



ELSEVIER

Contents lists available at [ScienceDirect](http://www.sciencedirect.com)

International Journal of Educational Research

journal homepage: www.elsevier.com/locate/ijedures

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

VU University Amsterdam, The Netherlands

ARTICLE INFO

Article history:

Received 9 April 2014

Received in revised form 2 August 2014

Accepted 4 August 2014

Available online

Keywords:

Word problem solving

Visual representations

Spatial ability

Reading comprehension

Item-level analysis

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.

© 2014 Published by Elsevier Ltd.

1. 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.

^{*} Corresponding author at: Department of Educational Neuroscience and LEARN! Research Institute for Learning and Education, Faculty of Psychology and Education, VU University Amsterdam, Van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands. Tel.: +31 20 59 83459.
E-mail address: A.J.H.Boonen@vu.nl (Anton J.H. Boonen).

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, 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.


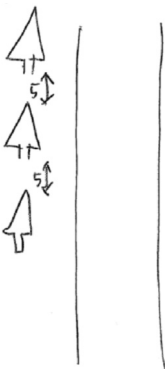
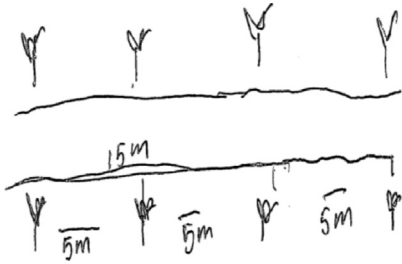
Hence, we can conclude that the way in which the relation between (type of) visual representation and word problem solving performance was investigated in past research does not provide a decisive answer to the question if, and to what extent, visual representation type affects the chance of producing 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.

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

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 m along the path he planted another tree. The length of the path is 15 m.		
How many trees were planted?		
Pictorial representation	Inaccurate visual-schematic representation	Accurate visual-schematic representation
		

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, Bakker Arkema, Horsley, & Van Lieshout, 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.

2. Method

2.1. Sample

One hundred twenty-eight Dutch sixth-grade students (64 boys, $M_{\text{age}} = 11.73$ years, $SD_{\text{age}} = 0.43$ years and 64 girls, $M_{\text{age}} = 11.72$ years, $SD_{\text{age}} = 0.39$ 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.

2.2. 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.

2.2.1. 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).

2.2.2. 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. Fig. 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 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). Fig. 2 shows an example of an inaccurate visual-schematic representation.

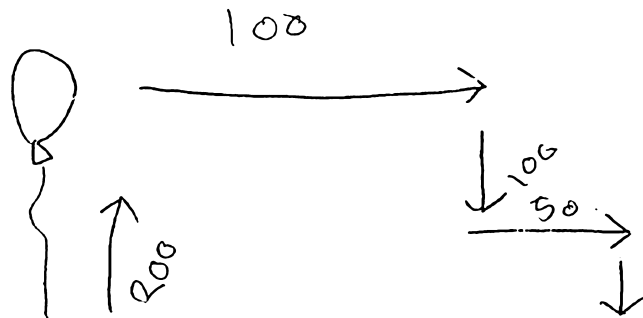


Fig. 1. An example of an accurate visual-schematic representation.

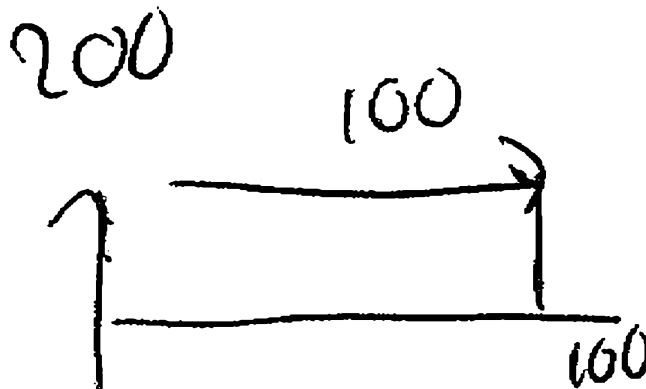


Fig. 2. An example of an inaccurate visual-schematic representation.

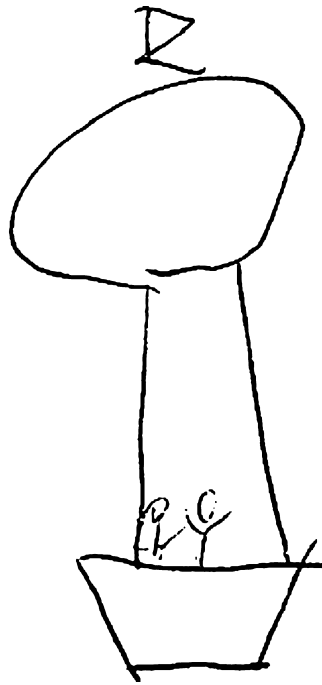


Fig. 3. An example of a pictorial representation.

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. Fig. 3 shows 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.

2.2.3. 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. Fig. 4 shows one of the 20 test items of the Paper Folding task. This task took 6 min and had a sufficient internal consistency coefficient in the present study (Cronbach's alpha = .70).

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. Fig. 5 shows one of the 30 test items of the Picture Rotation task. This task took 1.5 min and its internal consistency coefficient in this study was high (Cronbach's alpha = .93).

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 ($M = .00$, $SD = .84$).

2.2.4. 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

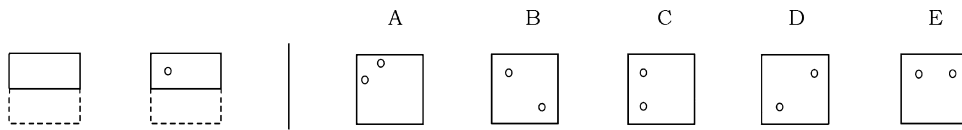


Fig. 4. The Paper Folding task (Ekstrom, French, Harman, & Derman, 1976).



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

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 = 15.00–95.00) 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).

2.3. Analyses

To replicate the findings of previous studies (e.g., Hegarty & 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.

3. Results

3.1. 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)

Table 2

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

	<i>M</i>	<i>SD</i>	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 \times 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	
Representation	Pictorial representation	Count	68	21	89
		Std. residual	3.2	-3.3	
Total		Count	936	856	1792

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: $r = .31, p < .001$; $r = .23, p < .05$).

3.1.1. 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: $\chi^2(3) = 220.10, p < .001$ (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.

3.1.2. 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 ($\text{Exp}(B)$) regression

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	$\text{Exp}(B)$	SE	$Wald$	B	$\text{Exp}(B)$	SE	$Wald$	B	$\text{Exp}(B)$	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*

* $p < .05$.

** $p < .001$.

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, $B = 1.77$, $\text{Exp}(B) = 5.85$, $SE = .17$, $p < .001$). 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: $B = -1.09$, $\text{Exp}(B) = 0.34$, $SE = .17$, $p < .001$; pictorial representation vs. no visual representation: $B = -1.02$, $\text{Exp}(B) = 0.36$, $SE = .26$, $p < .001$). *Model 2* shows 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: $B = 0.43$, $\text{Exp}(B) = 1.53$, $SE = .27$, $p = .11$). 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 ($B = 0.51$, $\text{Exp}(B) = 1.66$, $SE = .08$, $p < .001$). On the other hand, a 1 unit increase in reading comprehension skills increased the chance of solving a word problem correctly 1.02 times ($B = 0.02$, $\text{Exp}(B) = 1.02$, $SE = .00$, $p < .05$). 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 ($\text{Exp}(B) = 2.22$), whereas the role of spatial ability in pictorial representation was the lowest ($\text{Exp}(B) = 1.13$). In contrast to spatial ability, the relevancy of reading comprehension in problem solving success did not differ between the different representation types ($\text{Exp}(B) = 1.02$).

4. 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). Inline 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](#)).

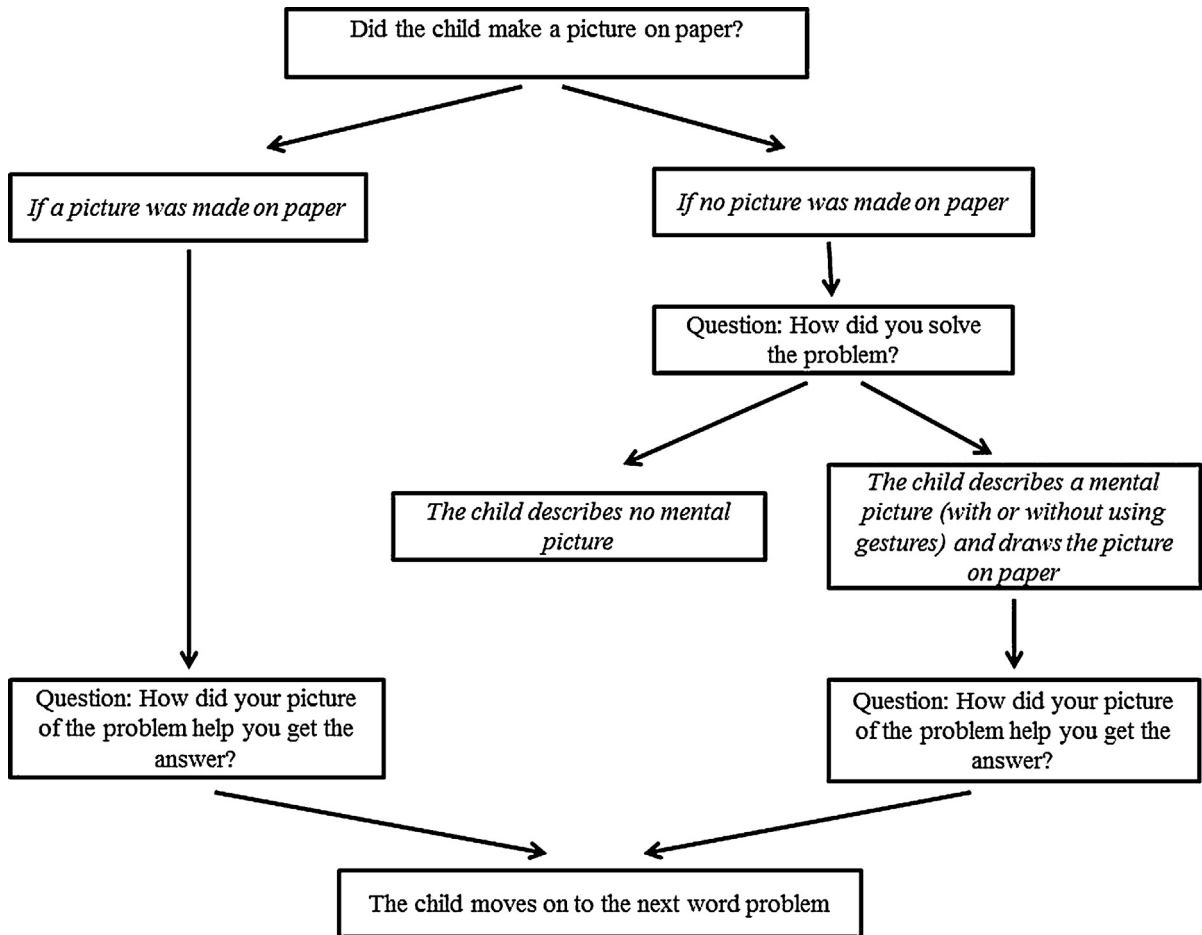
Appendix A

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 m along the path he planted another tree. The length of the path is 15 m. 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 m from the ground, then moved 100 m to the east, then dropped 100 m. It then traveled 50 m 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 3 m behind Jim. How far is Peter ahead of Tom?
5. A square (A) has an area of 1 m^2 . 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 m^2 . If its length is 10 m, 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 m 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° to the right. An hour later it made a turn through 45° to the left. In what direction was it then traveling?

Appendix B

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



References

- 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. <http://dx.doi.org/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. <http://dx.doi.org/10.1080/0144341990190203>
- 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. <http://dx.doi.org/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. <http://dx.doi.org/10.1016/j.cedpsych.2013.05.001>
- Bryant, D. J., & Tversky, B. (1999). Mental representations of perspective and spatial relations from diagrams and models. *Journal of Experimental Psychology*, 25, 137–156. <http://dx.doi.org/10.1037/0278-7393.25.1.137>
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159. <http://dx.doi.org/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. [http://dx.doi.org/10.1016/0010-0285\(88\)90011-4](http://dx.doi.org/10.1016/0010-0285(88)90011-4)
- 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. <http://dx.doi.org/10.1037/0022-0663.77.4.460>
- De Corte, E., Verschaffel, L., & Pauwels, A. (1990). Influence of the semantic structure of word problems on second graders' eye movements. *Journal of Educational Psychology*, 82, 359–365. <http://dx.doi.org/10.1037/0022-0663.82.2.359>
- 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. <http://dx.doi.org/10.1080/00220973.2012.745468>
- 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. <http://dx.doi.org/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.
- Eliá, I., Van den Heuvel-Panhuizen, M., & Kovolou, A. (2009). Exploring strategy use and strategy flexibility in non-routine problem solving by primary school high achievers in mathematics. *ZDM Mathematics Education*, 41, 605–618. <http://dx.doi.org/10.1007/s11858-009-0184-6>

- Guoliang, Y. (2003). Visual-spatial representations and mathematical problem solving among mathematical learning disabilities. *Acta Psychologica Sinica*, 35, 643–648.
- Hegarty, M., & Kozhevnikov, M. (1999). Types of visual-spatial representations and mathematical problem solving. *Journal of Educational Psychology*, 91, 684–689. <http://dx.doi.org/10.1037/0022-0663.91.4.68>
- 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. <http://dx.doi.org/10.1037/0022-0663.87.1.18>
- Jitendra, A. K. (2002). Teaching students math problem-solving through graphic representations. *Teaching Exceptional Children*, 34, 34–38.
- Jitendra, A. K., Star, J. R., Starosta, K., Leh, J. M., Sood, S., Caskie, G., et al. (2009). Improving seventh grade students' learning of ratio and proportion: The role of schema-based instruction. *Contemporary Educational Psychology*, 34, 250–264. <http://dx.doi.org/10.1016/j.cedpsych.2009.06.001>
- 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. <http://dx.doi.org/10.1080/10459880903493104>
- 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. <http://dx.doi.org/10.1016/j.intell.2007.11.001>
- 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. <http://dx.doi.org/10.1037/0033-295X.92.1.109>
- 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. http://dx.doi.org/10.1207/S1532690XCI2001_3
- Krawec, J. L. (2010). *Problem representation and mathematical problem solving of students with varying abilities (doctoral dissertation)*. Miami: University of Miami.
- Krawec, J. L. (2012). Problem representation and mathematical problem solving of students of varying math ability. *Journal of Learning Disabilities*, XX, 1–13. <http://dx.doi.org/10.1177/0022219412436976>
- 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. <http://dx.doi.org/10.1016/j.chb.2008.12.010>
- Lewis, A. B., & Mayer, R. E. (1987). Students' miscomprehension of relational statements in arithmetic word problems. *Journal of Educational Psychology*, 79, 363–371. <http://dx.doi.org/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.
- 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. http://dx.doi.org/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.
- Montague, M. (2003). *Solve It! A practical approach to teaching mathematical problem solving skills*. VA: Exceptional Innovations Inc.
- Montague, M., Warger, C., & Morgan, T. H. (2000). Solve It! Strategy instruction to improve mathematical problem solving *Learning Disabilities Research & Practice*, 15, 110–116. http://dx.doi.org/10.1207/SLDRP1502_7
- 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. <http://dx.doi.org/10.1177/0731948711421762>
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231–259.
- Pape, S. J. (2003). Compare word problems: Consistency hypothesis revisited. *Contemporary Educational Psychology*, 28, 396–421. [http://dx.doi.org/10.1016/S0361-476X\(02\)00046-2](http://dx.doi.org/10.1016/S0361-476X(02)00046-2)
- 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. http://dx.doi.org/10.1207/S15327574IJT0303_2
- Robinson, W. S. (2009). Ecological correlations and the behavior of individuals. *International Journal of Epidemiology*, 38, 337–341. <http://dx.doi.org/10.1093/ije/dyn357>
- Sanchez, C. A., & Wiley, J. (2006). An examination of the seductive details effect in terms of working memory capacity. *Memory & Cognition*, 34, 344–355. <http://dx.doi.org/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. <http://dx.doi.org/10.1007/s11251-007-9029-2>
- 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. <http://dx.doi.org/10.1016/j.jecp.2011.12.001>
- Tabachnick, B. G., & Fidell, L. S. (2006). *Using multivariate statistic (fifth edition)*. Boston: Allyn & Bacon.
- Van der Schoot, M., Bakker Arkema, A. H., Horsley, T. M., & Van Lieshout, E. D. C. 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. <http://dx.doi.org/10.1016/j.cedpsych.2008.07.002>
- Van Garderen, D. (2006). Spatial visualization, visual imagery, and mathematical problem solving of students with varying abilities. *Journal of Learning Disabilities*, 39, 496–506. <http://dx.doi.org/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. <http://dx.doi.org/10.1111/1540-5826.00079>
- Van Meter, P. (2001). Drawing construction as a strategy for learning from text. *Journal of Educational Psychology*, 93, 129–140. <http://dx.doi.org/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. <http://dx.doi.org/10.1007/s10648-005-8136-3>
- 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. <http://dx.doi.org/10.1037/0022-0663.84.1.85>
- Vilenius-Tuohimaa, P. M., Aunola, K., & Nurmi, J.-E. (2008). The association between mathematical word problems and reading comprehension. *Educational Psychology*, 28, 409–426. <http://dx.doi.org/10.1080/01443410701708228>
- 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.