Understanding and solving multistep arithmetic word problems

GURI A. NORTVEDT

This article discusses the findings of a study in which the interplay between reading, numeracy, and strategies for working on multistep arithmetic word problems was researched through two approaches. The first approach involved analysing results on national tests in reading and numeracy for a representative sample of 1,264 grade 8 (13-years-old) students. A scale of ten multistep arithmetic word problems was identified in the numeracy test. Proficiency in reading explained 44 % of the variance in scores on this scale, indicating a positive relationship between reading comprehension and success in word problem solving. The second approach involved analysing verbal protocol data for 19 grade 8 students who worked on a collection of multistep arithmetic word problems. Protocols consisted of both independent and scaffolded work. Interpretive analysis of student work on one of the eight word problems given in the protocol sessions revealed three main areas of difficulties: representing quantities in the word problem text, retrieving number facts from memory, and performing basic operations. Difficulties within more than one area were frequent. To students with below-average numeracy skills, executing the basic operations was the main obstacle for this particular word problem.

Word problems are widely used for both teaching and assessment purposes, hence, students face a wide spectrum of word problems for a number of purposes throughout their school years. The struggles and difficulties they face in solving these problems have intrigued researchers for a long time (for a review, see, for instance Lesh & Zawojewski, 2007; Reed, 1999; Verschaffel, de Corte & Greer, 2000). In this ongoing study, we research grade 8 students' competence in solving multistep arithmetic

Guri A. Nortvedt
University of Oslo

Nortvedt, G. A. (2010). Understanding and solving multistep arithmetic word problems.
Nordic Studies in Mathematics Education, 15 (3), 23–50.

word problems. The focus of interest is on how students' strategic actions towards understanding and solving the word problems are connected not only to each other, but also to the students' proficiencies in reading and numeracy. This article reports general findings regarding the interplay between reading and numeracy and solving multistep arithmetic word problems before aiming to answer the question of why or how a textbook word problem, assumed to be traditional and easy, was difficult to solve correctly. Underlying sources of the interplay between reading and numeracy and strategies for working on multistep arithmetic word problems are exemplified and highlighted through an interpretive analysis of students' work towards solving this word problem.

The interplay between reading, numeracy, and working on word problems

Word problems can be described as "verbal descriptions of problem situations wherein one or more questions are raised that the answer to which can be obtained by the application of mathematical operations to numerical data available in the problem statement" (Verschaffel et al., 2000, p. ix). Much of the research previously conducted within the field of problem solving is on solving word problems and in particular, on problems that can be solved by performing only one of the basic operations (Reed, 1999; Verschaffel, Greer & de Corte, 2007). Much emphasis is also devoted to investigating the differences in novice and expert behaviour (Lesh & Zawojewski, 2007). Throughout the past 40 years, an interest in researching the connection between reading comprehension and success in solving word problems has also intrigued a large number of researchers. This research has subsequently established that reading comprehension has a significant and positive correlation to solving word problems (Cook, 2006; Cummins, Kintsch, Reusser & Wiemer, 1988; de Corte & Verschaffel, 1991; Nathan, Kintsch & Young, 1992; Nortvedt, 2009; Roe & Taube, 2006; Thevenot, Devidal, Barrouillet & Fayol, 2007; Verschaffel et al., 2000; Vilenius-Tuohimaa, Aunola & Nurmi, 2008). In their analysis of Programme for International Student Assessment (PISA) 2003 data, Roe and Taube (2006) found that overall scores in reading and mathematical literacy had a correlation coefficient of .57 for Norwegian and Swedish students. All PISA mathematics items are word problems, and Roe and Taube suggest that the "close relationship [...] may be explained by the fact that reading and mathematics are both parts of the general concept of 'literacy', which is dominant in each assessment area in PISA" (Roe & Taube, 2006, p. 131). Further analyses revealed that the crucial aspect was not text length; rather, Roe and Taube suggest that elements that influence text quality, such as hidden or misleading information or low-frequency words, might also influence mathematics performance. In an analysis of national test data for Norwegian grade 8 students, Nortvedt (2009) found a correlation of .714 ($p < .001$) between reading comprehension and numeracy. Further dividing comprehension into sub-constructs, i.e. the ability to retrieve, interpret, and reflect on text content, the strongest relationship was found between retrieving information and numeracy.

While no universally accepted definition of numeracy exists, several of the definitions share some notion of the student as someone who is "able to understand mathematical ideas in various contexts and to apply those ideas to learn more about the context in which they are embedded" (Hurst, 2008, p. 273). Askew, for instance, defined numeracy as "the ability to process, communicate, and interpret numerical information in a variety of contexts" (Askew, 1999, p. 93), i.e. to consist of the five strands: numbers and the number systems; measures, shapes and space; calculations; handling data; and solving problems (Department of Education and Employment (DfEE), 1999; Hughes, Desforges, Mitchell & Carré, 2000). Included in numeracy is proficiency in handling numbers and number operations and applying them to solve multistep arithmetic word problems. This definition of numeracy is close to the definition of "grunnleggende ferdighet i å kunne regne"[1] (Ministry of Education and Research & Norwegian Directorate for Teaching and Training, 2006).

Traditionally, authors divide solving word problems into a series of phases or steps through which solvers progress (cf. Mayer, 2003; Pólya, 1957; Reed, 1999; Schoenfeld, 1985; Verschaffel et al., 2000).[2] These models involve one or more steps in each phase and typically have four or five steps in total. Most models include a phase of reading or understanding the word problem, making a model or planning, and executing. Some models even include a phase of looking back or evaluating the answer. A simpler division is to divide the solving of word problems into two phases: *understanding* and *solving*. The first phase consists of reading and comprehending the word problem and includes forming a mental model of the problem. To understand means to create a mental representation of the problem situation to an extent that enables the solver to solve the problem (Cook, 2006; Nathan et al., 1992; Thevenot et al., 2007). This combines the phases of translating and integrating in Mayer's model (Mayer, 2003) or understanding and modeling in Verschaffel et al.'s model (Verschaffel et al., 2000). The second part, solving, includes all the other steps students make towards achieving a solution (planning, executing, evaluating, etc). It is important to note that whether the activity of solving a word problem is problem solving or a routine activity, is solely dependent on the solver: something that is well-known and routine to a proficient student can be unknown and novel to another student.

Put simply, the process of understanding and solving a word problem is considered problem solving if it requires the problem solver to reason (Mayer, 2003).

Understanding

When reading, we draw on prior knowledge to interpret and elaborate on text content (Snow, 2002). When reading word problems, our prior knowledge in part has to be knowledge regarding the social context in which the problem is set. However, the most crucial prior knowledge is the mathematical knowledge neccessary to model the given situation. This mental model has the form of a situation model stored in working memory. It contains mathematical as well as non-mathematical information connected to the context of the word problem. Recent research suggests that the situation model does not have the form of an activated schema, as is often suggested (cf. Coquin-Viennot & Moreau, 2007; Nathan et al., 1992). Rather, it is a more qualitative temporary structure that contains mathematical relationships embedded within the text (Thevenot et al., 2007).

During the initial reading and rereading of the word problem, students employ different reading strategies to discriminate between text elements and to identify both relevant information and relations between these pieces of information (Cook, 2006; Cook & Rieser, 2005). Number-grabbing or direct translation of keywords (Hegarty, Mayer & Monk, 1995) can be termed "surface-level strategies" since students, when employing such strategies, do not consider relationships between text elements (Alexander et al., 2004). It could be argued that when students misinterpret keywords or fail to comprehend a word problem, it is because they lack prior mathematical knowledge that would have allowed a more elaborate use of keywords and deep-level comprehension. Cummins et al. (1988) found that although students in their study made errors or arrived at an erroneous solution, they had correctly solved the problem as they comprehended it. It could be argued, however, that this is because the problem as it is captured in their model is a more simplistic problem due to a surface-level discrimination that fails to detect more complex mathematical relationships.

Planning – Part of understanding or solving?

Once a word problem is comprehended, the next necessary step is planning (Pólya, 1957). A clear-cut boundary between understanding and solving cannot be established, as planning can be part of understanding if a problem is a routine problem for the student (Schoenfeld, 1985). For struggling students and less proficient problem solvers, planning is more often expected to be part of solving the problem. Verschaffel et al. (2007) suggest that this group might detach from the problem context once they have identified the information needed to proceed towards execution.

Solving

Besides planning, solving includes all the other steps students make towards producing a solution to a word problem. Multistep arithmetic word problems can be solved by applying a combination of the basic operations (Reed, 1999). Thus, for solving a multistep word problem, planning includes identifying all the neccesary operations before carying them out in the given order. When students form inappropriate situation models during word problem comprehension, such as creating a model that only contains part of the problem or a simplified model, the plan is thus affected, as only incomplete or incorrect operations will be recognized as neccesary.

The basic operations consist of a series of single-digit or multidigit operations. Single-digit operations range from counting to retrieving number facts from memory (Geary & Hoard, 2003; Verschaffel et al., 2007). To perform multidigit calculations, students employ either mental calculation strategies or use standard algorithms. Like Verschaffel et al. (2007), we include informal strategies involving pen-and-paper use in mental strategies.[3] A large number of such strategies have been identified in prior research; these range from direct modeling strategies, in which students draw and count tallies, to complex and flexible compensating strategies (for a review, see Verschaffel et al., 2007). Students' flexibility and use of such strategies is closely connected to their number concepts. This is not always the case for students' mastery of the standard algorithms; these can be executed as a non-consious series of steps through which they must progress in a given order (Thompson, 1999), and students tend to use algorithms in a "stereotype, inflexible way" (Verschaffel et al., 2007, p. 576). A second concern is the number of errors students make when using them (Kilpatrick, Swafford & Findell, 2001). When executing multidigit calculations using standard algorithms, students need to produce number facts by either retrieving them from memory or by some other strategy. Struggling students rely more on counting strategies or decomposition strategies than on retrieving number facts from memory (Ostad & Sorensen, 2007). This lessens the attention and capacity that could otherwise be directed towards their performance and monitoring of progress in the overall work on the word problem.

Polya (1957) states that the last phase of solving a problem is to look back, to check the answer, or to validate. Validating answers involves not only checking calculations, but also evaluating whether the given answer makes sense in relation to the context situation given in the word problem and in relation to the solver's prior knowledge. However, students often accept nonsensical answers to word problems with an everyday or realistic context (Palm, 2008; Verschaffel et al., 2000).

Methodology

Individual students learn mathematics by participating in different mathematical practices, some of which are the mathematics classrooms they enter as part of their formal schooling (Cobb, Stephan, McClaim & Gravemeijer, 2001; Lesh & Zawojewski, 2007). Through their participation in the different practices they encounter, students develop concepts, beliefs, understandings, and ways of knowing and solving mathematical tasks (Lave, 1996; Lerman, 1998). The empirical object of study in this research project is student competence in the form of domain knowledge and strategies for solving multistep arithmetic word problems. This competence is displayed in part through the mastery that students demonstrate on national tests and in part through their thinking aloud and interactions while solving a collection of multistep arithmetic word problems. In short, students' responses are viewed as patterns of participation that at the same time represent students' competence within and across situations (cf. Bingobali & Monaghan, 2008; Greeno, 1997).

The framework of the study coordinates the social perspective of patterns of participation across situations with a more cognitive perspective focusing on the reasoning of the participating students. Even though the unit of analysis is connected to the individual, this is not an individual in the "pure" cognitive perspective tradition. This can be exemplified by the manner in which students' score patterns on national tests are viewed as patterns of participation as they describe participation ranging from peripheral to more engaged and complex (Bingobali & Monaghan, 2008).

The aim of the study is to investigate student competence for solving multistep arithmetic word problems; that is, exploring the interplay between mathematical knowledge and skills, text comprehension, and strategies for working on such problems. The overall study has a mixed method design consisting of a large-scale quantitative analysis of the relationship between aspects of reading comprehension and numeracy, and a qualitative analysis of protocols from task-based interviews with 19 students who worked on a collection of multistep word problems.

While the large-scale data gives information on general patterns among Norwegian grade 8 students, the protocols provide insights into students' individual patterns. Data were collected in June–October 2007. Access to data on national tests was granted by the Norwegian Ministry of Education and Research.

Measures

Two instruments for assessing students' competence in working on multistep arithmetic word problems have been employed. Students' proficiency levels in reading comprehension and numeracy were measured using national tests.[4] Students' competence in the form of their use of strategic actions was assessed using task-based interviews. As students were scaffolded by the researcher, these measures should not be seen as measures of students' strategy repertoires, but rather, as insights into what they can do when provided with the kind of scaffolding given in the interviews (Goldin, 2000).

National tests

National tests are developed according to a national framework[5] that identifies reading comprehension as the ability to retrieve information, interpret, and reflect on text content. Two text formats are used: continuous and non-continuous. Competence in numeracy, according to the national framework, is identified as the ability to apply knowledge in number, measurement, and statistics to solve routine and non-routine (embedded) problems (Norwegian Directorate for Education and Training, 2006). Tests were validated by both external expert groups and the research groups themselves. Test reliability for both tests was above .85, allowing for a correlation analysis (Cohen, Cohen, West & Aiken, 2003). A scale of ten multistep arithmetic word problems was identified among the 76 test items. This scale was identified as part of the initial analyses performed within this research project. Cronbach's alpha for the word problems scale was .716, indicating that this scale gives reasonable reliable information on students' performance on multistep arithmetic word problems.

Verbal protocols

The verbal protocol session was a task-based interview (Goldin, 2000) designed for this study to assess students' emerging competence in solving multistep arithmetic word problems. The eight items could all be solved using a combination of basic arithmetic operations. All problems had a few lines of text, and some contained extraneous or irrelevant information. Items represented different complexity levels to both challenge proficient students and provide struggling students with word problems that fell within their levels of proficiency.

Scaffolding can be described as ”controlling those elements of the task that are initially beyond the learner’s capacity, thus permitting him to concentrate upon and complete only those elements that are within his range of competence” (Wood, Bruner & Ross, 1976, p. 9). Hence, scaffolding enabled a larger portion of students’ competence to be displayed than it otherwise would have been if a student had been left alone and unable to solve the problem (Goldin, 2000). Wood et al. (1976, p. 96) claim that ”well executed scaffolding begins by luring the child into actions that produce recognizable-for-him solutions”. A possible consequence of this might be that in situations where students had formed a situation model that was not appropriate due to comprehension difficulties or lack of prior knowledge, scaffolding at times resulted in students’ correctly solving the problem as it was comprehended. In other cases, students, although scaffolded, performed a calculation error and arrived at an incorrect answer despite following an appropriate solution method.

Word problem 2 (WP2)

Word problem 2 (WP2) is a textbook problem with a shopping context that involves buying clothes. Prices are given as whole numbers close to realistic prices, e.g. 79 for 79.90 kroner. No extraneous or irrelevant information is given. The word problem is considered ”school authentic”, as the context situation describes an everyday situation assumed well-known to 13-year-old students. The word problem is not considered to be authentic (Palm, 2006), as the price of socks would be given on the receipt, and as realistic prices would include decimals. The mean score for the 19 students was .66 points (SD = 0.36). The correlation to the word problem sum score was .477, $p = .05$.

> **Oppgave 2 – sokker og t-skjorter**
> Aud kjøpte tre par sokker og fire t-skjorter. T-skjortene kostet 79 kroner per stk. Til sammen betalte hun 364 kroner. Hva kostet et par sokker?
>
> **WP2 – socks and t-shirts**
> Aud bought three pairs of socks and four t-shirts. The t-shirts were 79 kroner each. Altogether she paid 364 kroner. How much did one pair of socks cost?

Figure 1. *WP2 text body*

Samples

The sample for the national test data is a clustered random sample consisting of 1,264 students from 26 schools situated across Norway. The sample included a mix of small and large schools, with student enrolment in grade 8 ranging from 4 to 146 students. Only students present for both tests were included in the sample. Grade 8 students are 13-years-old.

Students were sorted into four achievement groups based on their test scores. As all items were dichotomous, students needed to score 39 points in numeracy and 28 in reading to be assigned to an above-average group; both cuts were approximately at the 51st percentile. Mean scores were M = 39.40 (of 76), SD = 15.92, for the numeracy test and M = 26.62, SD = 8.24, for the reading test.[6] It should be noted that the criteria for dividing students into achievement groups means that students quite similar to each other can be found in the four achievement groups as no middle band is used. However, for the analysis in this paper, the achievement groups will be used to illustrate how students with similar measures on one test (numeracy) but with different scores on the other assessment (reading) differ when it comes to solving multistep arithmetic word problems. The sorting is considered useful, although it limits possible further analysis.

Table 1. Assignment to achievement groups compared to sample means in national tests

Group	Numeracy	Reading
LL	Below	Below
LH	Below	Above
HL	Above	Below
HH	Above	Above

The 19 students in the verbal protocol sample came from two different combined primary and secondary schools situated in Oslo. Both schools have a mix of socio-cultural backgrounds and first- and second-language learners. On the basis of classroom observation, teacher judgment, and test scores on national tests, students were selected to take part in the study to ensure that from both schools, an even mix of boys and girls representing a wide range of mathematical proficiencies were included. Students with identified dyslexia were excluded. Students and their parents had given informed consent prior to the onset of the study.

Data collection procedures

Norwegian teachers score and report their students' national test results through a national web application. Test scores on item level for each student from the national database were provided for this study by the Norwegian Directorate for Education and Training. A missing data analysis was performed to ensure that the sample was still representative. Students sat for tests in September 2007. Data was handed over in October 2008.

The verbal protocols were collected during the last week of teaching in grade 7 for four of the students. The other interviews were collected during the autumn of 2007, when students were in their first months of grade 8. Test scores on item level were collected from students' national tests.

Verbal protocols were collected in an interview setting during school hours. The student and researcher were seated opposite each other in a secluded room at the student's school. Interviews were audio taped. Scaffolding was offered when students paused or asked for assistance. Each word problem was printed on top of a separate page, allowing more than sufficient space for students to write notes or perform written arithmetic. Students progressed through the word problems at their own pace, deciding when to turn the page and move on to the next problem. Also, the choice between solving problems using mental calculation strategies or standard algorithms was left to the student, as the instruction for the students was to solve each problem as they would if they were working on their own. Students were not allowed to use a calculator.

Data analysis

National tests were linked at the student level. A group of ten multistep arithmetic word problems were identified in the numeracy test, allowing for a correlation analysis involving numeracy, reading, and multistep arithmetic word problems. An ANOVA with post hoc Scheffe test was used to test for significant differences between groups.

Interview protocols were scored for overall success, the appropriateness of the situation model, reading, planning, calculation methods, students' technical skills, students' ways of checking answers, and causes for errors (model errors and calculation errors). Scaffolded workings were scored as correct or partially-correct workings in the same manner as the work of students who worked independently, as the purpose of this assessment was to allow students to complete the parts of the process that were within the reach of their competence. Later, protocols were coded. Codes were partially developed from findings in other research studies and partially developed from the protocol data after the protocols were scored. It should be noted that both scoring and coding were done in an interpretative manner based on students' protocols and worksheets (Clement, 2000), where students' actions were seen as strategic actions taken towards solving the word problem. That is, students actions were viewed as embedded in a local context in which the aim is to accomplish something. Further analysis was interpretive. The aim of the analysis was to understand what this "something" was and to search for patterns in individual students' work as well as in groups of students. As the protocol consisted both of independent work and dialogues between student and researcher, the researchers' scaffolding remarks and these dialogues also formed part of the local context for the students' work.

Results regarding the interplay between reading, numeracy, and solving multistep arithmetic word problems

Solving multistep arithmetic word problems had a strong positive correlation both to numeracy and to reading comprehension, though it was stronger for numeracy ($r = .859$, $p = .01$) than for reading ($r = .631$, $p = .01$). Students' reading levels explained 44 % of the variability in their scores on the multistep arithmetic word problem scale, $F(1,1262) = 836.950$, $p < .001$. All of these word problems can be solved using a combination of basic operations. On average, Norwegian students solved approximately half of the problems successfully. The sample mean score on this scale was 5.24 (of 10), $SD = 2.55$. For the 19 students who gave verbal protocols, the mean score was 5.16, $SD = 2.77$. The two mean scores are comparable in size; both means are approximately 52 % of the possible full score.

Table 2. Mean scores on numeracy test word problem scale

		Numeracy	
		Low	High
Reading	Low	LL 3.15 (1.77) $n = 499$	HL 6.37 (1.60) $n = 154$
	High	LH 4.20 (1.52) $n = 144$	HH 7.39 (1.65) $n = 472$

Table 2 displays the mean scores on the word problem scale on the national test for the four achievement groups in the national sample. A one-way ANOVA with a Scheffe post hoc test revealed significant differences between all groups, $F(3, 1260) = 558.454$, $p < .001$. However, given the criteria for assigning students to assessment groups, the differences in means on this scale reflect the strong positive correlation between reading, numeracy, and solving multistep arithmetic word problems. As can be expected, students with below-average numeracy scores were more similar to each other, considering multistep arithmetic word problems as well, than they were to students with above-average scores. Given the strong relationship to reading, students with higher reading scores outscored students with lower reading scores.

Of the different sub-constructs in the reading test[7], the ability to retrieve information had the highest positive correlation to solving multistep word problems, $r = .625$, $p = .01$. Word problems often comprise elements other than pure text, and, as expected, a higher correlation was found for non-continuous ($r = .640$, $p = .01$) than for continuous texts ($r = .530$, $p = .01$) (for a full correlation table, see appendix C). These patterns were consistent with findings for overall numeracy (Nortvedt, 2009).

Table 3. *Results for the verbal protocol sample*

Student	Achievement group[1]	Reading Percentile	Numeracy percentile	Verbal protocols[2]	WP2	Model error[3]	Calculation error[4]
Adam	LL	4	8	2	0.5	M4	R
Abraham	LL	14	8	3.5	0.5	M1	A
Benjamin	LL	14	50	2.5	0.5	M3	
Anita	LL	17	22	3	0,5	M2	A
Adeela	LL	23	38	2.5	1		
Anna	LL	17	50	4.5	1		
Berit	LL	35	11	2.5	0	M5	RA
Anders	LL	51	31	5.5	0.5		A
Bashir	LL	51	33	4	1		
Bente	LH	80	25	4	0,5		A
Alice	LH	85	49	3.5	0.5	M1	R
Billy	HL	35	85	6.5	0.5	M3	R
Benny	HL	38	87	7.5	1		
Burhan	HH	68	62	4.5	0.5	M3	
Anton	HH	76	81	6.5	1		
Britt	HH	80	78	6.5	0.5	M3	
Babette	HH	85	73	5.5	0.5	M1	R
Annette	HH	97	99	7.5	1		
Belinda	HH	100	98	8	1		

Notes. 1) Students are sorted into four achievement groups based on scores on national tests (see description of sample). 2) Scores on verbal protocols are raw scores on word problems. One score point is assigned for each correctly solved word problem. Partial scores are assigned for partially-correct solutions. No deduction is made for scaffolding. The maximum score is eight. 3) A classification of model errors can be found in appendix A. 4) R – retrieval errors, A – algorithmic errors.

Table 3 displays results for the 19 students who have given verbal protocols. Column 2 displays the achievement groups to which students were assigned. In columns 3 and 4, students' scores on national tests in reading and numeracy are given as the percentile at which they scored compared to the national sample. Most students were assigned to either the LL or the HH group. As in the national sample, the tendency was that high scores on one assessment could be observed with high scores on the other; and accordingly, lower scores on one assessment were mostly observed with lower scores on the second assessment. Only four of the 19 students were assigned to the LH and HL group. The verbal protocols (column 5) were scored as though students did not have access to scaffolding prompts and remarks. Still, the same general tendency can be seen: when students had high scores on one assessment, they often had high scores on the other two as well. The picture was slightly modified by the HL and LH groups.

Results of word problem 2 (WP2)

Column 6 in table 3 displays student scores on WP2, and in the last two columns, a classification of their errors can be found. WP2 can be solved by performing three multidigit calculations: first, to calculate the price of four t-shirts, then to subtract this sum from the total cost, and finally, to divide by three. The mathematical model for this word problem can be written as – [(79 x 4) – 364]: 3. WP2 was solved correctly only by seven of the 19 students. Most of the other students were assigned partial scores. Causes for errors can be connected both to the model (three students; see Understanding WP2 section), and to calculation errors (two students; see Solving WP2 section), or to a combination of model and calculation errors (seven students). Students with below-average reading scores appeared to make "double" errors more often; that is, they made a model error and one or more calculation errors.

Students with below-average numeracy scores made more calculation errors, as can be expected. Of the eleven LL and LH students, only three were assigned a full score for WP2; these three students were all scaffolded. Of the other eight LL and LH students, seven made at least one calculation error. These students, to a large extent, made errors that displayed a difficulty in understanding or mastering the basic operations (calculation error A; see Solving WP2 section).

Understanding WP2

If students fail to identify necessary information or relationships between numbers in the text, their model of the problem situation will be insufficient or incorrect, which will consequently prevent the student from solving the word problem correctly (Cummins et al., 1988; Reed, 1999) When ten of the 19 students made some sort of modelling error, it became crucial to investigate the nature and ramifications of the error. Protocol analysis revealed that students' errors were not linked to understanding or relating to the social context of the word problem. Buying clothes was a situation they recognised. They also correctly identified 364 kroner as the total amount of money that Aud paid. All errors were connected to socks and t-shirts.

Five different model errors were observed for WP2 (M1–M5[8]). Most of the mistakes were connected to linking the correct quantity to the correct object. Alice, Babette, and Abraham operated with three t-shirts as well as three pairs of socks (M1). One model error that was mainly displayed among HL and HH students was skipping the last step in the calculation; Britt and Burhan correctly identified 48 as the cost of three pairs of socks ("tre par sokker koster 48 kroner"), but failed to notice that they were asked for the price of *one* pair of socks (M3). Billy made the same assumption but made retrieval errors as well. The LL students demonstrated a wider variety of model errors. Abraham (M1) and Benjamin (M3) made model errors that were also made by high-achieving students. Anita's error (M2) might be caused by a lack of understanding of the keyword "each" and treating 79 as the price for all t-shirts. Adam's model (M4), on the other hand was difficult to judge; his mistake might be connected to the fact that he counted up to meet the sum of 364 and accidentally counted 20 six times, only to give 20 as the price of one sock. His understanding of WP2 might be appropriate but not transparent due to his lack of mathematical proficiency. Although difficult to judge, his model contained relevant elements and operations. The assumption that Berit made when she multiplied the difference between the total price and the cost of t-shirts by three (M5) is a model error that it is difficult to explain.

When turning the page and facing WP2, three students read the problem in silence. Anton probably read silently as well, as he stated the quantity (numbers) of socks and t-shirts before reading or rereading. Ten of the students only read through the problem text once during the phase of Understanding. Equal numbers of struggling and proficient readers could be found in this group. Six of them stated a plan before executing necessary calculations. All six students identified the appropriate first step of calculations according to their mental model. Plans typically were short, "da regner jeg 79 ganger fire [so I will do 79 times four]," and they mostly contained only the first step of calculations. Rarely did students state more holistic plans incorporating more steps. However, stating a plan is one of the characteristics of success; six of the seven students that succeed in solving WP2 all stated a plan during the Understanding phase. This might indicate that they recognized similarities between the mathematical situation embedded in the context of WP2 and other word problems previously experienced, and that the activity of solving WP2 was more of a routine activity than problem solving to them.

Scaffolding students' situation models

Two students (Adam and Bente) were scaffolded during Understanding. However, more students adjusted their first situation model during mediated work while solving WP2. Adam's (LL) scaffold was a neutral rephrasing of the question in the text. Bente's (LH) scaffold was a direct intervention towards her original situation model and plan. Britt (HH) received a similar scaffold when she was about to start solving. Both girls asked if they were to divide, Britt 79 by four and Bente three by 79. When confronted with their plan, they both appropriately identified 79 x 4 as the first step. Burhan, while proficient in handling numbers, revised and adjusted his SM several times during solving. Most of these interventions were directed scaffolds.

Solving WP2

About half of the students had an appropriate understanding of WP2 that served as a solid basis for working towards a solution. While the four HH students who solved WP2 correctly did so without scaffolding, the three LL and LH students did not.

Nine of the students made some sort of calculation error, even though all numbers were whole numbers and the multiplication and division was performed with relatively small numbers (3 and 4). According to the Norwegian curricula, multidigit operations are introduced and practiced in primary level, grades 5–7, including applying basic multidigit operations to solve embedded problems. Developing awareness towards recognising the appropriateness of a specific operation to a given situation is a primary goal (Ministry of Education and Research, 1996; Ministry of Education and Research & Norwegian Directorate for Teaching and Training, 2006).

Two main areas of difficulties could be observed: retrieving number facts from memory (R) and executing algorithms, including mental calculation strategies (A). Several students across proficiency levels relied on counting strategies, like group counting or counting up to produce number facts. Other students made retrieval errors, like Billy did when he thought that 4 x 80 equals 240 (R). Five students in total made recall errors (R). Retrieved numbers were often neighbouring number facts; for instance, 7 + 8 was recalled as 14, which is the sum of 7 + 7.

Students who struggled to apply both informal and standard algorithms (A) all had numeracy scores below the 32nd percentile. Misconceptions such as believing that if you carry a number it will always be 1, or that you cannot divide 130 by three because ”tre går ikke opp i en [three will not go into one]” were displayed through student talk. Some students used more simple informal methods, like repeated addition instead of multiplication, i.e. adding 79 four times to calculate 79 x 4 (Bashir and Bente). However, while such informal methods work well with small numbers, they will not be efficient when students are to execute complex calculations. Both Bashir and Bente signalled that the size of the numbers was what caused the difficulty. When it was suggested she work with 80, Bente knew that she could use repeated doubling. It might be that this calculation exceeded Bashir's independent capabilities. He, too, at first commented on the size of the numbers. Adding the numbers in each column he then wrote 2836 (see Fig. 2). The transcript displays what followed:

B: Twenty-eight plus thirty-six?

I: That is thirty-six kroner, but is that twenty-eight kroner?

B: No.

I: Is it that, these are tens are they?

B: Then it is twenty-eight tens.

I: Mm.

B: So I am to add, or?

Bashir knew that 2836 was not the final sum, and that he should do something with these numbers to arrive at the sum. However, he needed scaffolding to monitor his execution as well as to tie operations to the magnitude of the involved numbers in order to arrive at 316.[9]

Most students, at some point during the solving of WP2, stated what they planned to do next or identified the next step of calculations. This could be viewed as planning or monitoring one's own process. Such ”plans” were short, and they drove the calculation forward. Two short utterances made by Anne can serve as illustrations of how these statements kept students on track while working on solving WP2: ”[...] and

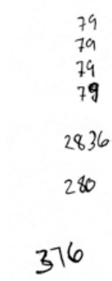

Figure 2. *Bashir's execution of 79* x *4*

then minus 364", which identified the next step, followed by "must put 364 first", which guided the performance. None of the students checked his/her answer. However, most of them (16) stated what they identified as their answer to demonstrate that "this is it". Typical outbursts were "16, et par sokker koster 16 kroner [16, a pair of socks costs 16 kroner]".

Scaffolding students' handling of numbers

Ten students were scaffolded at some point during the phase of Solving. The aim of the scaffolding prompts was not to have every student solve the problem correctly, rather, it was designed to assist the students by aiding and controlling some parts of the process and to allow them to demonstrate what they could do under these circumstances (Goldin, 2000; Wood et al., 1976). For this reason, solving WP2 correctly was outside the reach of some students, and they did not succeed. However, Bashir, who might not have progressed past 79 x 4 if he had been left to work independently, ultimately succeeded in solving WP2. To these ten students, scaffolding was mainly directed towards performing multidigit calculations and keeping on track. Some of the struggling students were provided with a number fact (direct intervention), which they needed in order to proceed towards solving WP2 and were unable to produce themselves. While HH students were proficient in handling numbers, eight of the eleven LL and LH students needed scaffolding on several occasions to produce number facts, to help monitor and keep on track while executing informal strategies as well as standard algorithms, or to choose the correct operation to perform the calculations needed for the next step. These students often also needed emotional support or verification of their actions by the researcher.

Discussion

Is WP2 a real problem in the sense that it required students to reason (Schoenfeld, 1985), or is it, as we claim, a stereotypical textbook problem? We will argue that the main obstacle to the students in solving WP2 was performing the multidigit calculations. However, to students who adjusted their model of the problem due to scaffolding during the problem solving, reasoning was involved in this process, and we will argue that WP2 was a real problem to them. Maybe WP2 also represented a problem to the many students who reread and elaborated on text content during Understanding. If this is the case, we will claim that the students in the protocol sample need more experience in solving multistep arithmetic word problems.

All but one student were assigned a full or partial score for WP2. Why do we then claim that they struggled and that too many failed? When students erred in solving WP2, this was caused both by modelling and calculation errors. Unlike in Cummins et al. (1988), students did not successfully solve the problem as they comprehended it. Seven of the ten students who made a model error also made calculation errors. Double errors were more common among students with below-average numeracy scores, as can be expected. At surface-level, 14 students formed appropriate or close-to-appropriate situation models, – [(79 x *a*) – 364]: *b*, where *a* and *b* are the number of t-shirts and pairs of socks, respectively. However, five of these students misrepresented one of the relationships between objects and quantities in the word problem text.

Moving below the surface-level of score points, students with below-average numeracy scores demonstrated that they were more at risk. Being proficient in reading probably helped as the two LH students formed more appropriate situation models for WP2. The LH students also demonstrated a higher mean score on the multistep arithmetic word problem scale on the national tests. Nevertheless, both LL and LH students still needed scaffolding to support their work towards solving WP2, and both groups made calculation errors. Ostad and Sorensen (2007) found that the use of counting strategies rather than number fact retrieval was one of the main differences between struggling and proficient students. It is worrying that both proficient and struggling students made recall errors or relied on counting to produce number facts. It might be that when a student needed to produce number facts during work on more complex problems, then focusing on the overall activity demanded too much of the student's attention, resulting in recall errors or difficulties remembering number facts.

A third worry is connected to the lack of mastery of standard algorithms. This is well-known from prior research both internationally

(Verschaffel et al., 2007) and among Norwegian students (Grønmo & Bergem, 2009). Many students counted or applied naive mental calculation strategies. It is thus possible that these students were not confident enough in their mathematical knowledge to use it in a flexible way. They instead "[sought] security in counting procedures which work promisingly in simple tasks [...]" (Gray & Tall, 1993, p. 6).

As indicated by both the national test data and the protocol data, proficiency in numeracy and reading comprehension are related. As the context given in WP2 is well-known and contains no irrelevant, extraneous, or hidden information, we will suggest that the ability to retrieve information from the text is the most relevant part of a student's reading competence when working on this specific word problem. While students of both samples had comparable overall scores in reading, students in the protocol sample for all achievement levels scored approximately one point higher in "retrieve information". In terms of effect size, this difference was between .43 and .7 (Cohen's d). This might have contributed to students' comprehensions of WP2.

While solving WP2, some students reread parts of the text. Prior research suggests that unsuccessful problem solvers who use surface-level strategies when comprehending word problems mainly focus on the numbers in the text when rereading. Consequently, they rarely discover instances in which their model representation differs from the problem given in the text (Hegarty et al., 1995). It could be argued that when students who apply deep-level strategies reread the problem, they then are more likely to refine and expand their situation model due to their strategy use and can consequently discover possible misinterpretations and errors. We also suggest that when some of the more struggling students did not reread, even though they experienced substantial difficulties solving WP2, it was because they detached their work from the context situation and worked on only using the numbers, as described in Verschaffel et al. (2007).

Conclusions

Three main areas of difficulties were found in the work of the 19 students solving WP2. To them the crucial parts of solving WP2 were, firstly, to form a situation model, and then to perform the calculations including retrieving number facts from memory. While one of the obstacles might be seen as more closely related to reading, the other two obstacles are closely related to students proficiency in number and number operations (Kilpatrick et al., 2001; Verschaffel et al., 2007).

The result regarding the verbal protocol data is in line with the main findings from the analysis of the national tests. Students reading level explained 44 % of the variability in scores on multistep arithmetic word problems identified in the national test, indicating a positive relationship between reading level and success in word problem solving. Although students at all achievement levels made model errors on WP2, the struggling students made these errors to a larger extent. It is not apparent why some of the high-achieving numeracy students failed to recognise the last step of calculation and gave 48 as the price of three pairs of socks. It could be that these students failed to notice that they were asked to find the price of one pair of socks, or it could be that real-world knowledge interfered, as socks are often sold in packs of three (Palm, 2008). Other students identified the necessary steps according to their model.

Relationships between reading and numeracy are recognised in the Norwegian mathematics curricula, and a special emphasis towards reading as well as problem solving is thus identified (Ministry of Education and Research & Norwegian Directorate for Teaching and Training, 2006). Given that reading has a positive correlation to numeracy in general (Nortvedt, 2009; Roe & Taube, 2006) and solving word problems in particular, it is probable that the students who struggle in numeracy and word problem solving are also the struggling readers. Hence, when students and teachers are working together on word problems, we suggest that special attention is given to recognising relationships between text elements, such as quantities and objects. This attention might be beneficial also to proficient students to strengthen their deep-level processing.

The two other main obstacles found regarding student difficulties in solving WP2 were retrieving number facts from memory, and performing basic operations. Students in all achievement groups experienced difficulties with retrieving number facts, but while above-average numeracy students typically retrieved wrong numbers, below-average numeracy students also in part relied on counting strategies to produce number facts. Even though primary school focuses on developing students' number concepts, increasing their knowledge of number facts, and performing basic operations as well as solving applied problems (Ministry of Education and Research & Norwegian Directorate for Teaching and Training, 2006), students in this study with below-average numeracy scores were not proficient enough to apply their knowledge on number and number operations when solving WP2; therefore needed "a lot of" scaffolding. For these students, applying basic operations was the main obstacle when trying to solve WP2. Consequently, we recommend that students be given the opportunity to practice these skills in classroom activities that allow them to expand their number concepts, to tie these concepts to their strategies for working on multidigit operations (Verschaffel et al., 2007), and to practice their strategies on meaningful word problems, not only as separate activities but also in integrated activities.

Acknowledgement

Data from national tests were provided by the Norwegian Ministry of Education and Research and the Norwegian Directorate for Education and Training.

References

Alexander, P. A., Sperl, C. T., Buehl, M. M., Fives, H. & Chiu, S. (2004). Modeling domain learning: profiles from the field of special education. *Journal of Educational Psychology*, 96 (3), 545–557.

Askew, M. (1999). It ain't (just) what you do: effective teachers of numeracy. In I. Thompson (Ed.), *Issues in teaching numeracy in primary school* (pp. 91–102). Buckingham: Open University press.

Bingobali, E. & Monaghan, J. (2008). Cognition and institutional setting. In A. Watson & P. Winbourne (Eds.), *New directions for situated cognition in mathematics education* (pp. 233–259). New York: Springer.

Clement, J. (2000). Analysis of clinical interviews: foundations and model viability. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 547–589). Mahwah: Lawrence Erlbaum.

Cobb, P., Stephan, M., McClaim, K. & Gravemeijer, K. (2001). Participating in classroom mathematical practices. *The Journal of the Learning Sciences*, 10 (1&2), 113–163.

Cohen, J., Cohen, P., West, S. G. & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah: Lawrence Erlbaum.

Cook, J. L. (2006). College students and algebra story problems: strategies for identifying relevant information. *Reading Psychology*, 27 (2–3), 95–125.

Cook, J. L. & Rieser, J.-J. (2005). Finding the critical facts: children's visual scan patterns when solving story problems that contain irrelevant information. *Journal of Educational Psychology*, 97 (2), 224–234.

Coquin-Viennot, D. & Moreau, S. (2007). Arithmetic problems at school: When there is an apparent contradicion between the situation model and the problem model. *British Journal of Educational Psychology*, 77, 69–80.
Corte, E. de & Verschaffel, L. (1991). Some factors influencing the solution of addition and subtraction word problems. In K. Durkin & B. Shire (Eds.), *Language in mathematics education. Research and practice* (pp. 117–130). Philadelphia: Open University Press.
Cummins, D. D., Kintsch, W., Reusser, K. & Wiemer, R. (1988). The role of understanding in solving word problems. *Cognitive Psychology*, 20 (4), 405–438.
Department of Education and Employment (DfEE). (1999). *The national numeracy strategy: framework for teaching mathematics from reception to year 6.* London: DfEE.
Geary, D. C. & Hoard, M. C. (2003). Learning disabilities in basic mathematics: difficulties in memory and cognition. In J. M. Royer (Ed.), *Mathematical cognition* (pp. 93–116). Greenwich: Information Age Publisher.
Goldin, G. A. (2000). A scientific perspective on structured, task-based interviews in mathematics education research. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 517–545). Mahwah: Laurence Erlbaum.
Gray, E. & Tall, D. (1993). Success and failure in mathematics: the flexible meaning of symbols as process and concepts. *Mathematics Teaching*, 142 (March), 6–10.
Greeno, J. G. (1997). Response: on claims that answer the wrong questions. *Educational Researcher*, 26 (1), 5–17.
Grønmo, L. S. & Bergem, O. K. (2009). Prestasjoner i matematikk [Mathematical outcomes]. In L. S. Grønmo & T. Onstad (Eds.), *Tegn til bedring. Norske elevers prestasjoner i matematikk og naturfag i TIMSS 2007* [It is going better now. Norwegian students' results in mathematics and science in TIMSS 2007] (pp. 49–111). Oslo: Unipub.
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.
Hughes, M., Desforges, C., Mitchell, C. & Carré, C. (2000). *Numeracy and beyond. Applying mathematics in the primary school.* Buckingham: Open University Press.
Hurst, C. (2008). Using task-based interviews to assess mathematical thinking of primary school students. In M. Goos, R. Brown & K. Makar (Eds.), *Proceedings of the 31st Annual Coference of the Mathematics Education Research Group of Australasia* (pp. 273–280). Brisbane: MERGA.
ILS/UiO. (2008). *Den nasjonale prøven i lesing på 8. trinn* [The national test in reading comprehension grade 8]. University of Oslo.

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

Lave, J. (1996). Teaching, as learning, in practice. *Mind, Culture, and Activity,* 3 (3), 149–164.

Lerman, S. (1998). Learning as social practice: an appreciative critique. In A. Watson (Ed.), *Situated cognition and the learning of mathematics* (pp. 33–42). Centre for Mathematics Education Research, University of Oxford.

Lesh, R. A. & Zawojewski, J. (2007). Problem solving and modeling. In F. K. J. Lester (Ed.), *Second Handbook of research on mathematics teaching and learning* (Vol. 2, pp. 763–804). Charlotte: NCTM.

Mayer, R. E. (2003). Mathematical problem solving. In J. M. Royer (Ed.), *Mathematical cognition* (pp. 69–92). Greenwich: Information Age Publishing.

Ministry of Education and Research. (1996). *Læreplanverket for den 10-årige grunnskolen* [Curriculum for the 10-year compulsory school in Norway]. Oslo: Nasjonalt læremiddelsenter.

Ministry of Education and Research & Norwegian Directorate for Teaching and Training. (2006). *Læreplanverket for kunnskapsløftet* [National Curriculum for Knowledge Promotion in Primary and Secondary Education and Training].

Nathan, M. J., Kintsch, W. & Young, E. (1992). A theory of algebra-word-problem comprehension and its implications for the design of learning environments. *Cognition and Instruction*, 9 (4), 329–389.

Nortvedt, G. A. (2009). The relationship between reading comprehension and numeracy among Norwegian grade 8 students. In M. Tzekaki, M. Kaldrimidou & C. Sakonidis (Eds.), *Proceedings of the 33rd conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 233–240). Thessaloniki: PME.

Norwegian Directorate for Education and Training. (2006). *Rammeverk for nasjonale prøver 2007* [Framework national tests 2007]. Retrieved June 1, 2010 from http://www.udir.no/upload/Nasjonale%20prøver/Rammeverk_for_nasjonale_prover_2007.pdf.

Ostad, S. A. & Sorensen, P. M. (2007). Private speech and strategy-use patterns – Bidirectional comparisons of children with and without mathematical difficulties in a developmental perspective. *Journal of Learning Disabilities*, 40 (1), 2–14.

Palm, T. (2006). Word problems as simulations of real-world situations: a proposed framework. *For the Learning of Mathematics*, 26 (1), 42–47.

Palm, T. (2008). Impact of authenticity on sense making in word problem solving. *Educational Studies in Mathematics*, 67(1), 37–58.

Pólya, G. (1957). *How to solve it* (second edition). London: Penguin books.

Ravlo, G. (2008). *Rapport. Nasjonal prøve i regning 8 trinn 2007* (Report. National test in numeracy grade 8 2007). Trondheim: Matematikksenteret (Norwegian centre for mathematics education).

Reed, S. K. (1999). *Word problems: research and curriculum reform*. Mahwah, NJ.: Erlbaum.

Roe, A. & Taube, K. (2006). How can reading abilities explain differences in math performance? In J. Mejding & A. Roe (Eds.), *Northern lights on PISA 2003 – a reflection from the Nordic countries* (pp. 129–142). Oslo: Nordic Council of Ministers.

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

Snow, C. E. (2002). *Reading for understanding: toward a R&D program in reading comprehension*. Pittsburg: RAND Education.

Thevenot, C., Devidal, M., Barrouillet, P. & Fayol, M. (2007). Why does placing the question before an arithmetic word problem improve performance? A situation model account. *Quarterly Journal of Experimental Psychology*, 60 (1), 43–56.

Thompson, I. (1999). Written methods of calculation. In I. Thompson (Ed.), *Issues in teaching numeracy in primary school* (pp. 169–183). Buckingham: Open University press.

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

Verschaffel, L., Greer, B. & Corte, E. de (2007). Number concepts and operations. In F. K. J. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 2, pp. 557–628). Charlotte: Information Age Publishing.

Vilenius-Tuohimaa, P. M., Aunola, K. & Nurmi, J.-E. (2008). The association between mathematical word problems and reading comprehension. *Educational Psychology*, 28 (4), 409–426.

Wood, D., Bruner, J. S. & Ross, G. (1976). The role of tutoring in problems solving. *Journal of Child Psychology and Psychiatry*, 17, 89–100.

Notes

1 Translated to numeracy by the author. The Norwegian Directorate for Education and Training have sometimes used the phrase "the ability to perform calculations". However, this is conceived as limiting with regards to the wide scope of the described sub-competences including, for instance, problem solving and solving applied problems.

2 See appendix B for an overview of the different phases in these models.

3 It should also be noted that standard algorithms can likewise be performed mentally; they do not necessarily involve using pen and paper.

4 National tests in reading comprehension and numeracy are developed by research groups at the University of Oslo and The Norwegian University of Technology and Science on behalf of the Norwegian Directorate for Education and Training.

5 For more information about framework and measures of test constructs, reliability, and validity, please see one or more of the following: (ILS/UiO, 2008; Norwegian Directorate for Education and Training, 2006; Ravlo, 2008).

6 For more information on test scores, see Ravlo (2008), ILS/UiO (2008), or Nortvedt (2009).

7 Retrieve, interpret, and reflect.

8 See also appendix A.

9 A full transcript of Bashir's work to find this sum is given in appendix D.

Appendix A

Model errors

M1: 3 t-shirts: – [(79 x 3) – 364) : 3
M2: 79 is the cost of four t-shirts: (364 – 79) : 3
M3: 48 is the cost of three pair of socks: – [(79 x 3) – 364)
M4: 20 is the cost of one sock: – [(79 x 3) – 364) : 6
M5: You multiply the difference by three to calculate the cost of socks: – [(79 x 3) – 364) x 3

Calculation errors

R: Difficulties retrieving number facts
A: Difficulties applying algorithms

Appendix B

Phases of problem solving

Author	**Understanding**		**Planning**	**Executing**	**Evaluating**
Mayer (2003)	Translation	Problem integration	Solution planning	Solution execution	
Polya (1957)	Understanding the problem		Devising a plan	Carrying out the plan	Reviewing/ extending the plan
Reed (1999)	Conceptual phase	Modelling phase	Planning	Executing or* solution actions	Checking*
Schoenfeld (1985)	Analyzing	Selecting knowledge	Making a plan	Carrying it out	Checking the answer against the question

Note. * Reed (1999) pp 8–10: Reed draws on Fuson, Hudson, and Pilar (1997).

Phases of problem solving in Verschaffel et al. (2000): This model is not linear; a possible cyclic process is suggested (p. 168). Phases include: Understanding, Modelling, Mathematical analysis, Interpretation and Communication

Appendix C

Correlations between reading sub-constructs, text types, and multistep arithmetic word problems

	Reading	**Retrieve**	**Interpret**	**Reflect**	**Continuous texts**	**Non-continuous texts**
Reading	–					
Retrieve	.910*	–				
Interpret	.891*	.771*	–			
Reflect	.839*	.660*	.646*	–		
Continuous texts	.924*	.727*	.916*	.824*	–	
Non-continuous texts	.938*	.957*	.752*	.742*	.734*	–
Multistep arithmetic WP	.631*	**.625***	.527*	.496*	.530*	**.640***
Sample mean	26.62	10.98	10.75	4.68	13.19	13.43
SD	8.23	3.69	3.27	2.25	4.21	4.64
Range	43	18	17	8	21	22

Note. * Correlation is significant at the .01 level (2-tailed)

Appendix D

Transcript, Bashir adding 79 four times under scaffolding. Translation by the author.

B: Yes. (Reads silently) (Writes 79 four times beneath each other) (Sits quietly)
I: Was this difficult?
B: Yes!
I: Yes! You have written 79 and then you have added it four times in total?
B: Mm.
I: Mm. And then it is difficult to find the sum?
B: Yes!
I: Yes?
B: Much too big numbers.
I: Are the numbers to big, yes?
B: To calculate with, for me, in my head.
I: Oh, like that. But. We could try to do it together.
B: Yes.
I: Yes? What would you have done, ehm (pauses).
B: Started with the last number.
I: Yes.
B: Eighteen.
I: Mm.
B: Eighteen plus eighteen it twenty-eight, twenty-nine (pause) thirty-six?
I: Mm. That is correct.
B: Fourteen, fourteen twenty-eight. (Writes 28 in from of 36)
I: Yes.
B: Twenty-eight plus thirty-six?
I: That is thirty-six kroner, but is that twenty-eight kroner?
B: No.
I: Is it that, these are tens are they?
B: Then it is twenty-eight tens.
I: Mm.
B: So I am to add, or?
I: Do you know how much money 28 tens are?
B: Eh. Two hundred or something.
I: Mm. Two hundred and?
B: Two hundred and eighty?
I: Mm.
B: So this is 280 then.
I: Yes, so just. You can just write it underneath, that is OK. [of topic – telling B the purpose of the setting is not to make him write in a style he would otherwise not employ]
I: No. (Deep breath). And then you must remember those, maybe
B: Multiply?
I: No, you should include
B: Add.
I: Add with 36 should you not?
B: (Unclear, whispering sound) (writes 316)

Guri A. Nortvedt

Guri A. Nortvedt works at the University of Oslo. Her main research interests are 1) relationships between language and (learning and doing mathematics) and 2) assessment studies within mathematics education. The research reported in this article was carried out at the Department of Special Needs Eduaction, University of Oslo.

g.a.nortvedt@isp.uio.no