CONTRIBUTION OF FLEXIBILITY IN DEALING WITH MATHEMATICAL SITUATIONS TO WORD-PROBLEM SOLVING BEYOND ESTABLISHED PREDICTORS

Laura Gabler & Stefan Ufer

LMU Munich

To solve mathematical word problems, students need to build appropriate models of the described situations, which they can describe with mathematical operations. Various studies have confirmed the importance of general cognitive skills, basic arithmetic skills, and language skills for word-problem solving. Beyond these, we investigate flexibility in dealing with mathematical situations, a new construct that describes the skill to re-interpret everyday situations from various perspectives. In a study with N = 113 second graders, an instrument to measure this flexibility construct has been developed and investigated. We find that the construct explains word-problem solving skills beyond the established predictors. Being able to flexibly re-interpret everyday situations may be beneficial for word-problem solving.

Students' skills to solve word problems diverge strongly. It has been well investigated particularly for additive one-step word problems, which predictors explain these differences. Additive one-step word problems are mathematical problems embedded in a verbally described everyday situation that can be solved with a single arithmetic operation (addition or subtraction) and do not contain irrelevant information (Verschaffel & De Corte, 1997). Recently, a new skill construct, *flexibility in dealing with mathematical situations*, has been proposed to support learning regarding additive (one-step) word problems (e.g., Gabler & Ufer, 2021). However, the role of this skill among other well-established predictors is unclear yet. This paper aims to fill this gap.

CURRENT STATE OF RESEARCH

Solving additive one-step word problems

Common theories on word-problem solving (e.g., Kintsch & Greeno, 1985) assume that learners construct two models when solving word problems: a situation model and a mathematical model. When learners encounter the text base (the verbal description of the mathematical situation), they construct a situation model based on this information. The situation model is the learner's internal, mental presentation of the given situation (Czocher, 2018). Learners then connect their situation model to mathematical concepts and transform it into a mathematical model. In the context of additive one-step word problems, students rely on conceptual knowledge on addition and subtraction, which needs to be available and activated to find an adequate mathematical model. For example, some word problems may refer to subtraction as the idea of "taking something away", while others may relate to a difference between two sets, making a connection to subtraction less salient. In literature, these different situations connected

2022. In C. Fernández, S. Llinares, A. Gutiérrez, & N. Planas (Eds.), *Proceedings of the 45th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 267-274). PME.

to addition and subtraction have often been classified into four different types ("semantic structures"; Riley, Greeno, & Heller, 1983): change, combination, comparison, or equalization of sets. Once the mathematical model is successfully constructed, learners can proceed with solving the word problem.

Individual predictors for solving additive one-step word problems

During this solution process, a number of individual predictors influence the students' performance when solving word problems (Daroczy et al., 2015).

For example, domain-general skills are discussed to predict students' word-problem solving skills. Solving word problems successfully depends on *general cognitive skills* (e.g., Jõgi & Kikas, 2016; Renkl & Stern, 1994), which may help with handling new, unfamiliar challenges (Warner et al., 2003). It is assumed that other domain-specific skills mediate the effects of general cognitive skills at least to some extent (Zheng, Swanson, & Marcoulides, 2011).

In addition, students' *language skills* play a role in word-problem solving (Daroczy et al., 2015). In particular, reading comprehension skills are considered crucial to decode the text base and derive an accurate situation model from this text (Vilenius-Tuohimaa, Aunola, & Nurmi, 2008). Indeed, studies have repeatedly identified reading comprehension skills as significant predictors of word-problem solving skills (e.g., Beal, Adams, & Cohen, 2010; Muth, 1984; Vilenius-Tuohimaa et al., 2008).

Besides such domain-general skills, students also need certain subject-specific, *basic arithmetic skills* for word-problem solving (Daroczy et al., 2015). In the context of additive one-step word problems, not only technical skills to solve additive equations are considered necessary, but also knowledge on number concepts (e.g., part-whole relationships, addition and subtraction as complementary operations; Renkl & Stern, 1994). This was confirmed by several studies, which report higher word-problem solving skills for students with higher basic arithmetic skills (e.g., Bjork & Bowyer-Crane, 2013 for grade 2; Muth, 1984 for grade 6).

Beyond these well-established predictors, it may play a role for students during word-problem solving, if they can deal flexibly with the given mathematical situation. This idea has first been suggested in the eighties and nineties (e.g., by Greeno, 1980; Stern, 1993) and conceptualized as a new skill construct within this project. Some learners struggle with constructing and mathematizing their situation model. In this case, it may help them to be able to add alternative perspectives to their situation model and further, to find mathematical operations that describe their situation model. In this sense, *flexibility in dealing with mathematical situations (FDMS)* can be defined as the skill to enrich their individual situation models of additive one-step word problems with further information, which is not verbalized in the text base. For example, learners could reinterpret compare problems as equalize problems: Additionally to the given description (e.g., "Susi has 2 marbles less than Max."), learners could imagine an equalization of Max's set: "If Max gets 2 more marbles, he has as many marbles as

Susi has.” (similarly suggested by Greeno, 1980). Another idea is to change the perspective on the situation: Instead of Susi’s perspective on the relation (“Susi has 2 marbles less than Max.”), learners could also add the perspective of Max: “Max has 2 marbles less than Susi.” (as suggested by Stern, 1993). Learners could integrate these different descriptions of the situation into a network of linked perspectives (Scheibling-Sève, Pasquinelli, & Sander, 2020). One basic assumption of this idea is that this skill complements the learners’ conceptual knowledge in word-problem solving. Learners with a high FDMS could then draw on the perspective that seems most helpful for them to find an adequate mathematical operation.

The suggested construct may be connected with other predictors. Handling new, unfamiliar challenges such as having to re-interpret a word problem (Warner et al., 2003) seems to be connected with general cognitive skills. Imagining different descriptions of mathematical situations is likely to be influenced by language skills and conceptual arithmetic knowledge. It is an open question, if FDMS can be operationalized and measured, and if this construct contributes to word-problem solving skills beyond the other mentioned predictors.

AIMS AND RESEARCH QUESTIONS

Although the idea behind FDMS has been suggested quite early, it has only recently been proposed as a skill construct. We investigated the following research questions:

RQ1: Is it possible to measure FDMS with sufficient reliability?

RQ2: How do general cognitive skills, basic arithmetic skills, and language skills explain inter-individual differences regarding FDMS?

RQ3: How does FDMS explain inter-individual differences in word-problem solving skills beyond general cognitive skills, basic arithmetic skills, and language skills?

Based on prior research, we expected general cognitive skills, basic arithmetic skills, and language skills to predict word-problem solving skills. Due to the reported theoretical foundations, we assumed FDMS to have a direct effect on word-problem solving skills beyond the other predictors.

METHOD

To answer the research questions, paper-and-pencil based tests were used in a cross-sectional study with second graders from ten classrooms in Germany (N = 113, 56 female, 57 male). The average age of the participating students was 7.7 years. There were 47% of students with German as their only family language, 19% with only non-German family language(s), and 34% of students with mixed family languages (at least German and another language). The study spans over two measurement times, between 6 and 21 days apart. On the first day, we measured the students’ language skills, their general cognitive skills, and their basic arithmetic skills. On the second day, we collected data on the students’ word-problem solving skills and their FDMS.

Instruments

Language skills were measured using the ELFE II reading comprehension test (Lenhard & Schneider, 2018). This test provides the opportunity to assess language skills based on reading fluency and accuracy with a larger sample. On average, the students achieved *M* = 45.03 raw points out of 111 total points with a standard deviation of *SD* = 15.5, which is in line with the average performance of the norm sample of the test. The reliability was excellent (α = .97).

General cognitive skills were measured by using the subscales "Similarities", "Classifications", and "Matrices" of the Culture Fair Intelligence Test "CFT 1-R" (Weiß & Osterland, 2013), which measure characteristics of general cognitive skills in a culturally fair, language-free setting. The reliability of the three subscales was acceptable (subscale "Similarities": α = .66; "Classifications": α = .73; "Matrices": α = .80). The three subscales were combined into one joint indicator. On average, the students scored *M* = 30.41 points out of 45 total points, with a standard deviation of *SD* = 5.98.

Basic arithmetic skills were measured with a test, which was developed for third graders within the LaMa project (Bochnik, 2017) and adapted for second graders in this study. Some of the tasks relate to technical skills in adding and subtracting numbers ranging until 100. Further tasks required conceptual knowledge, for example on the relationship between addition and subtraction (e.g., by asking for all four calculations that can be conducted with the numbers 7, 8, and 15). The reliability is satisfying (α = .82). On average, the students scored *M* = 7.49 points out of 16 total points with a standard deviation of *SD* = 3.80.

Word-problem solving skills were measured with a newly developed test ("word problem test"). This test was implemented in a multi-matrix-design: learners solved ten different word problems from a pool of 20 word problems based on the work of Stern (1993). The tasks systematically varied typical features (e.g., semantic structure), so that the whole range of possible types of additive one-step word problems was covered. The arithmetic and linguistic complexity of all 20 items was at a similar level. The data were scaled with a one-dimensional Rasch model. The WLE reliability of the instrument was .68. The average item difficulty was -1.04, indicating a relatively low difficulty of the test instrument.

Flexibility in dealing with mathematical situations (FDMS) was also measured with a test, which was newly developed within this project ("flexibility test"). The 20 items measuring FDMS were embedded into a story about twins, who tell the learners about a birthday party they visited. The learners were asked to decide, if the statements of the twins are equivalent or not (see Figure 1). The items emphasize different perspectives on mathematical situations in line with the ideas of Greeno (1980) and Stern (1993). For example, learners contrasted different perspectives on relations (as in Figure 1) or on actions (e.g., "Ben gave Alma 4 cards." vs. "Alma got 4 cards from Ben."). There were also items, in which two different semantic structures were contrasted (e.g., comparison: “There are 3 children more than adults at the party.” vs. equalization: “If 3 children leave, there are as many children as adults at the party.”). This facilitates the assessment of situational understanding and the skill to deal flexibly with such mathematical situations without the need to conduct mathematical operations.

Figure 1: Sample item for measuring FDMS

Statistical analyses

To answer the research questions, we estimated linear mixed models, taking into account that students were nested in classrooms. We calculated two models with the word-problem solving test score as a dependent variable, one with all independent variables (general cognitive skills, language skills, basic arithmetic skills, and FDMS) and one without FDMS for comparison. We also calculated a model with FDMS as a dependent variable to disentangle, which variables predict FDMS.

RESULTS

RQ1: One question was, if FDMS could be measured with sufficient reliability. The reliability of the flexibility test was satisfying ($\alpha = .80$). The participants scored $M = 14.04$ points on average out of 20 total points, with a standard deviation of $SD = 4.12$, showing that the test instrument was relatively easy.

RQ2: All three predictors significantly predicted FDMS (general cognitive skills: $F(106.36, 1) = 4.24$, $p = .042$, $\eta_p^2 = .04$; language skills: $F(108.97, 1) = 7.53$, $p = .007$, $\eta_p^2 = .06$; arithmetic basic skills: $F(107.74, 1) = 27.56$, $p < .001$, $\eta_p^2 = .20$). About 5.6% of the variance that was not explained by the predictors was attributable to class membership.

RQ3: As expected, language skills, basic arithmetic skills, and FDMS were significant predictors of word-problem solving skills (see Table 1). However, general cognitive skills were not significantly predictive beyond the other predictors. When including FDMS into the model, marginal R-square values increased substantially (without FDMS: marginal $R^2 = .38$; with FDMS: marginal $R^2 = .45$), and indeed, FDMS contributed significantly to variance explanation (see Table 1) with a medium to large effect size. This indicates that FDMS may explain differences in word-problem solving skills beyond the other variables. About 3.1% of the variance that was not explained by the predictors was attributable to class membership. Effect sizes for language skills and, in particular, for basic arithmetic skills reduced substantially when including FDMS, indicating that FDMS might mediate their effects on word-problem solving skills.

	Model 1			Model 2		
	F-value	p-value	η_p^2	F-value	p-value	η_p^2
General cognitive skills	0.63	.430	.00	0.00	.990	.00
Language skills	11.19	.001**	.10	6.21	.014*	.05
Basic arithmetic skills	21.83	<.001***	.19	7.99	.006**	.08
FDMS				14.67	<.001***	.12
Marginal R^2	.38			.45		

Table 1: ANOVAs based on linear mixed models with and without FDMS (*: $p < .05$; **: $p < .01$; ***: $p < .001$)

DISCUSSION

The main goal of this contribution was to investigate, if FDMS can be measured reliably, how it relates to established predictors of word-problem solving skills, and whether it contributes to variance explanation in word-problem solving skills beyond these established predictors (Daroczy et al., 2015).

Although the instrument turned out to be quite easy for second graders, it captured inter-individual differences in FDMS reliably. Currently, the instrument only assesses receptive FDMS by comparing *given* descriptions of situations. It would be important to include productive FDMS as well, for example, by asking students to construct alternative (written) descriptions of a mathematical situation (Gabler & Ufer, 2021). This would come closer to what we assume is required during word-problem solving.

Inter-individual differences in FDMS were related to all three established predictors. Beyond general cognitive skills, language skills contributed to variance explanation, which reflects the close connection of FDMS to language skills (e.g., Prediger & Zindel, 2017). The largest contribution came, however, from basic arithmetic skills. This is particularly remarkable, since the flexibility test does not address any arithmetic calculations, and instead focuses on situational understanding. This relation might go back to the part of the arithmetic test that covered conceptual understanding of addition and subtraction. Thus, developing FDMS could be connected closely to developing conceptual understanding of arithmetic operations in classroom practice, for example, by not only using situations from everyday contexts or manipulatives to reflect on mathematical structures, but also to compare and contrast different perspectives and verbal descriptions of these situations.

Regarding inter-individual differences in word-problem solving skills, general cognitive skills did not contribute under control of language skills and basic arithmetic skills. This contradicts some prior findings (e.g., Jõgi & Kikas, 2016). Possibly, the effect of general cognitive skills is fully mediated by the other variables. Replicating results from prior research (e.g., Bjork & Bowyer-Crane, 2013; Vilenius-Tuohimaa et al., 2008), language skills as well as basic arithmetic skills predicted word-problem solving skills. As expected, FDMS contributed to variance explanation beyond the other predictors. These results indicate that being able to re-interpret situations flexibly may support students' word-problem solving processes (e.g., Kintsch & Greeno, 1985) by allowing them to consider alternative perspectives on the described situation, which might be easier to mathematize. This means that the new construct has explanatory power for inter-individual differences beyond existing constructs. Supporting students to develop FDMS might be a way to support their word-problem solving skills.

Although the results on the new construct are promising, further research will have to clarify, if and how it can be fostered, and if this has effects on students' word-problem solving skills. Moreover, future research will need to consider how the construct can be conceptualized beyond additive situations, for example in the light of multiplicative situations, possibly including proportional relations.

References

Beal, C., Adams, N., & Cohen, P. (2010). Reading proficiency and mathematics problem solving by high school English language learners. *Urban Education, 45*(1), 58-74.

Bjork, I. M., & Bowyer-Crane, C. (2013). Cognitive skills used to solve mathematical word problems and numerical operations: A study of 6- to 7-year-old children. *European Journal of Psychology of Education, 28*(4), 1345-1360.

Bochnik, K. (2017). *Sprachbezogene Merkmale als Erklärung für Disparitäten mathematischer Leistung: Differenzierte Analysen im Rahmen einer Längsschnittstudie in der dritten Jahrgangsstufe*. Münster: Waxmann.

Czocher, J. A. (2018). How does validating activity contribute to the modeling process? *Educational Studies in Mathematics, 99*(2), 137-159.

Daroczy, G., Wolska, M., Meurers, W. D., & Nuerk, H.-C. (2015). Word problems: A review of linguistic and numerical factors contributing to their difficulty. *Frontiers in Psychology, 6*, 1-13.

Gabler, L., & Ufer, S. (2021). Gaining flexibility in dealing with arithmetic situations: A qualitative analysis of second graders' development during an intervention. *ZDM Mathematics Education, 53*(2), 375-392.

Greeno, J. G. (1980). Some examples of cognitive task analysis with instructional implications. In E. Snow, P.-A. Frederico, & W. E. Montague (Eds.), *Aptitude, learning, and instruction. Volume 2: Cognitive process analysis of learning and problem solving* (pp. 1-21). Hillsdale, NJ: Lawrence Erlbaum Associates.

Jõgi, A. L., & Kikas, E. (2016). Calculation and word problem-solving skills in primary grades: Impact of cognitive abilities and longitudinal interrelations with task-persistent behaviour. *British Journal of Educational Psychology, 86*(2), 165-181.

Lenhard, W., & Schneider, W. (2018). *ELFE II: Ein Leseverständnistest für Erst- bis Siebtklässler. Version II*. Göttingen: Hogrefe.

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review, 92*(1), 109-129.

Muth, K. D. (1984). Solving arithmetic word problems: Role of reading and computational skills. *Journal of Educational Psychology, 76*(2), 205-210.

Prediger, S., & Zindel, C. (2017). School academic language demands for understanding functional relationships: A design research project on the role of language in reading and learning. *Eurasia Journal of Mathematics, Science and Technology Education*, *13*(7b), 4157-4188.

Renkl, A., & Stern, E. (1994). Die Bedeutung von kognitiven Eingangsvoraussetzungen und Lernaufgaben für das Lösen von einfachen und komplexen Textaufgaben. *Zeitschrift für Pädagogische Psychologie, 8*(1), 27-39.

Riley, M. S., Greeno, J. G., & Heller, J. I. (1983). Development of children's problem-solving ability in arithmetic. In H. P. Ginsburg (Ed.), *The development of mathematical thinking* (pp. 153-196). New York: Academic Press.

Scheibling-Sève, C., Pasquinelli, E., & Sander, E. (2020). Assessing conceptual knowledge through solving arithmetic word problems. *Educational Studies in Mathematics, 103*, 293-311.

Stern, E. (1993). What makes certain arithmetic word problems involving the comparison of sets so difficult for children? *Journal of Educational Psychology, 85*(1), 7-23.

Verschaffel, L., & De Corte, E. (1997). Word problems: A vehicle for promoting authentic mathematical understanding and problem solving in the primary school. In T. Nunes & P. Bryant (Eds.), *Learning and teaching mathematics: An international perspective* (pp. 69-97). Hove: Psychology Press.

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

Warner, L., Alcock, L., Coppolo Jr., J., & Davis, G. (2003). How does flexible mathematical thinking contribute to the growth of understanding? In N. A. Pateman, B. J. Dougherty, & J. Zillox (Eds.), *Proceedings of the 27th Conference of the International Group for the Psychology of Mathematics Education* (pp. 371-378). Honolulu, Hawaii: PME.

Weiß, R. H., & Osterland, J. (2013). *Grundintelligenztest Skala 1-Revision: CFT 1-R*. Göttingen: Hogrefe.

Zheng, X., Swanson, H. L., & Marcoulides, G. A. (2011). Working memory components as predictors of children's mathematical word problem solving. *Journal of Experimental Child Psychology*, *110*(4), 481-498.