomstig onderzoek naar het belang van visuele representaties bij het oplossen van talige rekenopgaven. Op basis van het onderzoek in deze thesis zou toekomstig onderzoek zich in eerste instantie moeten focussen op de eigen competentie van leerkrachten en hun didactisch gebruik van visuele representaties tijdens hun rekeninstructie. Wanneer leerkrachten meer kennis hebben over het belang van visuele representaties, dan kunnen ze deze kennis gebruiken om de moeilijkheden van hun leerlingen bij het oplossen van talige rekenopgaven te verhelpen.

Het haalbaarheidsonderzoek dat tenslotte in ***Appendix I*** wordt beschreven onderzocht vier leerlingen uit groep 4 van het basisonderwijs die laag presteerden op talige rekenopgaven. Deze vier leerlingen ontvingen een instructie van hun leerkracht gedurende een vijf weken durende interventieperiode. De effectiviteit van deze instructie werd bepaald door de prestaties van de vier leerlingen op deel-geheel-, voor-na-, en vergelijkingsproblemen op de voormeting te vergelijken met de prestaties op de nameting. Daarnaast werd gekeken in hoeverre deze leerlingen de oplossingsstappen van de instructie op een correcte manier gebruikten. De resultaten van dit onderzoek lieten zien dat, ondanks dat de totale prestaties van iedere leerling omhoog ging, deze vooruitgang niet bij alle drie de type routine rekenopgaven zichtbaar was. De mate waarin de oplossingsstappen van de instructie correct werden uitgevoerd bleek een bepalende factor te zijn in het correct oplossen van de rekenopgave. Ondanks dat we niet willen suggereren dat de instructie die in dit onderzoek centraal stond zijn vruchten afwerpt voor alle leerlingen, geeft het wel belangrijke inzichten omtrent de verschillende wijzen waarop een instructie de oplossingsstrategieën en prestaties van leerlingen kan beïnvloeden.

PUBLICATIONS

Scientific publications

Boonen, A. J. H., De Koning, B., Jolles, J., & Van der Schoot, M. (*under review*). Word problem solving in contemporary math education: A plea for semantic-linguistic skills training. *Journal for Research in Mathematics Education*

Boonen, A. J. H., & Jolles, J. (*under review*). Teaching four less successful second grade students to solve combine, change and compare word problems: Results of a feasibility study. *Preventing School Failure.*

Boonen, A. J.H., & Jolles (*under review*). Second grade elementary school students' differing performance on combine, change and compare word problems. *International Journal of School and Cognitive Psychology*

Boonen, A. J. H., Kolkman, M. E., & Kroesbergen, E. H. (2011). The relation between teachers' math talk and the acquisition of number sense within kindergarten classrooms. *Journal of School Psychology*, *49*, 281-299.

Boonen, A. J. H., Reed, H. C., Schoonenboom, J. & Jolles, J. (*submitted*). It's not a math lesson, we're learning to draw! Teachers' use of visual representations in instructing word problem solving in sixth grade of elementary school. *The Journal of Mathematics Teacher Education*

Boonen, A. J. H., Van der Schoot, M., Van Wesel, F., De Vries, M., & Jolles, J. (2013). What underlies successful word problem solving? A path analysis in sixth grade children. *Contemporary Educational Psychology, 38,* 271-279.

Boonen, A. J. H., Van Wesel, F., Jolles, J., & Van der Schoot, M. (2014). The role of visual representation type, spatial ability, and reading comprehension in word problem solving: An item-level analysis in elementary school children. *International Journal of Educational Research, 68, 15-26.*

Oostermeijer, M., Boonen, A. J. H., & Jolles, J. (2014). The relation between children's constructive play activities, spatial ability and mathematical word problem solving performance: A mediation analysis in sixth grade students. *Frontiers in Psychology, 5*, 1-7.

Conference presentations

Boonen, A. J. H. (2012, July). Cognitive processes underlying mathematical word problem solving: A path analysis in sixth grade children. Poster presented at EARLI JURE 2012. Regensburg, Germany.

Boonen, A. J. H. (2012, August). Cognitive processes underlying mathemati-

cal word problem solving: A path analysis in sixth grade children. Poster presented at the Biennial Meeting of EARLI SIG 15. Utrecht, The Netherlands.

Boonen, A. J. H. (2010, September). The relation between teachers' math talk and the acquisition of number sense within kindergarten classrooms. Poster presented at the Biennial Meeting of EARLI SIG 15. Frankfurt, Germany.

Boonen, A. J. H. (2013, April). What underlies successful word problem solving: a path analysis in sixth grade students. Paper presented at SRCD Biennial Meeting. Seattle, United States.

Boonen, A. J. H. (2013, August). What underlies successful word problem solving: a path analysis in sixth grade students. Paper presented at EARLI Biennial Meeting. München, Germany.

Boonen, A. J. H. (2014, January). Comprehend, Visualize, and Calculate: the successful solution of mathematical word problems. Paper presented at 'Panamaconferentie'. Noordwijkerhout, The Netherlands.

Professional publications

Boonen, A. J. H. (2012). Ontwikkeling meetkundekennis van toekomstige leerkrachten [The development of geometrical knowledge of future teachers]. *Panama post, 2*, 36-42.

Boonen, A. J. H. (2014). Begrijpen, Verbeelden en Berekenen: Visuele representaties ondersteunen bij talige rekenopgaven [Comprehending, Visualizing and Calculating: visual representations support mathematical word problems]. *Volgens Bartjens, 3*, 25-27.

Boonen, A. J. H., & Haarsma, J. (2011). Het leren construeren en visualiseren in kleutergroepen: Ontwikkeling meetkundige kennis van de toekomstige leerkracht [Construction and visualisation in kindergarten: The development of geometrical knowledge of a future teacher]. *Panama post Praktijktip, 3*, 14-16.

Houben, M., & Boonen, A. J. H. (*under review*). Talige rekenopgaven: Het gebruik van het strookmodel bij het begrijpen en oplossen van drie typen talige rekenopgaven: een onderzoek bij 4 zwakke presteerders. *Volgens Bartjens*

DANKWOORD

Datum: 29-10-2010
Plaats: Adelaide, Australië

Geachte heer Meijer en heer Jolles,

Bijgevoegd mijn schriftelijke sollicitatie voor de vacature: Docent-onderzoeker Onderwijsinnovaties Rekenen en Bètavaardigheden in het Primair Onderwijs. Mocht mijn sollicitatiebrief u aanspreken, dan zou ik graag wat meer willen toelichten in een gesprek. Gezien mijn huidige situatie; ik bevind me tot 14 december in Australië, ben ik dankbaar dat u de mogelijkheid biedt om een eventueel gesprek via SKYPE te voeren. Gezien het tijdsverschil tussen Nederland en Australië, maar ook het feit dat ik afhankelijk ben van het internetgebruik in de accommodatie waar ik op dat moment verblijf, is het zaak om, mocht u geïnteresseerd zijn, het één en ander kort te sluiten. Daarnaast heb ik van de heer Van Luit begrepen dat de sollicitatiegesprekken op donderdag 4 november plaats vinden. Ik kom op die dag in Melbourne aan, maar kan niet 100 procent zeker zeggen op welk tijdstip. Omdat ik niet te veel op de zaken vooruit wil lopen, wacht ik eerst op uw reactie naar aanleiding van mijn schriftelijke sollicitatie. Bij voorbaat dank!

Met vriendelijke groet,

Anton Boonen

Het sollicitatiegesprek vond inderdaad plaats via SKYPE in een hostel in Melbourne, Australië. Ik zat aan de ene kant van het scherm, op het zuidelijk halfrond, Prof. dr. Jelle Jolles, dr. Erna van Hest, Drs. Bert Meijer en Drs. Jan Haarsma aan de andere kant van het computerscherm, op het noordelijk halfrond. Tijdstip van het gesprek: 01.30 uur Australische tijd!

Datum: 09-12-2014
Plaats: Wanaka, Nieuw-Zeeland

Beste Anton,

De leescommissie is akkoord. En ik heb 10 minuten geleden de formele brief naar de decaan gestuurd. Over 2 dagen spreek ik hem en zal ik checken of zijn handtekening is gezet. Nu lekker verder met je vakantie. Geniet ervan.

Beste groet,

Jelle

In het kort, komt het er dus op neer dat mijn ‘proefschrift-avontuur’ Down Under begon, en min of meer ook Down Under eindigde.

In dit dankwoord wil ik graag een aantal mensen ontzettend bedanken voor de steun en inspiratie die ze mij hebben gegeven om deze avontuurlijke reis tot een goed einde te brengen.

Allereerst mijn promotor. Jelle, jouw visie op onderzoek spreekt me nog steeds enorm aan en ik heb geleerd om mijn eigen draai aan deze visie te geven. Daarnaast waardeer ik je als persoon en heb ik veel geleerd van je wijze les om voor mezelf op te komen en niet altijd de wens te willen hebben om als een, zoals jij dat noemde ‘herdershond’, iedereen bij elkaar te houden.

Menno, ook jouw visie op onderzoek spreekt mij enorm aan en ontzettend bedankt dat ik gebruik heb kunnen maken van jouw kennis en ervaring. Ik ben blij dat je uiteindelijk mijn copromotor ben geworden.

Daarnaast wil ik mijn andere coauteurs ontzettend bedanken voor het werk dat ze verzet hebben. Floryt, dank voor je hulp bij mijn statistische problemen, maar vooral ook voor je gezelligheid en collegialiteit. Meinou, Björn, Judith en Meike, dank dat jullie mijn onderzoek interessant vonden en mij met, naar mijn idee, veel plezier hielpen met het schrijven van artikelen.

Een speciaal woord tot Helen. We hebben dezelfde researchmaster gevolgd, en zitten vaak op één lijn wanneer het over rekenonderwerpen gaat. Ik wil je persoonlijk bedanken voor alle tijd, energie en aandacht die je mij en mijn proefschrift gegeven hebt. Dank dat je me, op momenten dat ik daar aan twijfelde, erop wees trots op mezelf te zijn.

Vervolgens wil ik de leden van de leescommissie bedanken voor hun kritische blik en oordeel. In het bijzonder noem ik professor Hans van Luit. Hans, tijdens jouw colleges aan de Universiteit Utrecht begon mijn interesse voor rekenontwikkeling en rekenonderwijs. Dank dat je me attendeerde op de vacature voor docent-promovendus!

Ik bedank ook mijn mede PhD-ers en andere collega's van de VU voor hun collegialiteit en gezelligheid. Jeffrey, ook al was ik er niet vaak, op de momenten dat ik er was behaalden we altijd het meest abominabel lage humorniveau. Ik ga de Tour & WK poule, maar met name de stukken proza over de Ethiekcursus missen. Ik ben blij dat je mijn paranimf bent.

Ook wil mijn twee beste oud-studenten bedanken. Meike & Mirte: zonder jullie hulp en inzet had ik in geen mogelijkheid de interessante onderzoek van hoofdstuk 5, 6 en Appendix I kunnen doen. Dankjulliewel en we houden contact.

Naast mijn collega's bij de VU, bedank ik uiteraard ook mijn l ieve collega's van Windesheim. In het bijzonder de (oud-)vakgroep Rekenen (Daniel, Jarise, Jan H., Geert, Nathalie, Els en Merel) en de onderzoekscommissie (Job, Marjon, Mariëlle, Marjolein, Suzanne, Marlies A. & Martinette). Bert en Jan, dank dat jullie vertrouwen in mij hadden en nog steeds hebben en mij de prachtige kans hebben aangeboden om mijn werkzaamheden voort te zetten op Windesheim. Ik ben ontzettend gemotiveerd om het onderwijs en onderzoek binnen het domein Bewegen en Educatie op een hoger niveau te brengen.

Een speciaal bedankje voor Marlies kan niet ontbreken. Dank voor de mooie tekeningen die in dit proefschrift zijn terecht gekomen. Het is prachtig geworden.

Naast mijn collega's wil ik mijn vrienden van SVE bedanken. Daarbij wil ik in het bijzonder Jasper, Maarten, Sean, Timeas, Wouter, Robbert, Jan, Peter, Mark, Arno-Luc en Erik noemen.

Marloes, Marlies en Dennis, jullie jarenlange vriendschap doet me ontzettend veel; dank voor jullie interesse in mijn onderzoek en steun de afgelopen jaren.

Paul, ik ken je al meer dan 24 en je bent al vanaf dat wij leerden rekenen mijn maatje. We lopen de deur niet bij elkaar plat, maar we zijn er als we elkaar nodig hebben. Fijn dat jij mijn paranimf wil zijn.

Mijn schoonfamilie, Hans, Annemieke & Paul; jullie waren altijd geïnteresseerd. Dank voor jullie support!

Hanneke, lieve zus, ik ben blij dat jij jouw weg en geluk met Martin ook gevonden hebt. Ik weet dat je ontzettend trots op me bent, maar weet dat ik net zo trots op jou ben.

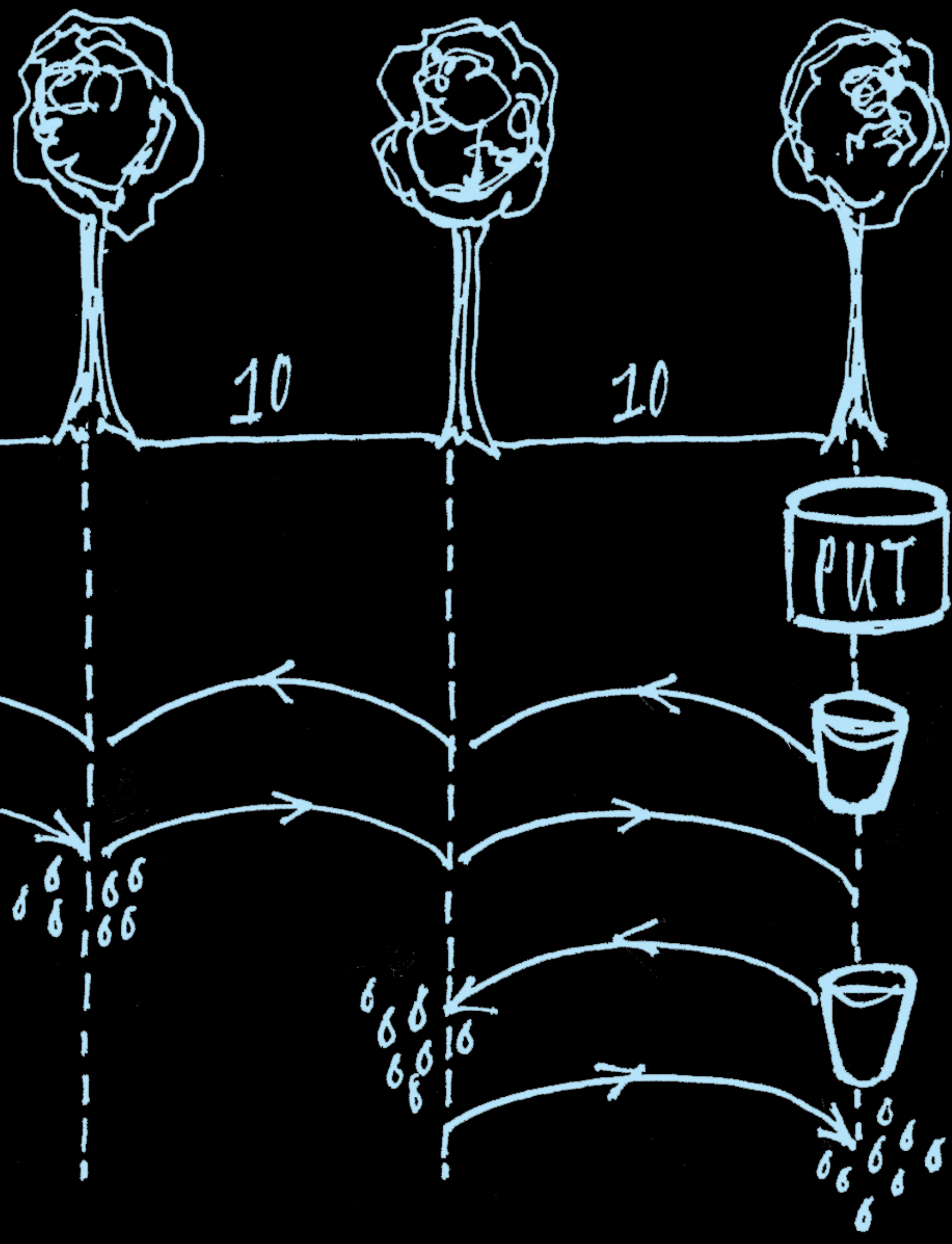
Mam en pap, door jullie heb ik door leren zetten, tevreden te zijn met minder en leren vechten om meer te kunnen bereiken. Ik ben jullie eeuwig dankbaar; hou van jullie.

En tenslotte, mijn maatje, mijn alles. Lieve Anne, na een orkaan te hebben weerstaan, ben ik zielsgelukkig dat jij nog steeds aan mijn zijde staat. Ik draag dit boek op aan jou en wil ik afsluiten met een regel uit één van de, voor mij, mooiste liedjes:

"Ben veel meer nog, ben de man die alles kan,
alles kan want ik ben jouw man (Acda, 2007)".

CURRICULUM VITAE

Anton Boonen was born on March 12, 1986 in Heeze, the Netherlands. He obtained his high school degree (VWO; gymnasium) in 2004 from 'Strabrecht College' in Geldrop. He obtained his bachelor degree in Pedagogical Sciences in 2008 and obtained his Research Master in Educational Sciences at Utrecht University (2008-2010, cum laude, Valedictorian). From January till May 2010 he did an internship at the University of Helsinki, Finland. During his period at Utrecht University, the mathematical development of students and the teaching of teachers became his primary research interests. Since December 2010, Anton has been employed for three days a week as a researcher at VU University Amsterdam, where he completed his PhD research. For two days a week, he worked at Windesheim, University of Applied Sciences, as a lecturer (mathematics & research) at the Department of Initial Teacher Education.


10
10
PUT