Irrelevant Information and Person Effects on Word-Problem Performance

By

Amber Yu Wang

Dissertation
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY
in

Special Education
August, 2016
Nashville, Tennessee

Approved:

Professor Lynn S. Fuchs, Ph.D.

Professor Doug Fuchs, Ph.D.

Professor Kimberly Paulsen, Ed.D.

Professor Bethany Rittle-Johnson, Ph.D.

ACKNOWLEDGEMENTS

I would like to thank Dr. Lynn S. Fuchs, Dr. Doug Fuchs, Dr. Kimberly Paulsen, and Dr. Bethany Rittle-Johnson. I am immensely grateful for your guidance and feedback throughout this project. I would like to express my gratitude to Rebecca Abramson, Ashlyn King, Grace Fletcher, Sarah Krowka, Will French, Sarah Imran, Asia Amos, Christina Bowden, Carissa Sacchetti, and Alex Damiano for their assistance in data collection and commitment to the success of this project. In addition, I am incredibly thankful for Jennifer K. Gilbert for her assistance and feedback in data analysis.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Irrelevant numeric information has been shown to diminish the word-problem performance of students with math learning disabilities and their typically-developing peers (e.g., Leon, 1992; Kinne, 2002; Muth, 1991; Parmar, Cawley, & Frazita, 1996). When a word problem (WP) includes irrelevant information, students often have greater difficulty understanding the problem and reaching a correct solution. This is because the inclusion of irrelevant information in WP narratives makes it more difficult for students to recognize a novel problem as belonging to a familiar problem type for which they know a solution strategy.

In many classrooms, students are taught a solution strategy for a WP type and then practice that solution strategy on WPs that vary only by cover story. Consider the following *combine* WP: Rebecca baked 5 cakes. Ashlyn baked 4 cakes. How many cakes did Rebecca and Ashlyn bake? When a student recognizes this problem type as a *combine* problem, she applies the corresponding solution strategy. Now consider this WP: Claudia bought 4 pairs of earrings. Joan bought 7 pairs of earrings. How many pairs of earrings did Claudia and Joan buy? While this problem is novel, it only varies from the first by its cover story and numerical values. The similarity between the two WPs makes it relatively easy for students to recognize the latter problem as a *combine* problem and apply the taught solution procedure.

If, however, irrelevant information was included in the WP narrative, students may have more difficulty recognizing a previously taught problem type and apply an appropriate solution strategy. This *combine* WP includes irrelevant information: During the summer, Eduardo read 2 books in English. He also read 8 books in Portuguese and 2 magazines. How many books did Eduardo read? To solve this problem correctly, students must discriminate the relevant from the irrelevant numbers. To do so, a student has to identify non-overlapping sets and match the values of the semantic features requested in the question (e.g., Eduardo, books) with those presented in the problem narrative. For example, in this WP, a student must recognize "2 magazines" is not a subset of the total set (e.g., books) requested by the problem question.

If a student cannot discriminate relevant from irrelevant information, the student may fail to recognize a problem type and thus fail to apply the correct solution strategy. Alternatively, when the student identifies the correct problem type but fails to identify the irrelevant information, the student may incorrectly compute a solution using *all* values presented in the WP. In still another scenario, the student does not even attempt to decipher the WP narrative and applies a rote computational strategy, such as adding all numbers presented (Goodstein, Cawley, Gordon, & Helfgott, 1971). In light of these different scenarios, one might suppose that different students experience difficulty with WPs containing irrelevant information for different reasons, leading to the use of different strategies or processes. While some students may not comprehend a WP and apply a rote strategy, others may comprehend the WP narrative but have difficulty because added steps are needed to identify relevant from irrelevant information (Littlefield & Rieser, 1993).

While research has shown that students at all grade levels have difficulty discriminating relevant from irrelevant information in mathematical WPs (e.g., Carpenter, Corbitt, Kepner, Lindquist, & Reyes, 1980; Kouba et al., 1988), it remains unclear under what conditions discrimination is easier or harder. Furthermore, little is known about what person characteristics account for variations in performance on WPs with irrelevant information. Few studies, however, have demonstrated that performance varies across WPs with varying subtypes of irrelevant information. This suggests that person characteristics affect information discrimination. In the following sections, we discuss irrelevant information and person effects.

Irrelevant Information Effects

The literature on mathematical problem-solving addresses the effect of irrelevant information on WP performance, but few studies have explicitly examined which *features* of irrelevant information make WPs more difficult. Yet, such studies are needed to understand how students selectively attend to information. A different pattern of effects across varying subtypes of irrelevant information may indicate that students attend to specific features of problem information, which leads them to apply particular strategies to represent problem information. Understanding which features are most salient to students can help practitioners improve instruction aimed at advancing students' word problem-solving skills. In the following studies, the position or placement of irrelevant information were systematically varied to examine whether these features produced differential effects on WP performance.

Position of irrelevant information. Englert et al. (1987) investigated the position of irrelevant information on the problem-solving performance of 48 second and fourth-grade students with learning disabilities (LD) and their typically-developing (TD) peers. Of the 24 students with LD, 12 were in the second grade; 12 were in the fourth grade. Similarly, TD students included 12 second graders and 12 fourth graders. Englert et al. operationally defined *position* to refer to the subject or object position of information; thus, *position* of irrelevant information referred to a quantity linked to the subject or object position of a sentence. For example, if a WP contained only an irrelevant subject, then the problem has a numerical quantity linked to the irrelevant subject, thereby making the numerical quantity irrelevant (e.g., Tim has 4 cats. Tim has 2 dogs. *Mary has 3 dogs*. How many animals does Tim have?). In contrast, a WP with an irrelevant object would contain a relevant subject with an irrelevant object quantity linked to it (e.g., Tim has 4 cats. Tim has 2 dogs. *Tim has 3 balls*. How many animals does Tim have?).

All students were prescreened for basic computational skills in addition; those that met 90% accuracy on the screening measure were included in the study. A set of addition WPs was administered to all students; all WPs were composed of four-sentence statements and required basic addition facts with sums less than 10. Four levels of irrelevant information were examined: (a) none, (b) irrelevant subject, (c) irrelevant object, and (d) linguistic information irrelevant to the problem text. During testing, students were cued that they might not need to use all of the information presented in a WP.

For both populations, WPs with an irrelevant object were more difficult than WPs with an irrelevant subject; both problem types were more difficult than WPs without irrelevant information. These results suggest that distractors in the subject position are easier to detect than distractors in the object position and that WP information associated with a subject serves as an important basis for identifying separate sets of information. Englert et al. (1987) stated, “the findings […] indicated that problem solvers relied most heavily on subject information as a basis for initially deciding object ownership and identifying non-overlapping subsets” (p. 34). Because of greater attention to subject information, students more readily recognized and excluded irrelevant information linked to an irrelevant subject. In contrast, students had greater difficulty recognizing separate objects linked to the same subject, and thereby excluding the irrelevant object associated with the subject. In general, when reading a problem and scanning the text for relevant information, students may cue in to a particular semantic category more strongly than other semantic categories.

This possibility is supported by Littlefield and Rieser's (1993) model of determining relevance, which posits that students notice major semantic categories. When solving WPs, students review semantic categories and their values and attempt to match the values of the semantic features requested in the problem question with those presented in the problem narrative. In the case of Englert et al. (1987), information linked to a subject semantic category may be more salient to students than information linked to an object semantic category. When reading a WP with irrelevant information linked to an irrelevant object (but relevant subject), students may not be able to switch their attention to the object category because they are focused on the subject category. Students may focus more on *who* the WP is about and how much that individual has, rather than on what the object is. By focusing on only one semantic feature to discriminate between irrelevant and relevant information, a student will less likely correctly identify all the information as relevant and reach a correct solution.

Placement of irrelevant information. Differences in the effect of placement of irrelevant information within WPs may indicate students' tendencies to focus on surface characteristics of WPs rather than engage in deeper analysis (examination of semantic categories and their values). Differential performance as a function of placement would suggest that some students use information placement as their strategy for discriminating relevant information, identifying the first two numbers in a WP as relevant for solution. For example, consider this WP: Mary has 5 apples. She buys 3 oranges. Then, Tom gives her 2 more apples. How many apples does Mary have now? In this WP, students who rely on information placement may identify *5* and *3* as relevant because they are the first two numbers presented in the WP.

In Blankenship and Lovitt (1976), *placement* referred to the sentence in which irrelevant information appeared within the WP. The researchers varied the placement of irrelevant information across the second, third, and fourth sentences of WPs. Using a single-subject design, a series of WP measures were administered to seven students with LD aged 9 to 12. The authors developed 12 classes of addition and subtraction WPs by varying number and type of nouns, type or tense of verb, form of verbal cue to operation, presence of an introductory sentence, placement of irrelevant information, form of question, and mode of numeral. Prior to the study, students were required to perform at 100% accuracy in three consecutive sessions on recall of the basic addition and subtraction facts and vocabulary to be presented within Class 1 WPs. Once a student met criterion, Class 1 problems were assigned. Students received a daily assignment of five problems each of addition, subtraction, and addition/subtraction problems. Criterion was operationally defined as 100% accuracy on each daily assignment across three consecutive sessions. Each class of problems was introduced with a baseline phase in which neither instruction nor feedback were provided to students.

Two classes of WPs focused on placement: Class 4, with irrelevant information in only the fourth sentence (excluding the question) and Class 5, irrelevant information in either the second, third, or fourth sentence (again excluding the question). The comparison of total mean baseline performance showed that when irrelevant information was first introduced in a constant position (i.e., Class 4), WP performance dropped (M = 79.6%). Once students mastered irrelevant information in a constant position, Class 5 problems were presented. With the placement of irrelevant information varying, WP performance was further diminished (M = 66.4%).

The authors concluded that student difficulty with irrelevant numerical information may be explained by the use of a rote strategy (Goodstein et al., 1971). Students do not read an entire WP, but instead read only the segments that include numbers and then use those numbers to reach a solution (Earp, 1970). Despite that the findings of this study suggest that placement of irrelevant information affects WP performance, it remains inconclusive whether the placement of irrelevant information has a significant effect on WP performance. Though the study provided examples of WPs containing irrelevant information in the object position, it is unclear if the *position* of irrelevant information was controlled for in Class 4 and Class 5 problems. If irrelevant information varied between the subject and object positions, then the effects of placement are confounded. Despite these limitations, findings are supported by Littlefield and Rieser (1993).

With 30 fifth-grade students, Littlefield and Rieser (1993) also examined the effect of placement of irrelevant information on problem solving. Performance, however, was operationalized as the percentage of numbers in each WP correctly identified as relevant or irrelevant. WPs varied in whether relevant or irrelevant information appeared first or last in the problem statement. Results indicated a statistically significant effect for the placement of irrelevant information, with greater accuracy on WPs containing relevant information first. While all WPs in Blankenship and Lovitt (1976) contained relevant information first and also varied irrelevant information across the second and third statements, findings across these two studies suggest that the relative placement of information within WPs affects a student's ability to discriminate relevant information.

In summary, this group of studies suggest irrelevant information subtypes exert a different pattern of effects on WP performance. Specifically, these studies demonstrate that varying positions or placements of irrelevant information differentially impact WP performance. Conclusions, however, are limited by the few available studies.

Person Effects

The process of solving a novel WP—interpreting text, identifying a schema, and applying a solution strategy—requires high-road transfer. During high-road transfer, individuals abstract connections (i.e., schemas) between familiar and novel tasks (Fuchs et al., 2003; Salomon & Perkins, 1989). In contrast, low-road transfer involves the automatic triggering of well-learned, stimulus-controlled behavior in a new context (Fuchs et al., 2003; Salomon & Perkins, 1989). Thus, when presented a novel WP, such as one that contains irrelevant information, a student must withhold an initial response and intentionally examine how the novel problem at hand connects with familiar problems. This process of high-road transfer is also known as "mindful abstraction," (Salomon & Perkins, 1989) or metacognition (Fuchs et al., 2003).

The process of solving a WP not only requires metacognition, but also makes strong demands on reasoning and working memory. Here, we highlight a model of mathematical WP solving based on the seminal work of Kintsch and colleagues (e.g., Cummins et al., 1988; Kintsch & Greeno, 1985) and theories of text comprehension and discourse processing (e.g., Van Dijk & Kintsch, 1983; Graesser, Millis, & Zwaan, 1997). This model (Fuchs et al., 2015) posits that WP solving involves problem-solving strategies that rely on working memory, language comprehension, and nonverbal reasoning.

Working memory. In Fuchs et al.'s (2015) model of WP solving, memory representations have three components. First, when reading a WP narrative, a student must construct a coherent structure to capture the text's essential ideas (i.e., the propositional text structure). Next, the student develops a situation model by supplementing the text with inferences based on his background knowledge. Then, the student coordinates this information with the third component, the problem schema. At this stage, the student formalizes conceptual relations among quantities and uses the schema to guide the application of solution strategies. Overall, the process of constructing a propositional text structure, inferencing, developing a situation model, identifying a schema, and finally, applying a solution strategy, places strong demands on working memory. This process involves not only the coordination of multiple pieces of information, but also the storing and manipulation of multiple steps in memory.

Language comprehension. Given the need to process linguistic information when solving a WP, the role of language comprehension is important to consider. Studies have demonstrated the role of language in WP solving (e.g., Fuchs et al., 2006; Fuchs et al., 2008). Furthermore, research has demonstrated that WPs differ from other forms of text comprehension. Fuchs et al. (2015) assessed second-grade students on general and WP-specific language comprehension. Path analytic mediation results showed that the effects of general language comprehension on informational text comprehension were entirely direct, whereas the effects of general language comprehension on WPs were partially mediated by WP-specific language.

Nonverbal reasoning. Previous research has identified nonverbal reasoning as a significant predictor of WP performance (e.g., Fuchs et al, 2006; Fuchs et al., 2015). Extending this line of research, Wang, Fuchs, and Fuchs (submitted) examined the performance of 701 2nd-grade students on WPs with and without irrelevant information. Nonverbal reasoning was a significant predictor of performance on WPs *with* irrelevant information, but not WPs without irrelevant information. This suggests that WPs with irrelevant information place stronger demands on an individual's reasoning ability than WPs without irrelevant information.

When solving a WP with irrelevant information, a student must (a) master problem-solution strategies, (b) categorize problems into problem-types or schemas, and (c) recognize how novel problems relate to taught problems (Cooper & Sweller, 1987). While the inclusion of irrelevant information makes a problem novel, it does not alter the underlying mathematical structure or the problem-solution method. The inclusion of irrelevant information, however, increases a WP's complexity. This places a strong demand on reasoning ability because to solve a WP, a student must reason analytically and make categorical judgments in distinguishing relevant from irrelevant features of classes of objects.

Rationale for Study

The purpose of this study is to extend the literature on WP performance by bringing together two lines of research, namely irrelevant information and person factors that affect WP performance, using a crossed random-effects item response model, also referred to in the literature as a cross-classification model (Goldstein, 1987; Raudenbush, 1993; Van den Noortgate, De Boeck, & Meulders, 2003). Two points of rationale drive this study. First, although prior research (e.g., Blankenship & Lovitt, 1976; Englert et al., 1987; Littlefield & Rieser, 2003) has documented the effects of irrelevant information subtypes on WP performance, the limited number of studies warrants replication studies. As documented by Englert et al. (1987), the position of irrelevant information creates a different pattern of effects on WP performance. This study, however, is the only one of its kind to examine directly the effect of information *position* on WP performance amongst students at risk for mathematics disabilities, while also accounting for person-level factors.

Furthermore, earlier studies did not examine what person-level factors influence performance on WPs *with* irrelevant information. Thus, the present study combined factors found to be significant in prior research into one model so that effects could be established in the presence of other potential influences. Specifically, the present study examined two levels of irrelevant information position (i.e., subject position vs. object position) and three person-level factors (i.e., working memory, language comprehension, and nonverbal reasoning). It is important to note here that WPs were constructed by varying the two levels of irrelevant information position with two levels of irrelevant information placement (i.e., relevant information first vs. irrelevant information first) in order to reflect real-world situations. Irrelevant information is not always presented or encountered after relevant information.

A crossed random-effects model allows for the simultaneous investigation of item covariates, person covariates, and item-by-person interactions while accounting for variance in responses that is due to clustering of items and persons. In a crossed random effects model, both person and item are treated as random effects. Modeling person effects as random is reasonable as it is unlikely any combination of person covariates will account for all variance in the outcome between people. Although treating item as a random effect is uncommon in item-response theory, it is a more realistic way to model the variation in item difficulty (Van den Noortgate et al., 2003) and is most useful when examining difficulties of item types rather than individual items. The analysis for the present study allows the effects of person covariates and item covariates (e.g., irrelevant information subtypes) to be assessed concurrently while accounting for the variance/residual variance due to items and persons, thereby providing unbiased estimates of covariates.

Research Questions and Hypotheses

In line with the purpose of the study, we had three research questions. The first question was, What is the effect of irrelevant information position on WP performance, after controlling for incoming basic arithmetic skill? Second, What are the effects of working memory, language comprehension, nonverbal reasoning, and grade in accounting for person variance in WP performance after controlling for irrelevant information position and basic arithmetic skill? Third, Are there significant person-by-item interactions? For this third question, we considered an interaction term between nonverbal reasoning and irrelevant information position as well as an interaction term between working memory and position.

Based on Englert et al.'s (1987) findings, our hypothesis regarding the first research question was that the predicted probability of WP performance (i.e., accuracy) is higher when irrelevant information is placed in the subject position than in the object position. For the second question, in accord with the prior research just summarized, we hypothesized unique predictive power on WP performance for each of the student covariates entered in the model. With respect to the third question, interaction terms were included for exploratory purposes. Based on prior research demonstrating the significant contribution of reasoning and the effect of information position in WP performance, we expected a significant interaction between nonverbal reasoning and irrelevant information position. Specifically, we expected students with low nonverbal reasoning to experience greater difficulty with WPs containing irrelevant information in the object position than WPs with irrelevant information in the subject position, whereas students with better reasoning to perform comparably in both conditions.

CHAPTER II

METHOD

Participants

Participants were from a larger study investigating the efficacy of fractions tutoring for English-speaking third- and fifth-grade students who were selected for a history of mathematics difficulties. Participants were identified in a large, metropolitan school district using a two-stage screening process. We recruited 26 teachers from 11 schools (6 elementary; 5 middle) who taught a total of 34 math sections to participate in the study. We conducted whole-class testing to identify at-risk students. For fifth graders, we defined risk as performance below the 27^{th} percentile at the start of the school year on a measure of broad-based calculations (Wide Range Achievement Test–4 [WRAT-4]; Wilkinson & Robertson, 2006). For third graders, students qualified as at-risk in one of two ways: (a) performance below the 30th percentile on the WRAT-4 or (b) performance at the 30th percentile on the WRAT-4 *and* a raw score of two or less on a measure of basic subtraction facts.

Because the present study (and larger study) was not about intellectual disability, we administered the *Vocabulary* and *Matrix Reasoning* subtests of the Wechsler Abbreviated Scales of Intelligence (WASI; Wechsler, 2011) to these students in individual testing sessions. We excluded students scoring below the 9^{th} percentile on *Vocabulary* and *Matrix Reasoning.* The remaining students (N = 140) were used for the present study and also participated in the larger study. Note that although students were randomly assigned at the individual level to tutoring or business-as-usual control conditions, the present study's data were collected before intervention began. So we included all 140 children, 90 third graders and 50 fifth graders.

Student demographics were as follows: 57% female; 60% African American, 16% Caucasian, 21% Hispanic, <2% Asian, 1 % other; 10% identified as receiving special education services; 74% free/reduced lunch; and 19% identified as ESL (English as a Second Language). On the WRAT-4 (Wilkinson, 2006), mean raw score performance was 22.77 (*SD* = 3.02); on WASI *Vocabulary* (Wechsler, 1999), 20.21 (*SD* = 5.68); and on WASI *Matrix Reasoning*, 11.30 (*SD* = 4.45). For third graders, mean raw score performance was 21.13 (*SD* = 2.01) on WRAT-4; on *Vocabulary*, 18.17 (*SD* = 4.35); and on *Matrix Reasoning*, 9.8 (*SD* = 3.79). For fifth graders, mean raw score performance was 25.72 (*SD* = 2.19) on WRAT-4; on *Vocabulary*, 23.88 (*SD* = 5.97); and on *Matrix Reasoning*, 14 (*SD* = 4.30).

Measures

Screening. The mathematics screening measure was WRAT-4 (Wilkinson & Robertson, 2006), for which students complete arithmetic problems of increasing difficulty. The subtraction measure administered to third graders consists of 25 single-digit problems with minuends up to 18. Students have 1 min to complete the measure. The general intelligence screening measure was the WASI (Weschler, 2011). WASI *Vocabulary* assesses expressive vocabulary, verbal knowledge, memory, learning ability, and crystallized and general intelligence. Students identify pictures and define words. Responses are awarded scores of 0, 1, or 2 based on quality. Testing discontinues after three consecutive scores of 0. WASI *Matrix Reasoning* measures nonverbal fluid reasoning and general intelligence. Students select 1 of 5 options that best completes a visual pattern. Reliability exceeds .92.

Working memory. We used two subtests of central executive working memory from the *Working Memory Test Battery for Children* (WMTB-C; Pickering & Gathercole, 2001)—*Listening Recall* and *Counting Recall*. Each subtest includes six dual-task items at span levels from 1-6 to 1-9. At each span level, the number of items to be recalled increases by one. Passing any four items within a span moves the student to the next level. Failing any three items within a span terminates the subtest. We used the trials correct score. For *Listening Recall*, the student determines if each sentence in a series is true and then recalls the last word in each sentence in the correct order. For *Counting Recall*, the student counts a set of 4, 5, 6, or 7 dots, each on a separate card. After the last card within an item, the student recalls the number of dots counted on each card in the correct order. Test-retest reliability ranges from .84-.93.

Language comprehension. WASI *Vocabulary* is a measure of expressive language ability and was entered as a language covariate in the data analysis models (see Screening section for description). In the larger study, neither receptive language ability nor reading ability were measured. During small-group testing, however, each WP was read up to two times for students.

Nonverbal reasoning. WASI *Matrix Reasoning* is a measure of nonverbal reasoning and was entered as a covariate in the data analysis models (see Screening section for description). *Matrix Reasoning* assesses nonverbal reasoning because a student must first identify the rules that govern the pattern. Then, the student must compare the rule of the stimulus pattern with the characteristics of the possible choices. Once a match in rules is made, the student applies the rule by choosing the correct answer.

Arithmetic. To control for initial arithmetic skill, we used *Math Fact Fluency* (Fuchs, Hamlett, & Powell, 2003) which has four subtests: 25 single-digit addition problems with sums from 6 to 12; 25 single-digit subtraction problems with minuends from 6 to 12; 25 single-digit addition problems with sums from 5 to 18; 25 single-digit subtraction problems with minuends from 5 to 18. Students have 1 min to write answers for each subtest. The score is the number of correct answers across subtests.

Word problems. To assess WPs with irrelevant information, we developed a measure of 20 problems. All WPs are combine story problems with one irrelevant numeric value. To assess the effect of position of irrelevant information, we created two versions of each WP (refer to Table 1). For one version, irrelevant information is in the subject position. In the other version,

Table 1

Examples of Word Problem Test Items

Item	Version	Position	Problem
1	A	Subject	A group of friends went to the zoo. Emma saw 3 brown monkeys. Liam saw 2 brown monkeys. Jenny saw 4 brown monkeys. How many brown monkeys did Emma and Liam see?
1	B	Object	A group of friends went to the zoo. Emma saw 3 brown monkeys. Liam saw 2 brown monkeys. Emma saw 4 gorillas. How many brown monkeys did Emma and Liam see?
2	A	Object	Ashlyn and Rebecca are making fruit salad. Ashlyn chopped 5 apples. Ashlyn chopped 8 pineapples and 3 melons. How many pineapples and melons did Ashlyn chop?
2	B	Subject	Ashlyn and Rebecca are making fruit salad. Rebecca chopped 5 melons. Ashlyn chopped 8 pineapples and 3 melons. How many pineapples and melons did Ashlyn chop?

irrelevant information is in the object position. Furthermore, irrelevant information is varied between two placements: (a) irrelevant value *before* all relevant values or (b) irrelevant value after *all* relevant values. Versions of each item are randomly distributed to Test A or Test B. Each test version includes 10 items with irrelevant information in the subject position and 10 items with irrelevant information in the object position; within each subtype, five items have irrelevant values placed before relevant information and five items with irrelevant values placed after relevant information. Cronbach's alphas for Test A and Test B were .92 and .90, respectively.

Procedure

As part of the larger study, students were administered measures of working memory during the individual screening for study inclusion. Prior to intervention, students were administered additional measures during two sessions of small-group testing. Small-group testing included measures of basic addition and subtraction skills and word problems. During administration of the WP measure, a test administrator read aloud each problem to students. Upon student request, a problem was re-read one time.

Testers were graduate research assistants at a local university. All research assistants received training on whole class and individual testing protocols during two four-hour sessions. Research assistants practiced administering the tests and passed a fidelity check before administering the tests in schools. Research assistants were also trained on scoring. They received a score sheet outlining scoring protocol and the project coordinator was available to answer any questions. Two independent research assistants scored and entered the data for each test. All scoring discrepancies were discussed and resolved. Research assistants received training on small-group testing protocols during one two-hour session. The above procedures were followed for small-group testing measures.

Data Analysis

In this study, the outcome was performance on WPs containing irrelevant information, with grade, language comprehension, working memory, and reasoning as person-level covariates and irrelevant information position as an item-level covariate. Data analysis consisted of cross-classified random-effects item response models. In a crossed random-effects item response model, both items and persons are treated as random parameters (Van den Noorgate et al., 2003); this assumes that the probability of a student answering a WP correctly is dependent on both person and item characteristics. In the present study, we test this assumption by calculating the proportion of variance in WP performance that is due to persons and items in an unconditional means model (i.e., Base model). Prior research (as cited in the introduction) suggests that this is a reasonable assumption for WP performance. All models were run using the *lmer* function from the *lme4* package in R (Bates et al., 2013).

In a crossed random-effects model, responses (i.e., units of analysis) are considered cross-classified in nature because they originate from the intersection of each student responding to each item in the same set of WPs. The lower-level units (i.e., responses) belong to upper-level units (i.e., persons and items) that are not nested within each other. Therefore, the data structure includes two levels. Problem-by-problem student responses are at Level 1. These responses are cross-classified by student and problem at Level 2. Furthermore, students were also nested in classrooms that are nested in schools. Each student was a member of only one classroom, and that classroom was a member of only one school. Clustering at those levels were taken into account by adding random effects for each. Figure 1 illustrates the 4-level data.

The first step in analyzing the data was to estimate a Base model without covariates. The purpose of the Base model was to determine the intraclass correlation (ICC) at each level. The first Base model (Full Base model) included all potential sources of variance (i.e., item, person, classroom, and school). For the sake of parsimony, we estimated a reduced model (Base without school). To determine whether each level of variance was necessary for accurately describing the

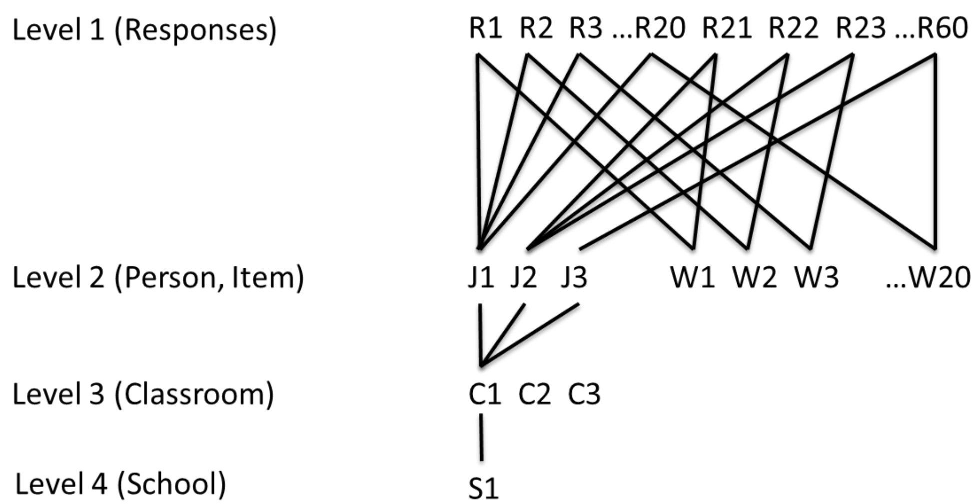

Figure 1. Cross-classified hierarchical data structure for school 1.

data, we compared the Full Base model to the reduced model. In addition to the Base model, we proposed three models to address the research questions. For analyses, all item- and person-level continuous predictors were centered at their respective means.

To address the first research question, we estimated a model including item subtype, while controlling for arithmetic skill, to provide information about the effect of irrelevant information position on WP performance. The effect of position was assessed with a dummy code at Level 2. Irrelevant information position was assessed with a dummy variable that equaled 0 if the WP contained an irrelevant number in the subject position and 1 if it contained an irrelevant number in the object position. To contextualize results, we converted the logit coefficient to a probability using the following formulas: $P = 1/(1 + \exp[-\gamma 0])$ and $P = 1/(1 + \exp[-\gamma 0 - \gamma 1])$. The calculated probabilities reflect how likely students were to correctly answer a problem containing irrelevant information in the subject position versus irrelevant information in the object position, respectively.

Next, we added predictors to Model 1 to determine which person factors affect WP performance above and beyond item subtype (Model 2). The aim of Model 2 was to estimate person effects on a student's probability of correctly solving a WP, controlling for arithmetic skill and irrelevant information position. As a final step, we estimated a model (Model 3) to test two person-by-item interactions. The first interaction was between reasoning and irrelevant information position and the second was between working memory and position.

In addition to the statistical analyses described above, we conducted an error analysis using 25% of the students' WP measures to examine whether students (a) were able to identify the two relevant numbers in a WP, but had performed an incorrect calculation; (b) relied on rote computation (i.e., addition of all three values); or (c) used a different or undetermined procedure.

CHAPTER III

RESULTS

Descriptive statistics on performance variables are provided in Table 2. Correlations among these scores are shown in Table 3.

Table 2

Description of Student Sample (N = 140)

	Mean	**SD**	**Minimum**	**Maximum**
WP Total	9.76	6.17	0	20
Listening Recall	8.94	3.89	0	19
Counting Recall	16.05	4.53	5	29
Language	20.21	5.68	6	40
Reasoning	11.30	4.45	2	22
Arithmetic	31.38	15.24	0	73

Table 3

Correlations for Student Assessment Scores (N = 140)

	1	**2**	**3**	**4**	**5**	**6**
1. WP Total	--					
2. Listening Recall	.39	--				
3. Counting Recall	.28	.52	--			
4. Language	.51	.40	.42	--		
5. Reasoning	.40	.50	.47	.44	--	
6. Arithmetic	.35	.29	.37	.36	.35	--

Note. Correlations are significant at $\alpha = .05$.

Base Model

Prior to estimating models for each research question, we estimated a potential Base model that included no covariates. For the sake of parsimony, we estimated a Reduced Base model by removing one random parameter (i.e., school). A likelihood ratio (LR) test indicated that the reduced model did not fit significantly worse when halving the *p*-value produced by the LR test; the χ^2 Test of Equivalence between the Full Base and Reduced Base models equaled 0. It has been shown that when comparing models that differ in the number of random parameters, the distribution of the LR test is a mixture of χ^2 distributions with different degrees of freedom (Stram & Lee, 1994). In contrast, the standard LR test assumes an asymptotic χ^2 distribution with degrees of freedom equal to the number of parameters being tested, which can yield conservative *p*-values. As such, accurate results for the present study can be obtained by halving the *p*-value of the LR test statistic.

Because the Reduced Base model fit just as well as the Full Base model, the random effect of School was removed in all conditional models. Estimates of the fixed and random effects of the Reduced Base model are shown in Table 4. For ease of interpretation, the logit coefficient from the results of the Reduced Base model was converted to a probability. The average predicted probability for an average person in an average classroom decoding a WP item of average difficulty was .45, 1/(1 + exp[-{-.19}]). The 95% plausible value range (Raudenbush & Bryk, 2002) for persons was .03 to .95 (1/1 + exp[-{-.19 ± 1.96*√2.70}]) for the predicted probability of getting an average WP item correct. For item, the range was .30 to .49 (1/1 + exp[-{-.19 ± 1.96*√.11}]) for the predicted probability of an average person. Though more variability in responses was associated with persons, the Reduced Base model still showed variability in responses across items, thereby warranting a person-and-item analysis.

Table 4

Results of Reduced Base Model (Unconditional Means Model) (N = 140)

Fixed Effects					
	Variable	Coefficient	*SE*	*z*	*p*
	Intercept	-0.19	0.22	-0.87	.39
Random Effects					
	Source	Variance			
	Person (intercept)	2.70			
	Classroom (intercept)	0.68			
	Item (Intercept)	0.11			

Note. Coefficients are in log-odds units.

Research Question 1

With research question 1, we assessed if the probability of correctly solving a WP differs depending on whether irrelevant information was in the subject position or in the object position. To address research question 1, we estimated a model (Model 1) including irrelevant information position while controlling for arithmetic skill. Both the intercept and the slope were allowed to vary randomly over persons and items. Results of Model 1 are shown in Table 5. The *z*-statistic for position ($z = 4.88$) indicated that there was a significant difference (at the $\alpha = .01$ level) in the probability of a correct response when irrelevant information was in the object position versus subject position, with the object position having a significantly higher probability of a correct response. In Model 1, the intercept represents the case when irrelevant information was in the subject position of a WP. Thus for subject position, the probability of correctly solving a WP was .39, (1/1 + exp[-{-.44}]), also controlling for basic arithmetic skill. For object position, the probability of correctly solving a WP was .52, (1/1 + exp[-{-.44 + .51}]), controlling for basic arithmetic skill. Overall, results indicate that irrelevant information position affects WP performance. Person factors were then assessed in research question 2 to help explain what else is important for explaining variability in accurate WP performance.

Table 5

Results of Model 1 (N = 140)

Fixed Effects					
	Variable	Coefficient	*SE*	*z*	*p*
	Intercept	-0.44	0.22	-2.06	.04
	Position	0.51	0.10	4.88	.00
	Arithmetic	0.04	0.01	3.82	.00
Random Effects					
	Source	Variance			
	Person (intercept)	2.49			
	Person (Position)	0.19			
	Classroom (intercept)	0.53			
	Item (Intercept)	0.12			

Note. Coefficients are in log-odds units.

Research Question 2

With research question 2, we estimated person effects on a student's probability of correctly solving a WP, while controlling for irrelevant information position and arithmetic skill. We added predictors to Model 1 to determine which person factors affected WP performance above and beyond irrelevant information position. At Level 2, the following person factors were entered: working memory for words and sentences, nonverbal reasoning, language comprehension, and grade. Results of this model (Model 2) are shown in Table 6. Controlling for irrelevant information position and arithmetic skill, language comprehension ($z = 4.48$, $p < .001$) was a significant predictor of WP performance, while working memory, reasoning, and grade were not. As hypothesized, language comprehension had a significant influence on WP performance. The predicted probability of solving a WP (with irrelevant information in the subject position) correctly for a person with language skills 1 *SD* below the mean (and average skills on the other person factors) was .26, (1/1 + exp[-{-.32 + .13*-5.68}]), but was .60, (1/1 + exp[-{-.32 + .13*5.68}]), for a person with language skills 1 *SD* above the mean.

Table 6

Results of Model 2 (N = 140)

Fixed Effects					
	Variable	Coefficient	*SE*	*z*	*p*
	Intercept	-0.32	0.26	-1.21	.22
	Position	0.51	0.11	4.88	.00
	Listening Recall	0.08	0.04	1.94	.05
	Language	0.13	0.03	4.48	.00
	Reasoning	0.07	0.04	1.80	.07
	Grade	-0.39	0.45	-0.88	.38
	Arithmetic	0.02	0.01	2.09	.04
Random Effects					
	Source	Variance			
	Person (Intercept)	1.68			
	Person (Position)	0.19			
	Classroom (Intercept)	0.59			
	Item (Intercept)	0.12			

Note. Coefficients are in log-odds units.

Research Question 3

As a final step in analyses, to address research question 3, we estimated a model (Model 3) testing two person-by-item interactions. This analysis was exploratory. The first interaction term was between nonverbal reasoning and irrelevant information position; the second, between working memory and position. To estimate this model, we increased the tolerance level for model convergence by .0001 to accommodate the most complex model (the default tolerance in R is .001). As results in Table 7 show, there were no significant interactions, perhaps due to inadequate statistical power to detect significant interaction effects.

Table 7

Results of Model 3 (N = 140)

Fixed Effects					
	Variable	Coefficient	*SE*	*z*	*p*
	Intercept	-0.32	0.26	-1.21	.23
	Position	0.51	0.10	4.87	.00
	Listening Recall	0.07	0.04	1.56	.12
	Language	0.13	0.03	4.46	.00
	Reasoning	0.07	0.04	1.80	.07
	Grade	-0.39	0.45	-0.88	.38
	Arithmetic	0.02	0.01	2.09	.04
	LR*Position	0.03	0.03	0.87	.39
	Reasoning*Position	-0.03	0.02	-1.05	.20
Random Effects					
	Source	Variance			
	Person (intercept)	1.68			
	Person (position)	0.17			
	Classroom (intercept)	0.59			
	Item (intercept)	0.12			

Note. Coefficients are in log-odds units.

Given the exploratory nature of these analyses, with the purpose of formulating hypotheses for subsequent research with larger samples, we depict the interaction between working memory (i.e., listening recall) and position in Figure 2. This graph suggests that with a larger sample, an interaction may be found in which students in general have a higher predicted probability of correctly solving a WP with irrelevant information in the object position than irrelevant information in the subject position, but that the difference between the two WP types is more prominent for students with greater working memory resources. Figure 3 depicts the interaction between reasoning and position. This graph suggests that with a larger sample, an interaction may be found in which students have a higher predicted probability of correctly solving a WP with irrelevant information in the object position than in the subject position, but that the difference between the two information subtypes is more prominent for students with lower reasoning ability.

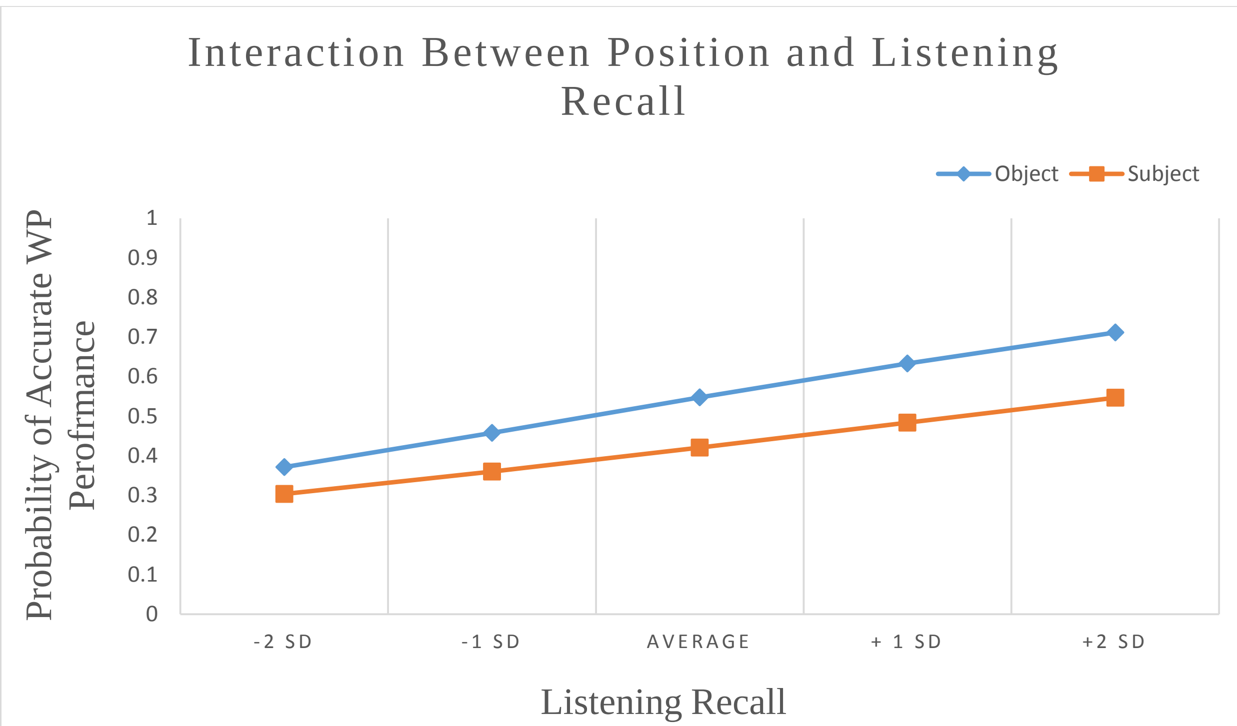

Figure 2. Graph of interaction between positioning and listening recall (i.e., working memory) when all other variables are held constant.

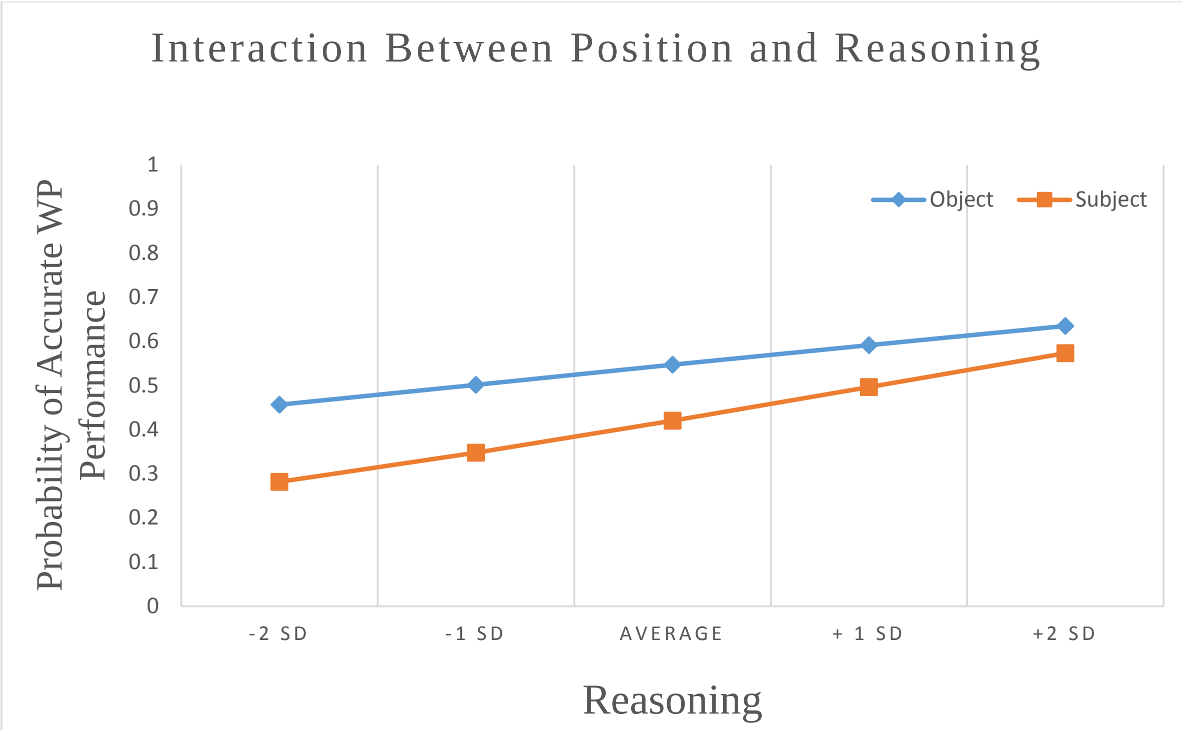

Figure 3. Graph of interaction between irrelevant information position and nonverbal reasoning when all other variables are held constant.

Error Analysis

We randomly selected 25% (N = 35) of student WP measures. Of 700 possible responses, students answered 410 problems incorrectly. Of incorrect responses, 17% were instances in which students correctly identified the two relevant values within the WP, but performed an incorrect calculation; 37% were instances of rote computation. Overall, students used rote computation 152 times out of 700 responses, performed an incorrect calculation using two relevant numbers 70 times out of 700 responses, and conducted a different process or undetermined strategy 188 times out of 700 responses.

CHAPTER IV

DISCUSSION

The purpose of the present study was to extend the literature on WP performance by using a crossed random-effects item response model to bring together two lines of research, namely irrelevant information and person factors that affect WP performance. This study combined covariates identified in prior research into one model so that effects could be evaluated in the presence of other potential influences. The study examined two levels of irrelevant information position (i.e., subject position vs. object position) while balancing the placement of irrelevant information (i.e., sentence containing irrelevant value placed before or after relevant information) and while considering four person-level factors (i.e., working memory, language comprehension, nonverbal reasoning, and grade).

In this section, we first discuss findings for the effect of irrelevant information position on WP performance when controlling for incoming basic arithmetic skill. Then, we consider the effects of the person-level variables on WP performance while controlling for irrelevant information position and incoming basic arithmetic skills and also consider possible person-by-item interactions. In conclusion, we suggest instructional implications and discuss study limitations.

Irrelevant Information Position

As expected, irrelevant information position influenced WP performance, when controlling for incoming arithmetic skill (Model 1), as well as language comprehension, working memory, reasoning, and grade (Models 2 and 3). As evidenced in our error analysis, students at times used rote computation when solving WPs with irrelevant information, but at other times, were aware of the presence of irrelevant information. Differences in student processes across responses may have been influenced by irrelevant information position.

Students did have greater difficulty with WPs containing irrelevant information in the subject position than in the object position. The probability of accurately solving a WP with irrelevant information in the object position was .52, whereas the probability was .39 for the subject position. These findings, however, were contrary to expectations and are inconsistent with those of Englert et al. (1987), who investigated the position of irrelevant information on the problem-solving performance of 48 second and fourth-grade students with learning disabilities and their typically developing peers. In that prior study, because WPs with an irrelevant object were more difficult than WPs with an irrelevant subject across student types, Englert et al. concluded that distractors in the subject position were easier to detect than distractors in the object position. Englert et al. further suggested that WP information associated with a subject serves as an important basis for identifying separate sets of information; students seek out *who* the WP is about and *how much* that individual has to identify relevant from irrelevant information. The difference in findings between Englert et al. and the present study may have occurred for three reasons.

First, in the present study, we examined the effects of irrelevant information position on WP performance while controlling for incoming arithmetic skill, working memory, language comprehension, nonverbal reasoning, and grade, whereas Englert et al. (1987) only controlled for initial arithmetic skill. Furthermore, in the present study, we controlled for superordinate set language within WPs; Englert et al. did not. For example, in Englert et al., a WP with irrelevant information in the object position contained superordinate language (see Problem A in Table 8).

Table 8

Word Problem Examples

Problem	
A[1]	Joy had 4 chairs. Joy had 2 tables. Joy had 4 pet frogs. How many pieces of furniture did Joy have
B[1]	Bill has 5 toy planes. Bill has 3 toy drums. Tom has 4 toy drums. How many toys does Bill have?
C	Bill has 5 toy planes. Bill has 3 toy drums. Marie has 4 toy drums. How many toys does the boy have?
D	Omega and Julian did school work on the weekend. On Saturday, Omega studied for 6 hours. On Sunday she studied for 2 hours. Omega studied for 3 tests. How many *hours* did *Omega* study for?
E	Omega and Julian did school work on the weekend. On Saturday, Omega studied for 6 hours. On Sunday she studied for 2 hours. Julian studied for 3 hours. How many *hours* did *Omega* study for?

Note. [1]These WPs are derived from Englert et al. (1987).

In Problem A, the question contains the word *furniture*, which requires a student to classify *chairs* and *tables* as pieces of furniture in order to identify the relevant information and reach a correct solution. Now consider Englert et al.'s example of a WP with irrelevant information in the subject position (Problem B in Table 8): "Bill has 5 toy planes. Bill has 3 toy drums. Tom has 4 toy drums. How many toys does Bill have?" (p. 31). This WP is likely less difficult for students than the former problem for a reason other than irrelevant information position. That is, a correct response depends on a student's ability to identify nonoverlapping sets of information linked to the relevant subject (i.e., Bill), without being further challenged by the presence of a superordinate category in the question. Thus, when a student searches the problem text for relevant information, the student does not have to recognize subsets of a larger or broader category. Now consider Problem C (Table 8), a re-wording of Problem B. To identify the relevant information and reach a correct solution, a student must still recognize a distractor in the subject position but at the same time, the student must categorize *Bill* and *Marie* as instances or non-instances of the superordinate category *boy*.

Because superordinate set language requires students to identify subsets of the superordinate set and then to identify relevant from irrelevant information, it places greater cognitive demands than do WPs without superordinate set language. Yet, Englert et al. (1987) concluded that students relied more heavily on subject information as a basis for identifying nonoverlapping subsets of information, even though findings appear to have been confounded by the presence of superordinate set language.

A third potential explanation for the difference in findings between Englert et al. and the present study is that: (a) students are attentive to and rely on object categories because they are preceded by numerals, and therefore (b) irrelevant information in the object position alerts students to check the object categories in the problem narrative against the information contained in the question. In the present study, students performed better on WPs containing irrelevant information in the object position than in the subject position. These results suggest that students are attentive to and rely on object categories.

In a study with 24 students with intellectual disabilities, Thibodeau (1974) also found that irrelevant information linked to the subject position posed greater difficulties than irrelevant information linked to the object position. He suggested that when students set up a problem, they tend to focus on object categories because these are typically preceded by numerals. This reasoning may account for the differential performance on WPs with irrelevant information in the subject position versus object position. For example, consider a WP with irrelevant information in the object position. Refer to Problem D in Table 8. As a student reads Problem D, the numerals precede the object categories of *hours* and *tests*. If a student ignores the question and computes all three numbers without recognizing *tests* as a different object category than *hours,* then the student will reach an incorrect solution. By contrast, if a student attends to the words following the numerals, then differing object categories may trigger the student to focus on the unknown in the question. That is, a student recognizes *tests* and *hours* as two different object categories and then proceeds to check this against the question's unknown, *hours*.

Now consider a WP with irrelevant information in the subject position. Refer to Problem E in Table 8. In problem E, a student has only one object category, *hours*, to process. Moreover, since all values have the same object category, a student may be more likely to compute with all three values, thereby ignoring the subject category of the question (*Omega*). Thus, students may perform better on WPs with irrelevant information in the object position because the object categories do not match, which alerts students to check the object categories against that contained in the question.

Based on this analysis, the following conclusion about the position of irrelevant information in WPs seems warranted: Irrelevant information position exerts a significant effect on student performance on WPs. Our findings suggest that students review semantic categories and their values and attempt to match the values of the semantic features requested in the problem question with those presented in the problem narrative. This supports Littlefield and Rieser's (1993) model of determining relevance.

Person Factors

Prior research has examined cognitive and linguistic predictors of WP solving but has not investigated the effects of person characteristics and item features concurrently. Nonverbal reasoning, sight word efficiency, phonological processing, working memory, and language comprehension have been identified as significant predictors of simple WP solving (e.g., Fuchs et al., 2005; Fuchs et al., 2006; Swanson & Beebe-Frankenberger, 2004). Item features such as complexity of semantic structures, extra step, and inclusion of irrelevant information have also been found to impact students' WP performance (e.g., Parmar et al., 1996; Russell & Ginsberg, 1984).

In the present study, person- and item-level factors were investigated concurrently. When controlling for arithmetic skill and irrelevant information position, we found that language comprehension uniquely contributed to WP performance, whereas working memory, nonverbal reasoning, and grade did not. Students with language scores 1 *SD* below the mean had a predicted probability of .26 for solving a WP with irrelevant information in the subject position correctly; for students with language comprehension scores 1 *SD* above the mean, the probability increased to .60. It is not surprising that language comprehension supports performance on WPs, given the need to process linguistic information when building a WP model. Our findings support previous work demonstrating the importance of language comprehension in WP solving (e.g., Fuchs et al., 2006; Fuchs et al., 2008; Tolar et al., 2012).

Interestingly, working memory for words and sentences (i.e., *Listening Recall* task) failed to achieve significance, when controlling for arithmetic skill and irrelevant information position. This is not consistent with previous studies in which WP solving depends on working memory (e.g., LeBlanc & Weber-Russell, 1996; Passolunghi & Siegel, 2001, 2004; Swanson & Beebe-Frankenberger, 2004). The effect of working memory is not, however, consistent in the literature. For example, Fuchs et al. (2006) examined the cognitive correlates of third-grade skill in arithmetic WPs and found that working memory did not uniquely contribute to WP solving.

Working memory only emerged as a significant predictor when the paths for phonological decoding and sight-word efficiency were set to zero. Fuchs et al. attributed these results to two possibilities: (a) reading or reading related processes influence the relations between cognitive abilities and WP solving, and (b) working memory was already captured within some of the other cognitive abilities entered in their models.

Our findings corroborate Fuchs et al. (2006). In the present study, working memory may have been captured by other factors entered simultaneously in our models. Furthermore, our findings may be attributable to the type of WP we used in this study. All the WPs in the present study were *combine* problems, in which two quantities are combined to form a total. Though all WPs contained irrelevant information, our WPs may have not been sufficiently complex to place a strong demand on working memory as compared to other problem types, such as *compare* problems, in which two quantities are compared to find a difference, or *change* problems, where a starting amount increases or decreases over time.

Despite similar simple computational demands requiring one-step addition or subtraction, combine, change, and compare problems are not equally difficult (Schumacher, 2012). Differential performance across these WP types may be associated with the semantic structure of the problem type (Briars & Larkin, 1984; Riley et al., 1983; Riley & Greeno, 1988). Further, only compare problems incorporate relational terminology, such as *more than*, *less than, fewer than* (Schumacher, 2012); change problems, while conveying a typically uncomplicated cause-effect story line, sometimes require addition, other times subtraction. This suggests that combine WPs (which do not incorporate relational terminology and always require addition) may lend themselves more easily to the use of a rote computational strategy, in which students add all numbers presented in a problem (Goodstein et al., 1971).

Without irrelevant information, a rote computational strategy will produce a correct answer when addition is required for accurate solutions in problems. For WPs with irrelevant information, however, students who typically rely on a rote computational strategy may be triggered to reexamine their computational processes. This triggering process may depend more on a student's language comprehension ability than working memory capacity because in order to build a situation model, a student must first recognize semantic categories and then identify disjointed sets of information to build a correct model of a WP.

As for other person effects, we expected nonverbal reasoning and grade to play key roles. However, these main effects were not significant. Prior research has identified nonverbal reasoning as a significant predictor of WP performance (e.g., Fuchs et al, 2006; Fuchs et al., 2015) and growth for high-complexity WPs (e.g., Tolar et al., 2012). It is possible our WP measure failed to capture a broad range of nonverbal reasoning demands across problems. As previously discussed, our WP measure contained only combine problems, which may not place strong demands on reasoning, even when the WPs contain irrelevant information. This may also explain why grade was not significant.

All this said, our main conclusion concerning person factors is that language comprehension uniquely contributes to student WP solving with irrelevant information. Present findings do not support a role for working memory, nonverbal reasoning, and grade.

Person-By-Item Interactions

For exploratory purposes, we examined the interaction between nonverbal reasoning and irrelevant information position as well the interaction between working memory and position. We did not find significant interactions. However, as depicted Figure 2, it appears that the difference in predicted probabilities of correctly solving a WP between the two subtypes of irrelevant information position is most prominent for students with higher working memory scores. In Figure 3, it appears that the difference in predicted probabilities of correctly solving a WP between the two irrelevant information positions is most prominent for students with lower nonverbal reasoning scores. These graphs suggest that with larger sample sizes, interactions may have reached statistical significance.

While findings of these exploratory analyses are inconclusive due to inadequate statistical power, it is interesting to note that the difference in WP performance between our two subtypes of irrelevant information was more prominent for students with greater working memory capacity, whereas the difference in WP performance between the two types of WPs was more prominent for students with lower nonverbal reasoning ability. Differential performance may be least prominent for students with lower working memory because they may not be able to store and manipulate semantic categories across the problem narrative and problem question. Therefore, irrelevant information position subtypes do little to alert students to check the semantic categories of numerals to the semantic categories contained in the problem question. By contrast, students with stronger working memory can store and manipulate information, thereby sorting relevant from irrelevant information using semantic categories. For these students, WPs with irrelevant information in the object position may provide an even greater advantage than those with irrelevant information in the subject position. As discussed previously, WPs with irrelevant information in the object position provide greater semantic clues that distinguish relevant from irrelevant numerals than WPs with irrelevant information in the subject position.

In the present study, we operationalized nonverbal reasoning with the WASI *Matrix Reasoning*, a task that first requires a student to identify the rules that govern a pattern and then compare the rule of the stimulus pattern with the characteristics of the possible choices. Thus, this measure taps categorical ability. Students who can reason analytically and make categorical judgments may be stronger in distinguishing relevant from irrelevant features of classes of subject or object within WPs. For students with lower nonverbal reasoning ability, irrelevant information in the object position may help students compensate for lower nonverbal reasoning abilities. As discussed above, when object categories do not match in a problem narrative, this alerts students to check the object categories against the unknown contained in the question.

In any case, the results of the exploratory analyses provide the basis for generating the following hypotheses concerning interactions between person effects and irrelevant information position: (a) the difference in WP performance between the two types of information position is more prominent for students with lower nonverbal reasoning ability; and (b) the difference in WP performance between the two types of information position is more prominent for students with higher working memory ability. Future research, with stronger statistical power, should address the tenability of these hypotheses.

Conclusions, Instructional Implications, and Limitations

In the present study, irrelevant information position uniquely contributed to WP performance, in the presence of nonverbal reasoning, language comprehension, working memory, and grade. Language comprehension also emerged as a significant factor across students, when controlling for all other person and item factors. No significant interactions were found; however, visual analysis of interaction graphs suggest that with larger sample sizes, interactions (i.e., working memory by position, nonverbal reasoning by position) may be detected.

These results suggest the need for two directions for WP instruction. First, WP solving interventions (e.g., Fuchs, Fuchs, Hamlett, & Appleton, 2002) that include instruction designed to help students solve WPs with irrelevant information should place an emphasis on varying irrelevant information features. Specifically, instruction on irrelevant information should focus on teaching students to compare the semantic categories asked in the question with those presented in the problem narrative.

For example, consider this WP: A softball team played in a tournament. On Saturday, Christina batted 2 homeruns. Molly batted 3 homeruns. On Sunday, Molly batted 1 homerun. How many homeruns did Molly bat? Here, the unknown refers to *homeruns*. To solve this WP, a student should be taught to focus not only on the unknown, which typically is an amount associated with an object, but also on the subject category that possesses or is associated with the unknown. In other words, a student should also attend to the subject (Molly) contained in the question, not only the object. Once a student identifies the subject (Molly) and object (homeruns) in the question, the student then checks this information against the subject categories (Christina, Molly) as well as object categories (homeruns) presented in the problem narrative to identify relevant from irrelevant information. Current interventions typically focus on teaching a student first to identify the unknown and from that unknown, to construct a label prior to solving the problem. In the above example, if a student were to focus only on the unknown (homeruns) and thereby the object category, the student may incorrectly add all three values presented in the problem.

The second instructional implication is that a focus on language comprehension instruction may represent a productive strategy for improving students' WP performance generally. Further, examining language comprehension as an overall predictor of responsiveness to WP intervention may be potentially informative in identifying directions for extending WP instruction, as shown in Fuchs et al. (2013). This is because language comprehension moderates the effect of different forms of WP intervention (Fuchs et al, 2013). Such research allows for the examination of factors associated with poor response to otherwise effective instruction or intervention, as well as the conditions that most effectively reduce the importance of language comprehension in learning to solve WPs

At the same time, it is important to consider our findings and conclusions with study limitations in mind. First, our WP measure included only one problem type. Because combine problems are generally less difficult for students than other problem types, our measure may have failed to capture a broad range of cognitive and linguistic demands. Furthermore, we operationalized each domain-general cognitive construct with a particular measure. Future research should consider using other measures or combining measures into a composite score to capture constructs. Changes in measures can alter patterns of student performance and study findings. Finally, our study was limited in sample size for the analyses conducted. Future research should include large samples to permit detection of interaction effects.

These limitations provide directions for extending research on WP solving. In future work, researchers interested in designing targeted interventions should explore additional person and item factors. Additional person and item factors may include reading skills, phonological processing, attention, and WP specific language. In extension, future work should not only incorporate these additional person and item factors, but also examine other person-by-word interactions to determine if certain types of children have particular difficulty with certain types of WPs. Future work should incorporate different problem types so that differential performance across grade levels may be detected. Overall, these studies may guide researchers in the development and testing of interventions that instruct students how to identify irrelevant information across varying subtypes of irrelevant information, as well as varying types of WPs.

REFERENCES

Bates, D., Chambers, J., Dalgaard, D. et al. (2013). R: A language and environment for statistical computing. *R Foundation for Statistical Computing.* Vienna: Austria.

Blankenship, C. S., & Lovitt, T. C. (1976). Story problems: Merely confusing or downright befuddling? *Journal of Research in Mathematics Education, 7*, 290-298. doi: 10.2307/748462.

Briars, D. J. & Larkin, J. H. (1984). An integrated model of skill in solving elementary word problems. Cognition and Instruction, 1, 245-296.

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

Englert, C. C., Culatta, B. E., & Horn, D. G. (1987). Influence of irrelevant information in addition word problems on problem solving. *Learning Disability Quarterly, 10*(1), 29-36. doi: 10.1037/0022-0663.97.2.224.

Fuchs, L. S., Fuchs, D., Compton, D. L., Hamlett, C. L., & Wang, A. (2015). Is Word-Problem Solving a Form of Text Comprehension? *Scientific Studies of Reading, 19*(3), 204-223. doi: 10.1080/10888438.2015.1005745.

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., . . . Fletcher, J. M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and word problems. *Journal of Education Psychology, 98*, 29-43. doi:10.1037/0022-0663.98.1.29.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Appleton, A. C. (2002). Explicitly teaching for transfer: Effects on the mathematical problem-solving performance of students with mathematics disabilities. *Learning Disabilities Research & Practice, 17*(2), 90-106.

Fuchs, L. S., Fuchs, D., Prentice, K., Burch, M., Hamlett, C. L., Owen, R., . . . Jancek, D. (2003). Explicitly teaching for transfer: Effects on third-grade students' mathematical problem solving. *Journal of Education Psychology, 95*, 293-305. doi:10.1037/0022-0663.95.2.293.

Fuchs, L. S., Fuchs, D., Stuebing, K., Fletcher, J. M., Hamlett, C. L., & Lambert, W. (2008). Problem solving and computational skill: Are they shared or distinct aspects of mathematical cognition? *Journal of Educational Psychology, 100*(1), 30-47. doi:10.1037/0022-0663.100.1.30.

Fuchs, L.S., Geary, D.C., Compton, D.L., Fuchs, D., Schatschneider, C., Hamlett, C.L., DeSelms, J., Seethaler, P.M., Wilson, J., Craddock, C.F., Bryant, J.D., Luther, K., & Changas, P. (2013). Effects of first-grade number knowledge tutoring with contrasting forms of practice. *Journal of Educational Psychology, 105*, 58-77.

Fuchs, L. S., Hamlett, C. L., & Powell, S. R. (2003). Math Fact Fluency and Double-Digit Additional and Subtraction Tests. Available from L.S. Fuchs, 228 Peabody, Vanderbilt University, Nashville, TN 37203.

Goldstein, H. (1987). Multilevel covariance components models. *Biometrika, 74*, 430-431. doi : 10.1093/biomet/74.2.430.

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology, 48*, 163-89. doi:10.1146/annurev.psych.48.1.163.

Kinne, L. J. (2002). Differential Memory of Irrelevant Information in Word Problems (Doctoral Dissertation). Available from ProQuest Dissertations and Theses Full Text. (Document ID 275672498).

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

LeBlanc, M. D., & Weber-Russell, S. (1996). Text integration and mathematics connections: A computer model of arithmetic work problem solving. Cognitive Science, 20, 357-407. doi:10.1207/s15516709cog2003_2

Leon, R. E. (1992). The effects of the presence of extraneous information in mathematical word problems on the performance of Hispanic learning-disabled students (Doctoral Dissertation). Available from ProQuest Dissertations and Theses Full Text. (Document ID 304006613).

Littlefield, J., & Rieser, J. J. (1993). Semantic features of similarity and children's strategies for identifying relevant information in mathematical story problems. Cognition and Instruction, 11(2), 133-188. doi: 10.1207/s1532690xci1102_2.

Muth, K. D. (1991). Effects of cuing on middle-school students' performance on arithmetic word problems containing extraneous information. *Journal of Educational Psychology, 83*(1), 173-174. doi: 10.1037/0022-0663.83.1.173.

Parmar, R. S., Cawley, J. R., & Frazita, R. R. (1996). Word problem-solving by students with and without mild disabilities. *Exceptional Children, 62*, 415-429.

Passolunghi, M. C., & Siegel, L. S. (2001). Short-term memory, working memory, and inhibitory control in children with specific arithmetic learning disabilities. Journal of Experimental Child Psychology, 80, 44-57. doi:10.1006/jecp.2000.2626

Passolunghi, M. C., & Siegel, L. S. (2004). Working memory and access to numerical information in children with disability in mathematics. Journal of Experimental Child Psychology 88, 348-367. doi:10.1016/j.jecp.2004.04.002

Pickering, S. & Gathercole, S. (2001). *Working Memory Test Battery for Children.* London: The Psychological Corporation.

Raudenbush, S. W. (1993). A crossed random effects model for unbalanced data with applications in cross-sectional and longitudinal research. *Journal of Educational Statistics*, *18*, 321-349. doi: 10.2307/1165158.

Raudenbush SW, Bryk AS. Hierarchical Linear Models: Applications and Data Analysis Methods. 2nd ed. Thousand Oaks, CA: Sage Publications, Inc; 2002.

Riley, M. S. & Greeno, J. G. (1988). Developmental analysis of understanding language about quantities and of solving problems. *Cognition and Instruction, 5*, 49 – 101.

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

Russell, R. L., & Ginsburg, H. P. (1984). Cognitive analysis of children's mathematical difficulties. *Cognition and Instruction, 1*, 217-244.

Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon. *Educational Psychologist*, *24*, 113–142. doi: 10.1207/s15326985ep2402_1.

Schumacher, R. F., & Fuchs, L. S. (2012). Does understanding relational terminology mediate effects of intervention on compare word problems? Journal of Experimental Child Psychology, 111(4), 607-628. doi:10.1016/j.jecp.2011.12.001

Stram, D. O. & Lee, J. W. (1994). Variance components testing in the longitudinal mixed effects model. *Biometrics, 50*, 1171-1177.

Tolar, T. D., Fuchs, L., Cirino, P., Fuchs, D., Hamlett, C. L., & Fletcher, J. M. (2002). Predicting development of mathematical word problem solving across the intermediate grades. *Journal of Educational Psychology, 104*(4), 1083-1093. doi:10.1037/a0029020

Van den Noortgate W, De Boeck P, Meulders M. (2003). Cross-classification multilevel logistic models in psychometrics. *Journal of Educational and Behavioral Statistics, 28*, 369–386. doi: 10.3102/10769986028004369.

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Wang, A., Fuchs, L. S., and Fuchs, D. (submitted). Cognitive and linguistic predictors of mathematical word problems with and without irrelevant information. Unpublished raw data.

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence.* San Antonio, TX: The Psychological Corporation.

Wilkinson, G. S., & Robertson, G. J. (2006). *Wide Range Achievement Test, 4th Edition.* Wilmington, DE: Wide Range.