Word problems: a review of linguistic and numerical factors contributing to their difficulty

Gabriella Daroczy[1,2], Magdalena Wolska[2], Walt Detmar Meurers[2,3] and Hans-Christoph Nuerk[1,2,4]*

[1] Diagnostics and Cognitive Neuropsychology, Department of Psychology, Eberhard Karls Universität Tübingen, Tübingen, Germany, [2] LEAD Graduate School, Eberhard Karls Universität Tübingen, Tübingen, Germany, [3] Department of Linguistics, Eberhard Karls Universität Tübingen, Tübingen, Germany, [4] Knowledge Media Research Center, Tübingen, Germany

Word problems (WPs) belong to the most difficult and complex problem types that pupils encounter during their elementary-level mathematical development. In the classroom setting, they are often viewed as merely arithmetic tasks; however, recent research shows that a number of linguistic verbal components not directly related to arithmetic contribute greatly to their difficulty. In this review, we will distinguish three components of WP difficulty: (i) the linguistic complexity of the problem text itself, (ii) the numerical complexity of the arithmetic problem, and (iii) the relation between the linguistic and numerical complexity of a problem. We will discuss the impact of each of these factors on WP difficulty and motivate the need for a high degree of control in stimuli design for experiments that manipulate WP difficulty for a given age group.

Keywords: word problems, linguistics complexity, numerical complexity, text properties, difficulty

Word Problems

Introduction

Word problems (WPs) are part of the school curriculum and are taught at all levels of education. In WPs, relevant information is presented in the form of a short narrative rather than in mathematical notation (Verschaffel et al., 2000). Sometimes WPs specifically encode a quantitative relation between objects (Boonen et al., 2013). Many children from kindergarten through adulthood have severe difficulties in solving WPs (Nesher and Teubal, 1975; Riley et al., 1983; Lewis and Mayer, 1987; Hegarty et al., 1992; Verschaffel et al., 1992). Both linguistic and numerical complexity contributes to the difficulty in solving WPs. However, researchers have so far often focused on the one or the other aspect, depending on which field they come from. Even within the respective fields, linguistics, and numerical cognition, some aspects have been studied extensively, while others have been (strangely) neglected. For instance, we will see that semantics and discourse structures have been frequently studied in the context of WP complexity, but systematic syntactic manipulations are scarce. As regards numerical cognition, number properties like parity and magnitude as well as the type of mathematical reasoning have often been studied, but the type and the form of operations (e.g., carry-over effects) have not been investigated thoroughly in WPs, although they play an important role in current numerical cognition research (Moeller et al., 2011; Nuerk et al., 2011, 2015).

In this review, as researchers from the field of linguistics and the field of numerical cognition we have collaborated to provide a systematic overview of linguistic and numerical aspects relevant to solving WPs as well as their interaction. To capture a broad range of relevant facets

in the review, we extended our view of the relevant literature with systematic keyword searches in several databases (Web of Science, Ebsco, Google Scholar, ScienceDirect) including the following terms: WPs, story problems in combination with situational model, performance, consistency hypothesis, language processing, relational terminology, semantic influence, rewording, semantic cues, number size and type, working memory, text comprehension, computational errors, operations, position of unknown. In **Table 1**, we present selected linguistic, mathematical and general factors investigated in previous studies.

Individual Differences and Social Factors

Individual differences and social factors must be also considered in WP research (Fite, 2002). For example, in the PISA studies—often measured with WPs—, mathematics literacy is a commonly used notion (Stacey, 2012). "It is defined as an individual's capacity to identify and understand the role that mathematics plays in the world, to make well-founded judgments and to use and engage with mathematics in ways that meet the needs of that" (OECD, 2010). Unsuccessful WP solvers can experience negative social health and life outcome (Schley and Fujita, 2014). Even beyond social consequences, numerous studies focused on individual differences and group differences, such

as students with and without learning disabilities (Kingsdorf and Krawec, 2014), and children with and without developmental disabilities (Neef et al., 2003). Hegarty et al. (1995) distinguished domain-specific strategies that successful and unsuccessful problem solvers develop with practice and how these strategies account for individual differences in performance. Different students – e.g., individuals with calculation difficulty, or WP difficulty (Powell and Fuchs, 2014) – may struggle with different types of WPs. Besides domain-general capabilities like IQ, the role of domain specific knowledge and processes were investigated to get a complete account of problem solving, basic cognitive abilities; visual, reading skills, mathematical skills, and metacognitive abilities involved in the solution process. For example Boonen et al. (2014) and Oostermeijer et al. (2014) explored the role of spatial ability and reading comprehension in WP solving, since good WP solvers do not select numbers and relational keywords but create a visual representation (Boonen et al., 2013).

Social factors like schooling, teachers and peers also deserve consideration because the way of responding [e.g., De Corte et al. (1988)], the scoring criteria, the presence of illustrations next to the text [e.g., Berends and van Lieshout (2009)], or solution models used by the teachers influence WP performance considerably. School WPs also support stereotypical thinking: WPs do not

TABLE 1 | Selected linguistic, mathematical, and general factors investigated in previous studies.

Linguistic factors	Mathematical factors	General factors
Structure Structural complexity of basic quantitative properties (e.g., Number of letters, word and sentence length, proportion of complex words) (Searle et al., 1974; Nesher, 1976; Lepik, 1990) Vocabulary level (e.g., polysemous words, prepositional phrases, passive voice, clause structure; Spanos et al., 1988; Abedi et al., 1997; Abedi and Lord, 2001; Shaftel et al., 2006; Martiniello, 2008) Question wording/placing (Cummins et al., 1988)	**Property of numbers** Single digit (Lean et al., 1990) Multi digit (Haghverdi et al., 2012) Type of number [e.g., fraction: (Raduan, 2010), decimal number: problem size, role of number (De Corte et al., 1988)] Number magnitude [e.g., range of number smaller than 100: (Nesher, 1976)]	**Skills and social aspects** Social consequences (Schley and Fujita, 2014) Learning disabilities (Kingsdorf and Krawec, 2014) Successful/unsuccessful problem solvers (Hegarty et al., 1995) Calculation/word problem (WP) difficulties (Powell and Fuchs, 2014) Children/adults (De Corte et al., 1990; Hegarty et al., 1992)
	Required operation Addition subtraction (De Corte and Verschaffel, 1987) Multiplication division (De Corte et al., 1988) Given number (De Corte et al., 1990; Vicente et al., 2007)	**Categorization** Semantic structure of arithmetic WPs (Riley et al., 1983) Algebra textbook frequency (Mayer, 1981) Standard/non-standard WP (Jimenez and Verschaffel, 2014)
Semantics Linguistics verbal cues (van der Schoot et al., 2009) Phrasing in cue words (LeBlanc and Weber-Russell, 1996) Conceptual rewording (Vicente et al., 2007) Semantic/Object relation (**?**) Presence of distractor (Muth, 1992)	**Mathematical solution strategy** Counting from larger number (De Corte and Verschaffel, 1987) Position of the unknown (Garcia et al., 2006) Arithmetic fact retrieval (Orrantia et al., 2010) Number combination (Fuchs et al., 2009) Situation/Mental arithmetic strategy (Brissiaud and Sander, 2010)	**Solution strategies** Algebra WP/arithmetic WP (Koedinger and Nathan, 2004) WP solving theory-models (Kintsch and Greeno, 1985) Translation strategies (Hegarty et al., 1995) Spatial/visual representation (Boonen et al., 2013) Situation model (Thevenot et al., 2007)
	Relevance of information (Terao et al., 2004) **Numerical processes and representation** (MacGregor and Price, 1999; Goebel et al., 2014)	**Other aspects** Pedagogical factors (Lean et al., 1990) Socio-mathematics (Reusser, 1988) Stereotypes thinking (Yee and Lee, 1997) Real-word knowledge (Verschaffel et al., 1997) Response mode (De Corte et al., 1988)
Consistency effect (Lewis and Mayer, 1987) e.g., revisited: (Pape, 2003)		Computer tutors (Nathan et al., 1992) Computer simulation (Dellarosa, 1986)
Basic linguistics influence on numerical cognition (Lachmair et al., 2014)		
Working memory (Swanson et al., 1993)		

resemble problems in real-world situations (Yee and Lee, 1997). In addition, there is a strong tendency among both students and teachers to exclude real-world knowledge from their WP solution (Verschaffel et al., 1997), which is consistent with the observation that the problem solving process is also influenced by social cognitive and epistemic behavior settings (Reusser, 1988). Linguistic and pedagogical factors also affect children's understanding of arithmetic WPs (Lean et al., 1990). Students' beliefs about what doing and knowing mathematics means are rather different from the ideals (Jimenez and Verschaffel, 2014) and shaped by "socio-mathematical norms." Resulting differences in motivation seem to influence the strategies used to solve WPs (Gasco and Villarroel, 2014). In sum, both individual differences and social factors contribute to WP performance and deserve consideration.

Subcategories of Word Problems and Solution Strategies

Several different types of WPs—e.g., in the underlying mathematical structure or solvability— are often presented intermixed in one study without acknowledging the problem type. This is problematic. Different types of WPs are presented for various student groups, in different schools or different age groups. For example, Swanson et al. (2013) investigated the role of strategy instruction and cognitive abilities on WP solving accuracy. The mathematical WPs they used were: addition, subtraction, and multiplication without any further description of the problem type. However, the available literature has already shown that different categories of WPs may lead to different solution strategies and different error types. For instance, different semantic problem types result in different errors (Vicente et al., 2007) and have a different difficulty level (LeBlanc and Weber-Russell, 1996). Obviously, different scientific studies reporting results for different student or age groups cannot be easily compared to one another when they use different WP types; it cannot be determined whether differences should be attributed to group or study manipulation or differences in the used stimulus material. In the following, we outline the major distinctions discussed in the literature. Besides the difficulty level, WPs have been categorized with regard to various other attributes. Based on standard algebra text books, Mayer (1981) categorized WPs according to their frequency. Riley et al. (1983) created four groups based on the *semantic structure* of additive arithmetic WPs (change, compare, combine, equalize) and 18 further subcategories. For instance, the change problem –where there is a start, a change, and a result state –can be subdivided into three subcategories depending on which state is the unknown.

The *mathematical content* of WPs can also serve as a basis for categorization. Algebra WPs typically require translation into a mathematical formula, whereas arithmetic WPs are solvable with simple arithmetic or even mental calculation. In contrast to arithmetic WPs, algebraic reasoning WPs share the same numerals and signs (Powell and Fuchs, 2014) and the manipulation of those numbers and signals differs based on the question or expected outcome (Kieran, 1990). However, the distinction is not that straightforward, as in some cases both methods can be applied. For instance, in a study by Van Dooren et al. (2002), future secondary school teachers preferred the use of algebra

even when an arithmetical solution seemed more evident, and some future primary school teachers rather applied arithmetical methods. Computer-aided environments have been introduced for algebraic WPs (Reusser, 1993) to support learning on "getting the formalism" and the "equation" (Nathan et al., 1992) and to allow students to generate, manipulate, and understand abstract formal expressions for WPs. However, solution approaches are not easily dissociable between arithmetic and algebraic problems. If a WP is intended to be solved with an equation, in some cases a simple arithmetic approach is enough (Gasco et al., 2014). Under some circumstances, it is even easier to solve WPs via alternative arithmetic strategies than by deriving algebraic equations. US children perform better on a story problem if it is in a money context and the numbers involve multiples of 25 (Koedinger and Nathan, 2004). While the distinction between algebra and arithmetic WPs is important for investigation and evaluation, in this review we concentrate mainly on arithmetic WPs.

Standardized phrases and the *idea that every problem is solvable* are other important attributes of many, but not all WPs. Textbooks generally suggest implicitly that every WP is solvable and that every numerical information is relevant (Pape, 2003). They usually provide standardized phrases and keywords that are highly correlated with correct solutions (Hinsley et al., 1977; **?**). There are so-called non-standard WPs (Jimenez and Verschaffel, 2014) which can be non-solvable WPs or if they are solvable some have multiple solutions and may contain irrelevant data. In the recent literature, non-standard WPs are getting more and more attention (Yeap et al., 2005; Csikos et al., 2011). Children give a high level of incorrect answers to non-standard WPS because these seem to contradict their mathematics-related beliefs learned in the classroom. Reusser (1988) presented 97 first and second graders with the following sentence: "There are 26 sheep and 10 goats on a ship. How old is the captain?" and 76 students "solved" the problem using the numbers in the task. The rationale behind such studies is that always-solvable textbook problems with standardized phrases and including only relevant numerical information are hardly ecologically valid. Real-life WPs are not standardized, contain irrelevant information, and a solution may not always exist.

The above subcategories, which essentially characterize specific sets of WP properties, have a direct impact on human performance in WP. For space limitations, we cannot discuss the impact of all subcategories in detail, but we illustrate their impact on performance and strategies with two examples: (i) different subcategories can result in different errors, and involve different representations and processes. For example, a familiar misconception is that multiplication (Vergnaud, 2009) always makes the result larger (which is not true for $n < 1$), that division makes the results smaller, and that division always involves division of the larger number by the smaller, (ii) addition problems are strongly influenced (De Corte and Verschaffel, 1987) by the semantic structure (change, compare, combine). Carpenter et al. (1981) reported that the dominant factor in determining the children's solution strategy was this semantic structure. For instance, *Change* problems [cf. the classification of Riley et al. (1983)] require the child to find the difference between the two numbers given in the problem; their nature influences the strategies

children adopt. Riley et al. (1983) illustrates this with the following examples: Change 2: "Joe had eight marbles. Then he gave five marbles to Tom. How many marbles does Joe have now?" Change 3: "Joe had three marbles. Then Tom gave him some more marbles. Now Joe has eight marbles. How many marbles did Tom give him?" Almost all the children used a subtraction strategy (e.g., counting up) to solve Change 2. For Change 3 almost all the children used an addition strategy (e.g., counting down). In sum, the subcategories introduced in this section influence both performance and the choice of solution strategies.

Indeed, solution strategies have systematically been in the focus of WP research and addressed the following questions: how do children and adults solve WPs? Why do they make different errors and at which level of the solution process they do so? Which kind of semantic representation do they create of the WP? Which skills are necessary for the solution process? The first theories on WP solution processes (Kintsch and Greeno, 1985) have drawn on the text comprehension theories of Mayer (1982) and Van Dijk and Kintsch (1983). When solving problems, the solver first integrates the textual information into an appropriate situation model or a mental representation of the situation being described in the problem, which then forms the basis for a solution strategy. This approach was further applied by (Thevenot and Oakhill, 2005; Jimenez and Verschaffel, 2014; Kingsdorf and Krawec, 2014). An important foundation of those approaches is that solving WPs is not a simple translation of problem sentences into equations (Paige and Simon, 1966). Often both WPs and the corresponding numerical problems are done without language translation (Schley and Fujita, 2014). Several researchers have focused on abstraction as a reductive process involved in the translation process in the WPs. Nathan et al. (1992) argue that WPs solving is an exercise in text processing required for understanding the problem (Cummins et al., 1988), which is highly dependent upon language comprehension skills. Successfully solving WPs has been argued to require at least three distinct processes (Nesher and Teubal, 1975): (i) understanding and constructing the relation between text and arithmetic task, (ii) linguistic understanding of the WP itself, and (iii) solving the arithmetic tasks. Typically only the latter process is assumed to be shared with common arithmetic tasks. Many students can successfully solve common arithmetic tasks and they show good text comprehension skills. Yet they fail to solve WPs correctly. This suggests that other factors like solution strategies and building up a mental model of the task also play a major role for the WP performance.

Linguistic Complexity and Linguistic Studies

In linguistics, the notion of complexity is discussed under a range of perspectives, with particularly fruitful definitions grounded in research on language evolution (Nichols, 1990) and language acquisition (Bulté and Housen, 2012). Following the latter, it is useful to delineate linguistic complexity from propositional complexity (the amount of meaning to be expressed) and discourse-interactional complexity (the interaction of participants in discourse). This makes it possible to zoom in on linguistic complexity as the degree to which a text at hand is elaborated and varied (Ellis, 2003, p. 340). Linguistic complexity can be analyzed with respect to all aspects of the linguistic system: from the words and their lexical and morphological aspects, via the way these words can be combined in syntax to form sentences, to the text structure, and overall discourse. Languages differ with respect to where in the linguistic system complexification is supported. For example, English makes use of word order to encode grammatical functions, whereas agglutinative languages such as Hungarian or Turkish make use of a rich morphological inventory for this and other uses. The implication of linguistic encoding differences is twofold: first, the difficulty of WPs is language-specific, thus linguistic manipulation leading to increased WP complexity in one language may not have an effect in another, more complex language. Second, the performance of language learners on WPs presented in a foreign language may be affected by the differences between the learner's mother tongue and the language of the problem presentation. In the following two sections, we briefly summarize the main findings on aspects of linguistic complexity that affect performance.

Structural Factors

Studies on the relation between linguistic structure and student performance on WPs have considered complexity at the micro-level of word and sentence forms as well as at the macro-level of the discourse structure of the WP passage. Early approaches addressed structural complexity in terms of basic quantitative properties of the WP text, such as the number of letters, words, sentences, mean word, and sentence length, or the proportion of complex (long) words (Searle et al., 1974; Nesher, 1976; Lepik, 1990). More linguistically motivated variables have been investigated in the context of comprehension difficulties in WPs for language learners, for the most part learners of English. At the vocabulary level, comprehension difficulties which result in problem solving difficulties for English language learners may stem from the presence of unfamiliar (low-frequency) words, polysemous words, idiomatic or culturally specific lexical references. At the sentence structure level, factors that have been shown to play a role include noun phrase length, the number of prepositional phrases and participial modifiers, the presence of passive voice and complex clause structure such as relative, subordinate, complement, adverbial, or conditional clauses (Spanos et al., 1988; Abedi et al., 1997; Abedi and Lord, 2001; Shaftel et al., 2006; Thevenot et al., 2007; Martiniello, 2008).

At the discourse structure level, specifically in terms of discourse ordering, the correspondence between the order in which numerical data is presented in the WP and the order in which it can be used to solve it has been shown to be a major predictive variable. Order-consistent problems result in better performance (Searle et al., 1974). Better performance has also been observed for simpler question wording or placing the question before the text results (Cummins et al., 1988).

Semantic Factors

A single factor that is straightforwardly related to WP difficulty and that has been widely investigated is the presence or absence

of explicit verbal cues whose semantics hint at the expected operation and thus directly lead toward the solution. Verbal cues include words and phrases of different categories: conjunctions ("and" for addition), adverbs ("left," "more than," "less than" for subtraction), or determiners ("each" for multiplication). Eye tracking studies have shown that subjects tend to focus on linguistic verbal cues and perform translation directly to the mathematical operation (e.g., Hegarty et al., 1992; van der Schoot et al., 2009).

Because verbal cues so often lead to default mathematical interpretation (Nesher, 1976), even small differences in phrasing in cue words can cause significant changes in performance (LeBlanc and Weber-Russell, 1996). This is especially relevant for young children (Lean et al., 1990), who in the course of development connect words such as "join," "add," "get," "find," or "take away" with concepts such as *putting together*, *separating*, *giving away*, or *losing*. A problem can thus be reworded by adding verbal clues which make the semantic relations more salient so that the underlying mathematical relation is more explicit. For example, the WP "There are five marbles. Two of them belong to Mary. How many belong to John?" can be reworded as "There are five marbles. Two of them belong to Mary. The rest belong to John. How many belong to John?" [from Cummins (1991)]. This kind of conceptual rewording has been shown to be useful to improve children's performance on WPs (Vicente et al., 2007). Thus changes in wording can influence representation (De Corte et al., 1985).

Semantic or object relations between the objects described in the problem also relate to difficulty. Division problems usually involve functionally related objects (e.g., *tulips–vases*) and rarely categorically related objects (e.g., *tulips–daisies*; **?**). By contrast, addition for the most part involves categorically related objects. The correlation between object relations and mathematical operations has been argued to reflect a structural correspondence between semantic and mathematical relations (Bassok et al., 1998). For this reason, the semantic structure properties of a WP have been emphasized as a more important factor contributing to difficulty than the syntactic structure (Yeap and Kaur, 2001; **?**). Interestingly, an effect related to information load has been observed; the presence of content irrelevant to the core solution, i.e., the presence of numerical or linguistic distractors, results in higher error rates (Muth, 1992). De Corte and Verschaffel (1987) found that the semantic structure of WPs influences children's choice of mathematical solution strategy. In terms of the broader task context, the required or expected way of responding to the WP has a big influence, especially for the domain of multiplication and division with rational numbers as argued in De Corte et al. (1988); for example, whether students are expected to answer the problem numerically or if they only have to indicate the required operation, or whether they respond in an open way or with multiple choice.

Numerical Complexity and Numerical Studies

Arithmetic WPs have to be usually transformed mentally into an arithmetic problem and usually require an arithmetic solution

(**?**). This means transforming word and numbers into the appropriate operation (Neef et al., 2003). Since the arithmetic problem has to be solved in the end, numerical representations and arithmetic processes will also play an important role in the solution process. In numerical cognition, different models and representations have been proposed (e.g., Dehaene and Cohen, 1995; Nuerk et al., 2011). However, the problem here is that the literature on WP often seems (with some exceptions) to be largely in a parallel research universe to the literature on numerical cognition and arithmetic processes, so that standard models of numerical cognition are hard to apply on the existing literature. What is more, WP research on numerical factors is also affected by the scoring criteria; in some studies on WP solving, computational errors are neglected, because in many studies researchers consider a solution as correct as long as the solver has chosen the correct mathematical model (Verschaffel and De Corte, 1990). This is not the case in behavioral numerical cognition research, where the correct result is usually essential and RTs, accuracies, error types, and solution types are analyzed based on the arithmetic problem and result.

Numerical complexity can influence WP performance via at least three routes (see **Figure 1**):

(1) Direct route: WPs with more complex arithmetic structure are more difficult independent of linguistic complexity.

(2) Cognitive load: more complex arithmetic problems involve a higher cognitive load. For instance, carry problems are supposed to require more working memory resources. If the linguistic properties are also complex and the built-up of a mental model also requires more working memory resources, high arithmetic, and linguistic complexities could lead to over additive difficulties which could neither be explained by main effects of linguistic or numerical difficulty.

(3) Solution strategies: multi-digit numbers are harder to process than single-digit numbers (Nuerk et al., 2011; for a review) and arithmetic complexity usually increases with numbers of digits. Thevenot and Oakhill (2005) compared the influence of processing three-digit numbers and two-digit numbers on WP solution strategies. They showed that processing numerically more complex three-digit numbers facilitated alternative strategies by the participants. The authors suggested that higher work load and working memory led to this facilitation. For our review and the model in the revised manuscript, the important point is that they resort to less effortful strategies. Similar results were observed by Brissiaud and Sander (2010), who manipulated the size and order of the numbers and operation thus resulting in two different solution strategies: (i) situation-strategy WPs that are easy to solve with informal strategies, e.g., double-counting, derived number fact, or trial-error strategy, (ii) mental arithmetic-strategy WPs are "easy to solve with mental calculation, but only when the relevant arithmetic knowledge is used." Number magnitude and order determined which strategy was used most likely. In this review, we suggest that resorting to alternative easier strategies is not restricted to number magnitude, but could be used with any numerical variable that allows simpler solutions. For instance, if a number
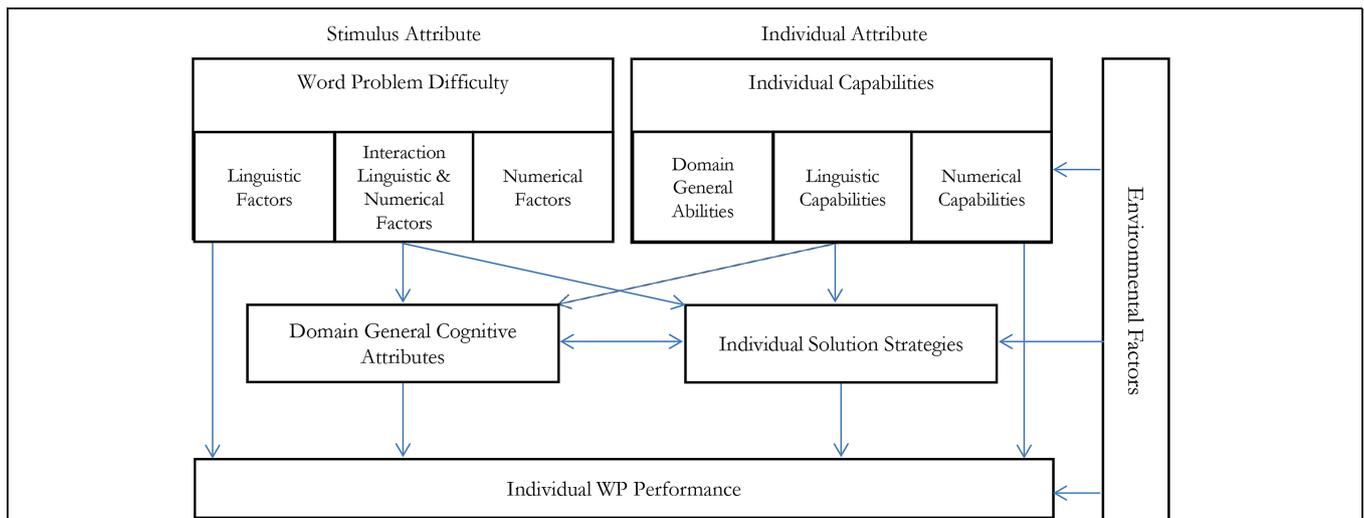
FIGURE 1 | This figure describes a possible theoretical process model of world problem solving based on this article and dissociating numerical and linguistic factors: Three general aspects are distinguished for predicting individual WP performance. Stimulus Attributes (WP difficulty), individual attributes (capabilities), and environmental factors (e.g., teaching). WP difficulty comprises linguistic factors (such as linguistic complexity of the WP text, Section 2 of this article), numerical factors (such as numerical difficulty of the numerical problem, Section 3), and their interaction (such as the relation between text and arithmetic problem, Section 4). Individual Capabilities can refer to linguistic and numerical capabilities and domain- general abilities such as individual working memory capacity. Stimulus attributes and individual attributes influence individual WP performance both directly and over two mediator variables. One mediator variable refers to domain-general attributes, such as cognitive load. Complex linguistic and numerical stimulus attributes can increase cognitive load and the impact of increased complexity may be overadditive, especially when the joint linguistic and numerical complexity exceeds the cognitive load of an individual. On the other hand, those domain-general attributes are influenced by individual capability. Cognitive load for an individual with high linguistic or numerical abilities may be lower for the same problem than for an individual with low linguistic or numerical abilities. The second mediator variable refers to specific solution strategies. If specific solution strategies can be applied to a particular WP problem, because the problem type allows this and because the individual knows the strategy, solution strategies can facilitate WP solving. Finally, environmental factors (e.g.: teaching, scoring system. . . etc.) influence individual capabilities, solution strategies, and also directly individual WP Performance.

bisection task were used in a text problem, we would also suggest that participants resort to easier strategies (e.g., checking the parities of the outer number), when the bisection problem gets more complex (e.g., larger interval, decade crossing etc.).

Nevertheless, some distinctions of numerical processes can be made in our review of the WP literature and are therefore proposed as an initial step in this review. Note that in our view this is not the end of the integration of numerical cognitive research and WP research, but rather just a beginning. For an overview of the investigation of specific numerical processes in current WP research, we suggest categorizing them into five categories:

 (i) the property of numbers (parity, single digit/multi digit, problem size, ties, type of number, role of the number, number magnitude),
 (ii) required operation (type, number)
(iii) mathematical solution strategies (larger number, place, automatic fact retrieval, position of the unknown),
 (iv) relevance of the information.
 (v) other numerical processes and representations

Number Properties

While some studies have shown an effect of numerical complexity, from a numerical cognition view it is surprising that actually the arithmetic complexity has rarely been systematically considered as an isolated factor in WPs, although it is frequently examined in other arithmetic problems or simply the description of numbers is missing, e.g., De Corte et al. (1990). For instance, *parity* attributes are rarely considered in WPs, although in children it influences task performance and strategy choice in arithmetic tasks. For instance, in the number bisection task (Is the middle number Y the exact mean of X and Z in X_Y_Z?), parity influences performance. Trials with unequal parities of X and Z are easier to solve than trials with equal parities (Nuerk et al., 2002). We suggested that this is due to a change in strategy. In trials with unequal parity (e.g., 21_25_28), it is impossible that the middle number is the mean, because the mean of numbers with unequal parity is not an integer number (and only integers were used in the experiment). Therefore, participants may change their strategy after they discovered unequal parities and may not compute further to find out whether the middle number is really the mean. A later fMRI study (Wood et al., 2008) corroborated this assumption. In the easier unequal parity ("impossible") condition, we observed more activation in the right ventrolateral prefrontal cortex, which is activated in cognitive set changes or when participants generate alternative solutions for a task. Thus, parity can influence performance and solution strategies in arithmetic. This seems not only the case in the bisection task, which is to our knowledge rarely used in WP research, but also in standard operations like addition and subtraction. A review by

Hines (2013) suggests that parity influences the difficulty of addition and subtraction, but not multiplication, and tasks containing odd numbers are more difficult than with even ones. Such parity effects have received little attention in WP research so far. Furthermore, it seems that most WPs, especially for children, contain *single-digit numbers*; e.g., each answer was in the range of 1–9, e.g., in Lean et al. (1990), or Powell and Fuchs (2014), only few use *multi-digit numbers* (Haghverdi et al., 2012). In Nesher (1976) the range of numbers is smaller than 100, contained division two-digit numbers into one-digit number.

Explanations why the studies have chosen specific numbers, e.g., mentioning problem size, are rare. De Corte et al. (1990) and Orrantia et al. (2010) controlled for the number of sentences; the size of the numbers given in the problems. In the study of van der Schoot et al. (2009) the final answers were between 14 and 40, included no fraction, no negative number, no numerical value twice, and none of the possible answers resulted in another. However, different types of numbers were presented in WPs in some more studies: (i) fraction (Raduan, 2010), (ii) whole number, (iii) decimal number (Haghverdi et al., 2012); and their effect has been rarely investigated. Koedinger and Nathan (2004) found an effect for decimal numbers: "however we also observed a smaller situation facilitation effect whereby story performance was better than word equation performance under certain conditions: namely dealing with decimal numbers."

The mixed use of single- and multi-digit numbers is problematic because in the last 15 years, numerous numerical cognition studies have shown that single-digit number processing cannot easily be generalized to multi-digit number processing, e.g., Nuerk et al. (2001); for reviews see Nuerk and Willmes (2005) and Nuerk et al. (2015). Nuerk et al. (2015) have identified 17 numerical effects linked to different numerical representation, which are specific for multi-digit number processing and which cannot be explained by single-digit number representations. Also even the same effects are different for single- and multi-digit numbers. For instance, Ashkenazi et al. (2009) have shown that the distance effect for two-digit numbers differentiates between dyscalculic and typically developing children. The sometimes seemingly arbitrary mix of single-digit and multi-digit number use in WP research is therefore not reasonable in our view given the state of numerical cognition research and the major differences between processing those different number types. The *role of a number* within an operation also influences WP complexity (De Corte et al., 1988). For example, in the case of addition the role means: addend, minuend or by multiplication: multiplicand, multiplier. One important finding from recent research on multiplication WPs is that children's performances are strongly affected by the nature of the multiplier whether, e.g., it is an integer, decimal larger than 1 or a decimal smaller than 1. On the other hand, the size of the multiplicand has little or no effect on problem difficulty. De Corte et al. (1988) stated that "two multiplication problems with the same mathematical, semantic, and surface structure but different in terms of the nature of the given numbers can elicit very distinct levels of problems difficulty." Indeed, this corresponds to recent findings that relatedness and consistency heavily influence the ease with which a multiplication problem can be solved cf. for relatedness (Domahs et al., 2006, 2007) and for consistency Verguts and Fias (2005).

Despite the major role of number properties in numerical cognition, number property has not been investigated extensively in the WPs (Fuchs et al., 2009). Nevertheless, numbers seem to play a major role. For instance, De Corte and Verschaffel (1986) observed that in their eye tracking study there was a relatively strong focus on the numbers in the problem. Twenty-five percent of the total solution time was spent in the two small number areas. However, major number properties of numerical cognitions research such as number magnitude are rarely systematically considered in WP research. In our view, more dialog between fields, – numerical cognition and WP research – seems necessary.

Required Operation

Carrying out operations are necessary steps in solving arithmetic WPs. Operations have been used extensively in WPs. Most errors seem to originate from people's failure to understand the language of WPs, i.e., the linguistic embedding of the calculation problem (Schumacher and Fuchs, 2012), and arithmetic computation errors themselves (Raduan, 2010; Kingsdorf and Krawec, 2014). Some errors may result from correct calculation performed on incorrect problem representation (Lewis and Mayer, 1987) and different operations may lead to different solution strategies. The most usual operation used in WP experiments are addition and subtraction (Carpenter et al., 1984; De Corte et al., 1988; Schumacher and Fuchs, 2012). Even the classification of Riley et al. (1983) was made for elementary addition and subtraction. Research in the 1980s and 1990s concentrated on how children learn to do one step addition and subtraction problems involving small whole numbers; see the review from Vicente et al. (2007). Later, the focus was more on the multiplication WPs or mixed WPs – e.g., Swanson (2004). Greer (1992) presented a framework categorization of multiplication and division WPs on the basis of the types of quantities involved (positive integers, fraction, and decimals) as models of situation. The semantic problem structure also influences the solution strategies for addition and subtraction.

Choosing the correct operation strongly depends on the type of the given numbers in the problem (De Corte et al., 1990). As already shortly outlined in above subsection on problem types, there is a huge body of research on what makes addition, subtraction, or multiplication problems difficult. Carry operations (e.g., $28 + 47$; the decade value 1 from the unit sum 15 has to be carried over to the decade sum) have long been known to make multidigit addition more difficult in children and adults; see Nuerk et al. (2015) for a review. However, solution strategies differ between children and adults – eye movement data suggest that in a choice reaction task elementary school children always compute and search for the correct results, while adults seem to also decide based on the rejection of the incorrect result. What is more, even within the carry operations at least three different cognitive processes can be identified for adults: unit sum calculation, carry detection, and carry execution (Moeller et al., 2011). Inability to execute one of these processes may lead to worse performance in carry problems in particular. Carry addition problems seem to

require larger working memory resources (Ashcraft, 1995; Furst and Hitch, 2000). If cognitive load/working memory demand is high, because both the linguistic and the numerical complexity of the WP are large, this may lead to over additive problems in the domain-general processing stages involved in WP solving — see **Figure 1**, for an elaboration. For multiplication, we know that relatedness, ties, whether a problem stems from the 0, 1, 2, 5, or 10 row (Josta et al., 2009), or consistency influence the difficulty of a multiplication problem (Domahs et al., 2006, 2007). Although such factors have been extensively studied in numerical cognition research, they are – to the best of our knowledge – rarely considered in WP research. Since we know that these factors make the arithmetic computation, which is part of the WP solution, this lack of consideration is again problematic in our view.

Mathematical Solution Strategies

Mathematical solution strategy variations have been studied extensively, and can be a function of linguistic factors like wording, semantic categories and propositions. However, how individuals come up with mathematical solution strategies can be also be influenced by numerical factors like number magnitude (Thevenot and Oakhill, 2005). Such variables, which are independent of other factors, make WPs harder and/or influence numerical representations, have rarely been studied. The position/place of the unknown variable has an effect on representation (Garcia et al., 2006). Even studies about working memory also investigated the position of the unknown variable (Swanson, 2004). The strategy of counting on from larger is easier if the bigger number is represented first (Wilkins et al., 2001). Even for adults: $4 + 2 = 6$, and $2 + 4 = 6$, which are mathematically equivalent, may psychologically imply different meanings (Kaput, 1979). The sequence of the numbers, e.g., whether a problem starts with the smaller or with the larger number (Verschaffel and De Corte, 1990), the position of the numbers and particular words (Schumacher and Fuchs, 2012) influence children's solution of elementary addition and subtraction problems. For example, in change problems children typically look for a specific number to begin with, depending on task features, like the first mentioned number (Lean et al., 1990; Wilkins et al., 2001), the type of problem (start or change set), and the size of the numbers (Verschaffel and De Corte, 1990).

Arithmetic fact retrieval is a well researched ubiquitous strategy in numerical cognition but less so in the domain of WPs. Orrantia et al. (2010) found that arithmetic fact retrieval is not limited to simple addition, but also possible in other tasks, such as single-digit arithmetic WPs. Fuchs et al. (2009) investigated so called "Number combination." This means simple arithmetic problems that can be solved via counting or decomposition strategies or committed to long term memory for automatic retrieval. Here, arithmetic fact retrieval had to be differentiated from other strategies on three levels: operational, items difficulty, and individual differences. These numerical factors influence solution strategies in arithmetic and WPs as well. Decomposition and counting require more working memory and therefore leave less resources for the built-up and maintenance of a text situations model. However, both individual and stimulus differences should also be considered. For instance, Grabner et al. (2009)

showed in an fMRI study that not only problem but also individual strategy choice contributed to fact retrieval processes when solving multiplications.

Information Relevance and Step-Wise Problem Processing

One relatively extensively studied factor in WPs is the relevance of the information. Individuals have to extract the relevant information from the text in order to carry out the correct solution. Secondary information distracts people from recognizing the underlying mathematical relations (Schley and Fujita, 2014). This extra information may also be presented in the form of an extra number or an extra operational step – one-step (i.e., one calculation step has to be performed) and two-step problems (i.e., two calculation steps have to be performed). Problem complexity increases with the addition of steps (Terao et al., 2004), as well as the addition of irrelevant information to the problem (Kingsdorf and Krawec, 2014) Presence of extraneous information and the need for an extra step reduced the accuracy of the students' solutions, because students believe that all of the numbers in a WP should be used. All other factors being kept constant, two-step problems are much more error-prone than one-step problems (Muth, 1992). However, it cannot be concluded that the reason for two-step problems being more difficult is arithmetic complexity, because in two-step problems, the WP has also become more difficult linguistically as it usually contains more phrases and semantic distractors.

Other Numerical Processes and Representation

Several other numerical processes and representations have not been investigated in WPs. For instance, as shortly outlined above, one major factor in simple calculation problems, which can be studied in isolation, is the presence or absence of a carry operation. Children and adults take longer and commit more errors when computing the solution to a sum for which adding the units leads to a change in the number of 10s (e.g., $14 + 9 = 23$; Furst and Hitch, 2000; Deschuyteneer et al., 2005) than when it does not (e.g., $11 + 12 = 23$). This effect is known as the carry effect; in carry problems, a one needs to be carried from the unit slot to the decade slot. The carry effect is influenced by various processes, but even by language structure (Goebel et al., 2014). Language influences on the difficulty of the numerical computations within a WP have to our knowledge not been studied. Other central topics of numerical cognition such as, e.g., number and symbol sense contribute to WP solving are also open questions (MacGregor and Price, 1999). We have chosen some selected variables/factors, which have been investigated in the WP research.

Connecting Linguistic and Mathematical Factors

There are so many linguistic influences on numerical cognition and arithmetic that this justifies a special issue like this. For instance, number word structure seems to play an essential

role. Children growing up with regular number word structure usually perform better in variety of numerical tasks from basic verbal counting up to arithmetic, e.g., Miller et al. (1995) or Dowker et al. (2008). In addition, the consistency of the order of the number word system and the Arabic number influences transcoding (Zuber et al., 2009; Pixner et al., 2011a; Imbo et al., 2014) number comparison (Nuerk et al., 2005; Pixner et al., 2011b; Klein et al., 2013; Moeller et al., 2014) calculation (Goebel et al., 2014); see also (Brysbaert et al., 1998; Colomé et al., 2010). In addition, reading direction influences numerical processes like the SNARC effect (Shaki et al., 2009; Fischer and Shaki, 2014); see Goebel et al. (2011) for reviews. Finally, grammatical and syntactic properties of elementary number words influence early number acquisition (Sarnecka, 2013) and spatial-numerical representations (Roettger and Domahs, 2015). The linguistic influence on numerical cognition is hardly debatable any more. In fact, Lachmair et al. (2014) argue for a connection of language and words, O´Neill (2013) states that the link between language and mathematics might originate from the same roots, and "required abilities are not that split up as we think," and MacGregor and Price (1999) also argue that between language and mathematics in WPs there is deep connection: "that the cognitive ability that drives symbol processing is the connection between language and maths." Nevertheless, systematic variation of both linguistic and numerical factors in WPs is scarce – though Bassok et al. (1998) already found that semantic relations between objects in the text of mathematical WPs were highly positively correlated with arithmetic operations that took these objects as arguments. Neural correlates of visualization and verbalization during arithmetic WP study also suggest that mental arithmetic in WPs is influenced by language processing (Zarnhofer et al., 2013).

Word problems require some connection between linguistic and mathematical understanding by the very nature of the task, because at least children do not have a repertoire of "highly automatized schemata" for representing the different problem types (Garcia et al., 2006). Therefore, it is not surprising that children make more errors when solving WPs compared to number problems (Geary, 1996; Koedinger and Nathan, 2004). Children are able to solve several types of addition and subtraction problems before they start formal schooling (De Corte and Verschaffel, 1987; Lean et al., 1990), and understand numerical concepts before seeing WPs in their curricula (Garcia et al., 2006). Therefore, most studies implicitly assume that problem solvers always have the necessary basic arithmetic skills, even in the case of children. This may lead to the misconception that numbers may play a lesser role than they actually do and factors other than computational skills are a major source of difficulty with WPs (Nesher, 1976; Reusser, 1993). In this aspect, it is also important to note that difficulties in solving WPs have been reported that could be neither attributed to the lack of general reading comprehension skills nor to the lack of general mathematical skills (Hegarty et al., 1995). Nevertheless, linguistics and numerical factors are usually not independently manipulated in WPs and not even dissociated by other means (e.g., regressions). What is more, their interaction is rarely studied [for an exception, see Verschaffel and De Corte (1990)].

Lexical Consistency Effect

One of the few frequently studied factors examining the relation between text and arithmetic problems is lexical inconsistency. Some WPs contain linguistic markers as "less" or "more." In the direct translation strategy (Hegarty et al., 1995) students simply associate "less" with subtraction and "more" with addition. They search for linguistic markers and keywords. In the problem model strategy, they construct a mental model of the problem and plan their solution on the basis of this model. Successful learners are more likely to employ the problem model strategy; they focus more on variables names and relational terms and successful problem solvers re-read the text less frequently (Pape, 2003) in the eye-tracking studies. Unsuccessful learners, on the other hand, seem to rely on the direct translation strategy; they focus on numerals and on relational terms, and linguistics markedness in the (Hegarty et al., 1992) eyetracking study. This leads to wrong solutions in lexically inconsistent texts, where "more" is associated with subtraction and "less" with addition. To give an example for lexical inconsistency, consider the following WP adapted from Boonen et al. (2013) "At the grocery store, a bottle of olive oil costs 7 €. That is 2 € more than at the supermarket. How much will [a bottle of olive oil] cost in the supermarket?" The anticipated difficulty in comprehension and finding the correct solution is due to the fact that the adverb "more" evokes the concept of addition, but the correct solution is not $7 + 2$ but $7 - 2$, given the way the text is organized. Verschaffel et al. (1992) found such a reaction time consistency effect for children but not for adults. Nesher (1976) and Lean et al. (1990) obtained similar results in experiments with groups of non-disadvantaged children and students, showing that linguistic semantic consistency with respect to the required mathematical operation is an important determinant of task difficulty. Inconsistent language results in a high error rate and longer response time (Hegarty et al., 1992), even in Verschaffel (1994) retelling one-step compared WPs showed a strong evidence for the consistency hypothesis. Students made ∼13% more reversal errors on inconsistent than on consistent language problems and the difficulty of comprehending inconsistent-language problems were increased when the correct arithmetic operation was an increase. However, the literature is inconsistent if the consistency effect is present in both students and children. Children find it easier to convert the relation term "more than" into subtraction operation than the relational term "less than" into an addition operation (Lewis and Mayer, 1987; Verschaffel et al., 1992; Pape, 2003; van der Schoot et al., 2009).

When neither reading comprehension nor arithmetic skills alone can explain failure to solve WPs, a possible explanation is that linguistic complexity and numerical complexity rely on the same resources (e.g., working memory). The premise is that there is not an absolute atomic concept of difficulty for WPs. Rather; there are multiple linguistic and numerical factors which contribute to a problem's complexity. It is a combination of these factors that might make a problem additively more or less difficult because they exert demands on more general resources like working memory. Generally, problem solving performance is related to the ability of reducing the accessibility of no target and irrelevant information in the memory

(Passolunghi and Siegel, 2001). Working memory contributes to early arithmetic performance, and studies also show that this extends to WP solving (Lee et al., 2004) due to semantic memory representation "less than" which is more complex than "more than." Changes in the structure of the text has more demand on the working. It has been suggested that WPs in general are related to working memory (Swanson et al., 1993). This will probably also be influenced by instruction specifying how participants have to solve a WP, and the method of evaluation, and scoring system. In Van Dijk and Kintsch's (1983) model of reading comprehension, working memory is used to keep a number of text propositions active simultaneously. In particular, working memory has been related to each single component mentioned above, such as text-problem relation, the linguistic complexity, and the arithmetic complexity.

Future Direction, Open Questions

Word problem difficulty is influenced by the complexity of linguistic factors, numerical factors, and their interrelation. To better understand the difficulty of WPs, it would be desirable to manipulate such variables and their interaction following the principle of isolated variation. To support a systematic investigation, the variables to be manipulated also need to be discussed against the backdrop of the relevant conceptual and empirical issues in the underlying fields, linguistics, and numerical cognition. This has too rarely been the case in the past. For instance, in the earlier studies on algebra WPs, the linguistic cues are of mixed categories (adverbs, verbs, nouns, etc.) and the effect of the complexity of syntactic structures is not taken into account. Similarly, numerical complexity like basic number properties (e.g., magnitude, place-value processing for multi-digit numbers) or the complexity of underlying arithmetic computations (e.g., carry effects for addition, relatedness, or consistency effects for multiplication) are often neglected. WP research would be well advised to take into account the foundational categories, properties and findings of both numerical cognition and linguistics when it examines which WPs are difficult for which groups and why. Not only the main effects of numerical and linguistic complexity should be studied, but also their interaction. To make

the relevant aspects explicit, **Figure 1** sketches an overall process model of WP solving.

The joint investigation of linguistic and numerical processes also needs to take into account joint moderator variables such as working memory in order to explore the possible interactions between them. Since working memory affects all components of complexity of a WP, the difficulties triggered may not be simply additive, but also interactive. The resolution of linguistic and numerical difficulties may rely on the same processing stages and resources (Sternberg, 1969). To investigate this, more collaboration between linguists and numerical cognition researchers would be desirable.

Finally, we suggest a differential-psychological approach to WP research. Different students may have a problem with different types of WPs. Linguistically rather weak students may have problems with linguistically complex WPs, and arithmetically rather weak students with arithmetically complex problems. Undifferentiated presentation of WPs in experiments will not provide sufficient information about which skills and processes an individual child should practice. Only with such differentiation on an item level (as regards linguistic and numerical complexity and their interrelation) and on an individual level (as regards linguistic and numerical skills and general cognitive abilities) will it be possible to understand why a particular child has its individual difficulties with particular WP types. Such an understanding, however, is essential to promote tailored learning of one of the most difficult arithmetic problem types that students encounter in school.

Acknowledgments

References

Abedi, J., and Lord, C. (2001). The language factor in mathematics tests. *Appl. Meas. Educ.* 14, 219–234. doi: 10.1207/S15324818AME1403_2

Abedi, J., Lord, C., and Plummer, J. (1997). *Final Report of Language Background as a Variable in NAEP Mathematics Performance.* Tech. Rep. No. 429. Los Angeles, CA: University of California, Center for the Study of Evaluation/National Center for Research on Evaluation, Standards, and Student Testing.

Ashcraft, M. H. (1995). Cognitive psychology and simple arithmetic: a review and summery of new directions. *Math. Cogn.* 1, 3–34.

Ashkenazi, S., Rubinsten, O., and Henik, A. (2009). Attention, automaticity, and developmental dyscalculia. *Neuropsychology* 23, 535–540. doi: 10.1037/a0015347

Bassok, M., Chase, V. M., and Martin, S. A. (1998). Adding apples and oranges: alignment of semantic and formal knowledge. *Cogn. Psychol.* 35, 99–134. doi: 10.1006/cogp.1998.0675

Berends, I. E., and van Lieshout, E. C. D. M. (2009). The effect of illustrations in arithmetic problem-solving: effects of increased cognitive load. *Learn. Instr.* 19, 345–353. doi: 10.1016/j.learninstruc.2008.06.012

Boonen, A. J. H., Van Der Schoot, M., Van Wesel, F., De Vries, M. H., and Jolles, J. (2013). What underlies successful word problem solving? A path analysis in sixth grade students. *Contemp. Educ. Psychol.* 38, 271–279. doi: 10.1016/j.cedpsych.2013.05.001

Boonen, A. J. H., Van Wesel, F., Jolles, J., and Van Der Schoot, M. (2014). The role of visual representation type, spatial ability, and reading comprehension in word problem solving: an item-level analysis in elementary school children. *Int. J. Educ. Res.* 68, 15–26. doi: 10.1016/j.ijer.2014.08.0010883-0355

Brissiaud, R., and Sander, E. (2010). Arithmetic word problem solving: a situation strategy first framework. *Dev. Sci.* 13, 92–107. doi: 10.1111/j.1467-7687.2009.00866.x

Brysbaert, M., Fias, W., and Noël, M. P. (1998). The Whorfian hypothesis and numerical cognition: is "twenty-four" processed in the same way as "four-and-twenty"? *Cognition* 66, 51–77. doi: 10.1016/S0010-0277(98)00006-7

Bulté, B., and Housen, A. (2012). "Defining and operationalising L2 complexity," in *Dimensions of L2 Performance and Proficiency – Investigating Complexity, Accuracy and Fluency in SLA,* eds A. Housen, F. Kuiken, and I. Vedder. (Amsterdam: John Benjamins), 21–46. doi: 10.1075/lllt.32.02bul

Carpenter, T. P., Hieben, J., and Moser, J. (1981). The effect of problem structure on first-grader's initial solution processes for simple addition and subtraction problems. *J. Res. Math. Educ.* 2, 27–39. doi: 10.2307/748656

Carpenter, T. P., Moser, T. P., and James, M. (1984). The acquisition of addition and subtraction concepts in grades one through three. *J. Res. Math. Educ.* 15, 172–202. doi: 10.2307/748348

Colomé, A., Laka, I., and Sebastiaen-Galles, N. (2010). Language effects in addition: how you say it counts. *Q. J. Exp. Psychol.* 63, 965–983. doi: 10.1080/17470210903134377

Csikos, C., Kelemen, R., and Verschaffel, L. (2011). Fifth-grade students' approaches to and beliefs of mathematics word problem solving: a large sample Hungarian study. *ZDM* 43, 561–571. doi: 10.1007/s11858-011-0308-7

Cummins, D. D. (1991). Children's interpretations of arithmetic word problems. *Cogn. Instr.* 8, 261–289. doi: 10.1207/s1532690xci0803_2

Cummins, D. D., Kintsch, W., Reusswer, K., and Weimer, R. (1988). The role of understanding in solving word problems. *Cogn. Psychol.* 20, 405–438. doi: 10.1016/0010-0285(88)90011-4

De Corte, E., and Verschaffel, L. (1986). Eye-movement data as access to solution processes of elementary addition and subtraction problems. *Paper Presented at the Annual Meeting of the American Educational Research Association,* San Francisco.

De Corte, E., and Verschaffel, L. (1987). The effect of semantic structure on first graders' strategies for solving addition and subtraction word problems. *J. Res. Math. Educ.* 18, 363–381. doi: 10.2307/749085

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

De Corte, E., Verschaffel, L., and Pauwels, A. (1990). Influence of the semantic structure of word problems on second graders eye movements. *J. Educ. Psychol.* 82, 359–365. doi: 10.1037/0022-0663.82.2.359

De Corte, E., Verschaffel, L., and Van Coillie, V. (1988). Influence of number size, problem structure, and response mode on children's solutions of multiplication word problems. *J. Math. Behav.* 7, 197–216.

Dehaene, S., and Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Math. Cogn.* 1, 83–120.

Dellarosa, D. (1986). A computer simulation of children's arithmetic word-problem solving. *Behav. Res. Methods Instrum. Comput.* 18, 147–154. doi: 10.3758/BF03201014

Deschuyteneer, M., De Rammelaere, S., and Fias, W. (2005). The addition of two-digit numbers: exploring carry versus no-carry problems. *Psychol. Sci.* 47, 74–83.

Domahs, F., Delazer, M., and Nuerk, H.-C. (2006). What makes multiplication facts difficult: problem size or neighbourhood consistency? *Exp. Psychol.* 53, 275–282. doi: 10.1027/1618-3169.53.4.275

Domahs, F., Janssen, U., Schlesewsky, M., Ratinckx, E., Verguts, T., Willmes, K., et al. (2007). Neighborhood consistency in mental arithmetic: behavioral and ERP evidence. *Behav. Brain Funct.* 3, 66. doi: 10.1186/1744-9081-3-66

Dowker, A., Bala, S., and Lloyd, D. (2008). Linguistic influences on mathematical development: how important is the transparency of the counting system. *Philos. Psychol.* 21, 523–538. doi: 10.1080/09515080802285511

Ellis, R. (2003). *Task-based Language Learning and Teaching.* Oxford: Oxford University Press.

Fischer, M. H., and Shaki, S. (2014). Spatial associations in numerical cognition – from single digits to arithmetic. *Q. J. Exp. Psychol.* 67, 1461–1483. doi: 10.1080/17470218.2014.927515

Fite, G. (2002). Reading and math: what is the connection? A short review of the literature. *Kansas Science Teacher* 14, 7–11.

Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., et al. (2009). Remediating number combination and word problem deficits among students with mathematics difficulties: a randomized control trial. *J. Educ. Psychol.* 101, 561–576. doi: 10.1037/a0014701

Furst, A., and Hitch, G. J. (2000). Separate roles for executive and phonological components of working memory in mental arithmetic. *Mem. Cogn.* 28, 774–782. doi: 10.3758/BF03198412

Garcia, A., Jimenez, L., and Hess, S. (2006). Solving arithmetic word problems: an analysis of classification as a function of difficulty in children with and without arithmetic LD. *J. Learn. Disabil.* 39, 270–281. doi: 10.1177/00222194060390030601

Gasco, J., and Villarroel, J. D. (2014). The motivation of secondary school students in mathematical word problem solving. *Electron. J. Res. Educ. Psychol.* 12, 83–106. doi: 10.14204/ejrep.32.13076

Gasco, J., Villarroel, J. D., and Zuazagoitia, D. (2014). Different Procedures for Solving Mathematical Word Problems in High School. *Int. Educ. Stud.* 7, 77. doi: 10.5539/ies.v7n7p77

Geary, D. C. (1996). *Children's Mathematical Development: Research and Practical Applications.* Washington, DC: American Psychological Association.

Goebel, S., Moeller, K., Pixner, S., Kaufmann, L., and Nuerk, H.-C. (2014). Language affects symbolic arithmetic in children: the case of number word inversion. *J. Exp. Child Psychol.* 119, 17–25. doi: 10.1016/j.jecp.2013.10.001

Goebel, S. M., Shaki, S., and Fischer, M. H. (2011). The cultural number line: a review of cultural and linguistic influences on the development of number processing. *J. Cross Cult. Psychol.* 42, 543–565. doi: 10.1177/0022022111406251

Grabner, R. H., Ansari, D., Koschutnig, K., Reishofer, G., Ebner, F., and Neuper, C. (2009). To retrieve or to calculate? Left angular gyrus mediates the retrieval of arithmetic facts during problem solving. *Neuropsychologia* 47, 604–608. doi: 10.1016/j.neuropsychologia.2008.10.013

Greer, B. (1992). "Multiplication and division as models of situations," in *HandBook of Research on Mathematics Teaching and Learning,* ed. D. A. Grouws (New York, NY: Macmillan Publishing Company), 276–295.

Haghverdi, M., Semnani, A. S., and Seifi, M. (2012). The relationship between different kinds of students' errors and the knowledge required to solve mathematics word problems. *Bolema Bol. Educ. Matem.* 26, 649–666. doi: 10.1590/S0103-636X2012000200012

Hegarty, M., Mayer, R. E., and Green, C. E. (1992). Comprehension of arithmetic word problems: evidence from students' eye fixations. *J. Educ. Psychol.* 84, 76–84. doi: 10.1037/0022-0663.84.1.76

Hegarty, M., Mayer, R. E., and Monk, C. A. (1995). Comprehension of arithmetic word problems: a comparison of successful and unsuccessful problem solvers. *J. Educ. Psychol.* 87, 18–32. doi: 10.1037/0022-0663.87.1.18

Hines, T. M. (2013). Parity influences the difficulty of simple addition and subtraction but not multiplication problems in children. *Psychol. Rep.* 113, 1048–1065. doi: 10.2466/10.11.PR0.113x16z4

Hinsley, D. A., Hayes, J. R., and Simon, H. A. (1977). "From words to equations: meaning and representation in algebra word problems," in *Cognition Processes in Comprehension,* ed. P. S. Carpenter (Hillsdale, NJ: Erlbaum), 89–100.

Imbo, I., Bulcke, C. V., De Brauwer, J., and Fias, W. (2014). Sixty-four or four-and-sixty? The influence of language and working memory on children's number transcoding. *Front. Psychol.* 5:313. doi: 10.3389/fpsyg.2014.00313

Jimenez, L., and Verschaffel, L. (2014). Development of children's solutions of non-standard arithmetic word problem solving non-standard arithmetic word problems. *Rev. Psicodidactica* 19, 93–123. doi: 10.1387/RevPsicodidact.7865

Josta, K., Khadera, P., Burkea, M., Bienb, S., and Röslera, F. (2009). Dissociating the solution processes of small, large, and zero multiplications by means of fMRI. *Neuroimage* 46, 308–318. doi: 10.1016/j.neuroimage.2009.01.044

Kaput, J. J. (1979). "Mathematics and learning: roots of epistemological status," in *Cognitive Process Instruction,* eds J. Clement and J. Lochhead (Philadelphia, PA: Franklin Institute Press), 289–303.

Kieran, C. (1990). "Cognitive processes involved in learning school algebra," in *Mathematics and Cognition,* eds P. Nesher and J. Kilpatrick (New York, NY: Cambridge University Press), 96–112.

Kingsdorf, S., and Krawec, J. (2014). Error analysis of mathematical word problem solving across students with and without learning disabilities. *Learn. Disabil. Res. Pract.* 29, 66–74. doi: 10.1111/ldrp.12029

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

Klein, E., Moeller, K., Kettl, D., Zauner, H., Wood, G., Willmes, K., et al. (2013). Object-based neglect in number processing. *Behav. Brain Funct.* 9, 5. doi: 10.1186/1744-9081-9-5

Koedinger, K. R., and Nathan, M. J. (2004). The real story behind story problems: effects of representations on quantitative reasoning. *J. Learn. Sci.* 13, 129–164. doi: 10.1207/s15327809jls1302_1

Lachmair, M., Dudschig, C., De La Vega, L., and Kaup, B. (2014). Relating numeric cognition and language processing: do numbers and words share a common representational platform? *Acta Psychol.* 148, 107–114. doi: 10.1016/j.actpsy.2013.12.004

Lean, G. A., Clemens, M. A., and Del Campo, G. (1990). Linguistic and pedagogical factors affecting children's understanding of arithmetic word problems: a comparative study. *Educ. Stud. Math.* 21, 165–191. doi: 10.1007/BF00304900

LeBlanc, M. D., and Weber-Russell, S. (1996). Text integration and mathematical connections: a computer model of arithmetic word problem solving. *Cogn. Sci.* 20, 357–407. doi: 10.1207/s15516709cog2003_2

Lee, K., Ng, S. F., Ng, E. L., and Lim, R. Y. (2004). Working memory and literacy as predictors of performance on algebraic word problems. *J. Exp. Child Psychol.* 89, 140–158. doi: 10.1016/j.jecp.2004.07.001

Lepik, M. (1990). Algebraic word problems: role of linguistic and structural variables. *Educ. Stud. Math.* 21, 83–90. doi: 10.1007/BF00311017

Lewis, A. B., and Mayer, R. E. (1987). Students' miscomprehension of relational statements in arithmetic word problems. *J. Educ. Psychol.* 79, 363–371. doi: 10.1037/0022-0663.79.4.363

MacGregor, M., and Price, E. (1999). An exploration of aspects of language proficiency and algebra learnin. *J. Res. Math. Educ.* 30, 449–467. doi: 10.2307/749709

Martin, S. A., and Bassok, M. (2005). Effects of semantic cues on mathematical modeling: evidence from word-problem solving and equation construction tasks. *Mem. Cognit.* 33, 471–478. doi: 10.3758/BF03195064

Martiniello, M. (2008). Language and the performance of english-language learners in math word problems. *Harv. Educ. Rev.* 78, 333–368.

Mayer, R. E. (1981). Frequency norms and structural analysis of algebra story problems into families, categories, and templates. *Instr. Sci.* 10, 135–175. doi: 10.1007/BF00132515

Mayer, R. E. (1982). Memory for algebra story problems. *J. Educ. Psychol.* 74, 199–216. doi: 10.1037/0022-0663.74.2.199

Miller, K., Smith, C. M., Zhu, J., and Zhang, H. (1995). Preschool origins of cross-national differences in mathematical competence: the role of number-naming systems. *Psychol. Sci.* 6, 6–60. doi: 10.1111/j.1467-9280.1995.tb00305.x

Moeller, K., Klein, E., Kucian, K., and Willmes, K. (2014). Numerical development - from cognitive functions to neural underpinnings. *Front. Psychol.* 5:1047. doi: 10.3389/fpsyg.2014.01047

Moeller, K., Klein, E., and Nuerk, H.-C. (2011). (No) small adults – children's processing of carry addition problems. *Dev. Neuropsychol.* 36, 702–720. doi: 10.1080/87565641.2010.549880

Muth, K. D. (1992). Extraneous information and extra steps in arithmetic word problems. *Contemp. Educ. Psychol.* 17, 278–285. doi: 10.1016/0361-476X(92)90066-90068

Nathan, M. J., Kintsch, W., and Young, E. (1992). A theory of algebra-word-problem comprehension and its implications for the design of learning environments. *Cogn. Instr.* 9, 329–389. doi: 10.1207/s1532690xci0904_2

Neef, N. A., Nelles, D. E., Iwata, B. A., and Page, T. J. (2003). Analysis of precurrent skills in solving mathematics story problems. *J. Appl. Behav. Anal.* 36, 21–33. doi: 10.1901/jaba.2003.36-21

Nesher, P. (1976). Three determinants of difficulty in verbal arithmetic problems. *Educ. Stud. Math.* 7, 369–388. doi: 10.1007/BF00452220

Nesher, P., and Teubal, E. (1975). Verbal cues as an interfering factor in verbal problem solving. *Educ. Stud. Math.* 6, 41–51. doi: 10.1007/BF00590023

Nichols, J. (1990). Linguistic diversity and the first settlement of the new world. *Language* 66, 475–521. doi: 10.2307/414609

Nuerk, H.-C., Geppert, B. E., Van Herten, M., and Willmes, K. (2002). On the impact of different number representations in the number bisection task. *Cortex* 38, 691–715. doi: 10.1016/S0010-9452(08)70038-8

Nuerk, H.-C., Moeller, K., Klein, E., Willmes, K., and Fischer, M. H. (2011). Extending the mental number line – a review of multi-digit number processing. *J. Psychol.* 219, 3–22.

Nuerk, H.-C., Moeller, K., and Willmes, K. (2015). "Multi-digit number processing: overview, conceptual clarifications, and language influences," in *The Oxford Handbook of Numerical Cognition*, eds R. C. Kadosh and A. Dowker (Oxford: Oxford University Press). doi: 10.1093/oxfordhb/9780199642342.013.021

Nuerk, H.-C., Weger, U., and Willmes, K. (2001). Decade breaks in the mental number line? Putting the tens and units back in different bins. *Cognition* 82, 25–33. doi: 10.1016/S0010-0277(01)00142-1

Nuerk, H.-C., Weger, U., and Willmes, K. (2005). Language effects in magnitude comparison: small, but not irrelevant. *Brain Lang.* 92, 262–277. doi: 10.1016/j.bandl.2004.06.107

Nuerk, H.-C., and Willmes, K. (2005). On the magnitude representation of two-digit numbers. *Psychol. Sci.* 42, 52–72.

OECD. (2010). *PISA 2009 Results. What Students Know and Can Do. Student Performance in Reading, Mathematics and Science,* Vol. 1. Paris: OECD Publishing.

O´Neill, D. K. (2013). Preschool children's narratives and performance on the peabody individualized achievement test – revised: evidence of a relation between early narrative and later mathematical ability. *First Lang.* 24, 149–183. doi: 10.1177/0142723704043529

Oostermeijer, M., Boonen, A. J. H., and Jolles, J. (2014). The relation between children's constructive play activities, spatial ability, and mathematical word problem-solving performance: a mediation analysis in sixth-grade students. *Front. Psychol.* 5:782. doi: 10.3389/fpsyg.2014.00782

Orrantia, J., Rodriquez, L., and Vicente, S. (2010). Automatic activation of addition facts in arithmetic word problems. *Q. J. Exp. Psychol.* 63, 310–319. doi: 10.1080/17470210902903020

Paige, J. M., and Simon, H. A. (eds). (1966). *Cognitive Processes in Solving Algebra Word Problem.* New York, NY: John Wiley & Sons.

Pape, S. J. (2003). Compare word problems: consistency hypothesis revisited. *Contemp. Educ. Psychol.* 28, 396–421. doi: 10.1016/S0361-476X(02)00046-2

Passolunghi, M. C., and Siegel, L. S. (2001). Short-term memory, working memory and inhibitory control in children with difficulties in arithmetic problem solving. *J. Exp. Child Psychol.* 80, 44–57. doi: 10.1006/jecp.2000.2626

Pixner, S., Kaufmann, L., Moeller, K., Hermanova, V., and Nuerk, H.-C. (2011a). Whorf reloaded: language effects on non-verbal number processing in 1st grade - a trilingual study. *J. Exp. Child Psychol.* 108, 371–382. doi: 10.1016/j.jecp.2010.09.002

Pixner, S., Zuber, J., Hermanova, V., Kaufmann, L., and Nuerk, H.-C. (2011b). One language, two number-word systems and many problems: numerical cognition in the Czech language. *Res. Dev. Disabil.* 32, 2683–2689. doi: 10.1016/j.ridd.2011.06.004

Powell, S. R., and Fuchs, L. S. (2014). Does early algebraic reasoning differ as a function of students' difficulty with calculations versus word problems? *Learn. Disabil. Res. Pract.* 29, 106–116. doi: 10.1111/ldrp.12037

Raduan, I. H. (2010). Error analysis and the corresponding cognitive activities committed by year five primary students in solving mathematical word problems. *Procedia Soc. Behav. Sci.* 2, 3836–3838. doi: 10.1016/j.sbspro.2010.03.600

Reusser, K. (1988). Problem solving beyond the logic of things: contextual effects on understanding and solving word problems. *Instr. Sci.* 17, 309–338. doi: 10.1007/BF00056219

Reusser, K. (1993). "Tutoring systems and pedagogical theory: representational tools for understanding, planning and refllection in problem solving," in *Computers as Cognitive Tools,* ed. S. L. S. Derry (Hillsdale, NJ: Lawrence Erlbaum).

Riley, M. S., Greeno, J. G., and Heller, J. I. (1983). "Development of children's problem-solving ability," in *The Development of Mathematical Thinking,* ed. H. P. Ginsberg (San Diego, CA: Academic Press), 153–196.

Roettger, T. B., and Domahs, F. (2015). Grammatical number elicits SNARC and MARC effects as a function of task demands. *Q. J. Exp. Psychol.* doi: 10.1080/17470218.2014.979843 [Epub ahead of print].

Sarnecka, B. W. (2013). On the relation between grammatical number and cardinal numbers in development. *Front. Psychol.* 5:1132. doi: 10.3389/fpsyg.2014.01132

Schley, D. R., and Fujita, K. (2014). Seeing the math in the story: on how abstraction promotes performance on mathematical word problems. *Soc. Psychol. Personal. Sci.* 5, 953–961. doi: 10.1177/1948550614539519

Schumacher, R. F., and Fuchs, L. S. (2012). Does understanding relational terminology mediate effects of intervention on compare word problems? *J. Exp. Child Psychol.* 111, 607–628. doi: 10.1016/j.jecp.2011.12.001

Searle, B. W., Lorton, P. Jr., and Suppes, P. (1974). Structural variables affecting CAI performance on arithmetic word problems of disadvantaged and deaf students. *Educ. Stud. Math.* 5, 371–384. doi: 10.1007/BF01424555

Shaftel, J., Belton-Kocher, E., Glasnapp, D., and Poggio, J. (2006). The impact of language characteristics in mathematics test items on the performance of english language learners and students with disabilities. *Educ. Assess.* 11, 105–126. doi: 10.1207/s15326977ea1102_2

Shaki, S., Fischer, M. H., and Petrusic, W. M. (2009). Reading habits for both words and numbers contribute to the SNARC effect. *Psychon. Bull. Rev.* 16, 328–331. doi: 10.3758/PBR.16.2.328

Spanos, G., Rhodes, N. C., and Dale, T. C. (1988). "Linguistic features of mathematical problem solving: insights and applications," in *Linguistic and Cultural Influences on Learning Mathematics,* eds R. R. Cocking and J. P. Mestre (Hillsdale, NJ: Lawrence Erlbaum), 221–240.

Stacey, K. (2012). "The international assessment of mathematical literacy: PISA 2012 framework and items," in *Proceedings of the 12th International Congress on Mathematical Education,* Seoul.

Sternberg, S. (1969). The discovery of processing stages: extensions of Donders' method. Attention and performance II. *Acta Psychol.* 30, 276–315. doi: 10.1016/0001-6918(69)90055-9

Swanson, H. L. (2004). Working memory and phonological processing as predictors of children's mathematical problem solving at different ages. *Mem. Cognit.* 32, 648–661. doi: 10.3758/BF03195856

Swanson, H. L., Cooney, J. B., and Brock, S. (1993). The influence of working memory and classification ability on childrens word problem solution. *J. Exp. Child Psychol.* 55, 374–395. doi: 10.1006/jecp.1993.1021

Swanson, H. L., Lussier, C., and Orosco, M. (2013). Effects of cognitive strategy interventions and cognitive moderators on word problem solving in children at risk for problem solving difficulties. *Learn. Disabil. Res. Pract.* 28, 170–183. doi: 10.1111/ldrp.12019

Terao, A., Koedinger, K. R., Sohn, M. H., Qin, Y., Anderson, J. R., and Carter, C. S. (2004). "An fMRI study of the interplay of symbolic and visuo-spatial systems in mathematical reasoning," in *Proceedings of the 26th Annual Conference of the Cognitive Science Society*, Chicago, 1327–1332.

Thevenot, C., Devidal, M., Barrouillet, P., and Fayol, M. (2007). Why does placing the question before an arithmetic word problem improve performance? A situation model account. *Q. J. Exp. Psychol.* 60, 43–56. doi: 10.1080/17470210600587927

Thevenot, C., and Oakhill, J. (2005). The strategic use of alternative representation in arithmetic word problem solving. *Q. J. Exp. Psychol.* 58, 1311–1323. doi: 10.1080/02724980443000593

van der Schoot, M., Bakker Arkema, A. H., Horsley, T. M., and Van Lieshout, E. C. D. M. (2009). The consistency effect depends on markedness in less successful but not successful problem solvers: an eye movement study in primary school children. *Contemp. Educ. Psychol.* 34, 58–66. doi: 10.1016/j.cedpsych.2008.07.002

Van Dijk, T., and Kintsch, W. (1983). *Strategies of Discourse Comprehension.* New York, NY: Academic Press.

Van Dooren, W., Verschaffel, L., and Onghena, P. (2002). The impact of pre-service teachers' content knowledge on their evaluations of students' strategies for solving arithmetic and algebra word problems. *J. Res. Math. Educ.* 33, 319–351. doi: 10.2307/4149957

Vergnaud, G. (2009). The theory of conceptual fields. *Hum. Dev.* 5, 83–94. doi: 10.1159/000202727

Verguts, T., and Fias, W. (2005). Interacting neighbors: a connectionist model of retrieval in single-digit multiplication. *Mem. Cogn.* 33, 1–16. doi: 10.3758/BF03195293

Verschaffel, L. (1994). Using retelling data to study elementary school children's representations and solutions of compare problems. *J. Res. Math. Educ.* 25, 141–165. doi: 10.2307/749506

Verschaffel, L., and De Corte, E. (1990). "Do non-semantic factors also influence the solution process of addition and subtraction word problems?," in *Learning and Instruction. European Research in an International Context,* eds H. Mandl, E. De Corte, N. Bennett, and N. H. H. F. Friedrich (Oxford: Pergamon), 415–429.

Verschaffel, L., De Corte, E., and Borghart, I. (1997). Pre-Service teachers conceptions and beliefs about the role of real-world knowledge in mathematical modelling of school word problems. *Learn. Instr.* 7, 339–359. doi: 10.1016/S0959-4752(97)00008-X

Verschaffel, L., De Corte, E., and Pauwels, A. (1992). Solving compare problems: an eye movement test of Lewis and Mayer's consistency hypothesis. *J. Educ. Psychol.* 84, 85–94. doi: 10.1037/0022-0663.84.1.85

Verschaffel, L., Greer, B., and De Corte, E. (2000). Making sense of word problems. *Educ. Stud. Math.* 42, 211–213.

Vicente, S., Orrantia, J., and Verschaffel, L. (2007). Influence of situational and conceptual rewording on word problem solving. *Br. J. Educ. Psychol.* 77, 829–848. doi: 10.1348/000709907X178200

Wilkins, J. L., Baroody, A. J., and Tiilikainen, S. (2001). Kindergartners' understanding of additive commutativity within the context of word problems. *J. Exp. Child Psychol.* 79, 23–36. doi: 10.1006/jecp.2000.2580

Wood, G., Nuerk, H.-C., Moeller, K., Geppert, B., Schnitker, R., Weber, J., et al. (2008). All for one, but not one for all. How multiple number representations are recruited in one numerical task. *Brain Res.* 1187, 154–166. doi: 10.1016/j.brainres.2007.09.094

Yeap, B. H., Ho, S. Y., Kaur, B., and Lee, N. H. (2005). Children making sense during word problem solving. *Papers from the International Conference "Redesigning Pedagogy: Research, policy and practice,"* Singapore. Available at: http://math.nie.edu.sg/T3/downloads/rprpp.pdf (accessed March 2015).

Yeap, B. H., and Kaur, B. (2001). "Semantic characteristics that make arithmetic word problems difficult," in *Proceedings of the 24th Annual Conference of the Mathematics Education Research Group of Australasia Incorporated (MERGA) on "Numeracy and Beyond,"* Australia.

Yee, F. P., and Lee, K. P. (1997). School word problems and stereotyped thinking. *Teach. Learn.* 18, 73–82.

Zarnhofer, S., Braunstein, V., Ebner, F., Koschutnig, K., Neuper, C., Ninaus, M., et al. (2013). Individual differences in solving arithmetic word problems. *Behav. Brain Funct.* 9, 28. doi: 10.1186/1744-9081-9-28

Zuber, J., Pixner, S., Moeller, K., and Nuerk, H.-C. (2009). On the language-specificity of basic number processing: transcoding in a language with inversion and its relation to working memory capacity. *J. Exp. Child Psychol.* 102, 60–77. doi: 10.1016/j.jecp.2008.04.003