THE ROLE OF SITUATION MODEL DIMENSIONS IN MATH WORD PROBLEMS

CHERYLL FITZPATRICK

The Role of Situation Model Dimensions in Math Word Problems

by

Cheryll Fitzpatrick

A Thesis submitted to the

School of Graduate Studies

in partial fulfillment of the requirements for the degree of

Masters of Science

Psychology Department

Memorial University of Newfoundland

July 2012

St. John's Newfoundland

Abstract

Math word problems can be quite challenging to students. In recent years, the math word problem literature has emphasized the importance of *situation models* (SM). SMs are a way of building a mental representation (using the text and previous knowledge) organized around five dimensions: protagonist, causal, motivational, temporal, and spatial. Wording of math word problems was manipulated to reflect two of these dimensions – spatial and motivational. Spatial (associated and dissociated) and motivational (motivational and neutral) questions were given to grades 3 and 5 students. Gender and grade were found to interact with performance in the spatial dimension, while only gender interacted with performance in the motivational dimension.

Keywords: word problems, situation models, motivational dimension, spatial dimension

Acknowledgements

First and foremost I would like to thank my supervisor, Dr. Darcy Hallett, whose knowledge and guidance this project would not have been possible. I would also like to thank the members of my committee, Dr. Carole Peterson and Dr. Toni Doyle. Your input has proved to be invaluable and I thank you for your time and commitment to this endeavor. The amount of time and dedication I have been able to devote to my research is in great part owed to the funding provided by the Social Sciences and Humanities Research Council. I owe many thanks to all the members of the Research Centre for the Development of Mathematical Cognition for your help with creating questions, data collection, and most of all having to listen to numerous presentations on the same topic! Finally, I would like to thank my family and friends for all of your support and words of encouragement, especially Neill, for the months of cooking and cleaning with no complaints while I collected data and wrote my thesis.

Table of Contents

List of Tables

List of Figures

The Role of Situation Model Dimensions in Math Word Problems

Math is considered to be a problematic subject for most children (Clements, 2010; Ginsberg, 2010; Cohen Levin, 2010) and continues to become more problematic as students progress through their education, especially when concepts such as word problems are introduced (Aiken, 1972; Ballew & Cunningham, 1982; MacLeod, 1992 as cited in Verschaffel & DeCorte, 1997; Son, Thai, Burke, & Kellman, 2010). *Word problems* are short, made-up stories using quantitative relations between various objects/characters that require a mathematical solution (Martin & Bassok, 2005). The National Mathematics Advisory Panel (U.S. Department of Education, 2008) reports that word problems are one of the three areas for which students have the poorest preparation. Yet, despite the development of specific teaching instruction to contend with students' lack of success, the problem still remains.

The literature on success in formulating and solving mathematical word problems in both children and adults has been examined in many different ways. Issues that have been considered include computation ability (Muth & Glynn, 1985), reading comprehension (Chase, 1960; Muth & Glynn, 1985) as well as reading ability of mathematical word problems (O'Mara, 1981); syntactic structure of the word problem (Linville, 1970 as cited in Aiken, 1972); vocabulary level of the individual (Dunlap & McKnight, 1978; Linville, 1970 as cited in Aiken, 1972); translation ability (Dark & Benbow, 1990); verbal-logical and visual strategies (Kaizer & Shore, 1995); linguistic knowledge, and knowledge about the schema of school word problems (Verschaffel & DeCorte, 1997). Though there are many diverse factors affecting a person's ability to approach math word problems, this list is by no means exhaustive and each factor plays an essential role in an individual's ability to master math word problems.

The recent literature in this field, however, has shifted focus to the importance of situation models in the comprehension of word problems (Vincente, Orrantia, & Verschaffel, 2008). *Situation models* (SM) are integrated mental representations of described states of relationships that are amalgamated from information stated explicitly in the text and information already known. These models are designed to explain text comprehension as well as comprehension in other modalities (Zwann & Radvansky, 1995). Situation models specify that any given situation can be described by five dimensions (protagonist, causation, motivational, temporal and spatial) and the associations and disassociations with these dimensions. Although SMs are used to help explain general comprehension (e.g., Zwann & Radvansky, 1998), they have also been fruitfully applied to the understanding of word problems (Coquin-Viennot & Moreau, 2007; Vincente et al., 2008; Thevenot, Devidal, Barrouillet, & Fayol, 2007).

The main objective of this research was to further explore the potential of SMs in understanding the solving of math word problems. Recent research has only begun to apply SMs to word problem understanding, so there remain many gaps in this effort as well as a need for more specificity. The present study attempts to fill two such gaps. First, although research has investigated the role of the temporal dimension (Moreau & Coquin-Viennot, 2003; Thevenot & Oakhill, 2005, 2006) and the protagonist dimension (Stern & Lehrndorfer, 1992), the motivational dimension has so far been overlooked in word problem research. The study below manipulates motivational aspects of word problems to see if increasing a character's motivation will make a problem easier to understand (or perhaps more salient), and therefore easier to solve.

The second aspect of SMs that this study explores is children's understanding of the association between the spatial dimension and the protagonist dimension. Recent research concerning the protagonist dimension has shown promising results by demonstrating not only the importance of the protagonist (e.g., Davis-Dorsey, Ross, & Morrison, 1991), but also how the protagonist can cause problem interference by being spatially associated with distracting information (e.g., Mattarella-Micke & Beilock, 2010). This last effect, however, has only been demonstrated in adults, and only in terms of interference slowing reaction time. The present study instead investigates children and whether their actual performance, and not reaction time, can be affected by interfering information that is either spatially associated or spatially disassociated.

Importance of Situation Models in Math Word Problems

Before describing the dimensions of SMs in more detail, it is worthwhile to describe how SMs have been used to explain word problem comprehension. Zwann and Radvansky (1998) have detailed four reasons as to why SMs are necessary for the comprehension of text in general.

First, the integration of information across sentences is necessary to understand a verbal math problem. For example, take the following word problem from Verschaffel and DeCorte (1997):

> Joe had some marbles. Then Tom gave him 5 more marbles. Now Joe has 8 marbles. How many marbles did Joe have in the beginning? (p. 72)

The integration of information across sentences allows the reader to create a SM that includes both Joe and Tom. On reading the first sentence Joe is placed in a person's short-term memory as a potential protagonist. When the second sentence is read ('Tom gave *him* 5 more marbles'), then Joe is retrieved from short-term memory as being the 'him' to whom Tom is referring (Zwann & Radvansky, 1998). The reader has created a mental representation in their mind based on the initial information of Joe having some marbles; when 'him' is referred to in the second sentence the reader knows 'him' is the character, Joe. The ability to integrate sentence information is vital for creating a SM.

Being able to integrate information across many different modalities (e.g., visual, written, and audio) is the second reason why SMs are necessary for comprehension. Visual images such as graphs, for example, can aid in comprehension by being mutually incorporated with information taken from a text into an integrated SM (Zwann & Radvansky, 1998). Skills in comprehending written media are strongly related to skills in comprehending auditory and visual media (Gernsbacher, Verner, & Faust, 1990), which supports the notion that multi-modal information is combined into a single SM. Nevertheless, word problem research suggests that highly able math students with relatively equally developed verbal and visual abilities tend to subordinate their visual analyses to verbal-logical analyses (Kaizer & Shore, 1995). In other words, less skilled comprehenders may have trouble with math word problems because of a lack of verbal-logical skills, and not because of a lack of visual skills (Gernsbacher et al., 1990; Kaizer & Shore, 1995).

Third, SMs explain how previous content knowledge can aid comprehension of new text. High-knowledge readers (i.e., those with domain expertise) with less verbal ability than low-knowledge readers can more easily construct a SM by retrieving information from their long-term memory whereas low-knowledge readers may have to produce a SM from no prior knowledge (Zwann & Radvansky, 1998). This role of expertise has also been demonstrated in the math problem literature. On a task that involved writing equations to represent sentences, Dark and Benbow (1990) found that mathematically talented junior high students performed better than their average peers, and college undergrads, but only performed better than their verbally-talented peers when the equations were complex. Dark and Benbow proposed that, when the problems became more complex, the mathematically talented students performed best because of their domain specific knowledge.

Translation from other languages represents the last demonstration of the role of SMs in text comprehension. If meaning was solely connected to words, then story elements from different languages could not be integrated into a more complete understanding. A SM provides a way for elements of meaning in different languages to be combined into one representation (Zwann & Radvansky, 1998). Translation could also be pertinent for the successful comprehension of a math word problem, as math has been argued to be a distinct language (Aiken, 1972; Austin & Howson, 1979; Hall & Fuson, 1986; Kane, 1968; Rothman & Cohen, 1989). Undergraduate students, for example, are quicker to solve numerical equation problems in comparison to solving word problems (Mayer, 1982) and have an increased performance when solving a word problem in comparison to formulating an equation from a word problem (Martin & Bassok, 2005). Essentially, it is easier to solve a math word problem than it is to restate the word problem in the language of mathematics, which is in the format of an equation.

If SMs are a vital part of comprehension tasks in general, SMs should also be just as important for the comprehension of math word problems. The next section reviews the literature to date that has applied SMs to math word problem understanding.

Dimensions of Situation Models

As previously mentioned, there are five dimensions used to describe SMs. These dimensions are protagonist, causal, motivational, temporal and spatial (Zwann & Radvansky, 1998). Each dimension enhances a SM through the use of *foregrounding* – highlighting, or changing the accessibility of content in a text (Mattarella-Micke & Beilock, 2010) by creating and maintaining a retrieval cue to the information in short-term working memory (Zwann & Radvansky, 1998). Using these dimensions as an organizational framework, the following subsections will review the various studies that apply SMs to math word problem comprehension.

Classifications of these studies into these dimensions, however, are not straightforward. Although a few studies actively aim to represent one or more of the five dimensions, most make no reference to the dimensions at all, instead referring to SMs more generally. Most of the studies reviewed below will therefore be included on the respective dimension based on an analysis of their method, even though the authors themselves did not identify a dimension.

Lastly, it should be noted that the authors in these research studies use different terms for SMs – qualitative situation model (Coquin-Viennot & Moreau, 2007), episodic situation model (Reusser, 1989; Coquin-Viennot & Moreau, 2003), mental model (Thevenot, 2010), and problem model (Hegarty, Mayer, & Monk, 1995). Nonetheless, all of these variants, regardless of differences in terminology, can be said to represent situation models as it has been defined in this paper.

Protagonists and Objects

The first and perhaps most fundamental dimension in situational models is the protagonist. The protagonist is typically the main character in a story, though the "main character" can be an object as well as a person, and there can also be more than one protagonist. Protagonists are given more attention and become more accessible in memory if they are given a proper name. Comprehension of objects in a text for the purpose of creating a SM is typically related to how an object is coupled with a protagonist, the intentionality of the object or the inferred intention, and the protagonist's goal while using the object (Zwann & Radvansky, 1998).

In the math word problem literature, only one study seems to deal directly with the protagonist dimension. Davis-Dorsey et al. (1991) wanted to examine applications to mathematics problem solving as being specifically conveyed through *context personalization* – information and specific referents about the individual learner. In creating a personalized condition the authors were able to introduce a protagonist to the word problem that was familiar (in name) to the individual. To study this question, they had 2nd and 5th graders answer a short biographical questionnaire with information such as the child's name, the name of their pet or best friend, etc. The authors then created specialized math word problems for all of the children based on information from their biography. That is, they gave children problems where the protagonist had the same name as the child or one of his or her friends (personalized problems) and compared their performance on these problems to ones that had not been personalized. Both groups of children performed better on personalized problems, suggesting that the act of giving a protagonist a name that is significant to the child somehow helps that child to do better on the problem in question.

The SM dimension of protagonist provides a focus in the story problem, not to mention that SMs are built primarily on the protagonist in the story. There is much promise for research in the protagonist/object dimension in word problems beyond this one study, especially given that each of the other four dimensions can affect the way a protagonist/object is viewed by the reader, as well as how a protagonist is viewed in combination with an object (Mattarella-Micke & Beilock, 2010; Zwann & Radvansky, 1998).

Causal

The causal dimension of SMs reflects an individuals' comprehension of causal information within a text. Relationships can be inferred or openly stated using words such as *because, so,* or *therefore* (Zwann & Radvansky, 1998). For example, Coquin-Viennot and Moreau (2007) wanted to determine if 3^{rd} and 4^{th} grade students used *Qualitative Situation Models* (which we will simply call Situational Models or SM) or *Problem Models* (PM – formalized representations that include texts, logico-mathematical relations and numerical values). To test this hypothesis, they devised word problems, for example:

> At the beginning of the year, a farmer had a flock of 22 sheep. By the end of the summer, the size of the flock had increased by 9, and by the following spring, **as a lot of lambs had been born**, the flock totalled 42 sheep. Did the size of the flock increase or decrease during the winter? (see Appendix, Coquin-Viennot & Moreau, 2007)

The PM states information regarding the numerical change in the size of the flock from 22 to 42, while the SM describes the change in the size of the flock using everyday terms, such as "*a lot of lambs had been born*". The problems were either consistent (PM and SM matched) or inconsistent (PM and SM did not match). Although Coquin-Viennot and Moreau (2007) did not mention any of the SM dimensions, their manipulation of the language in the word problems suggested that they were using a conflict in causality to make the SM inconsistent with the PM. In the example above, the phrase *"as a lot of lambs had been born"* makes the SM consistent with the PM, as the action of a lot of lambs being born should cause an increase in the number of sheep. The inconsistent version of this problem replaces this phrase, *"as a lot of lambs had been born"* with, *"as the wolf had devoured some of the sheep"*, which provides SM information that would contradict the PM. In other words, the action of the wolf devouring some of the sheep should cause the number of sheep to decrease (SM), but this is inconsistent with the increase in the numbers in the problem (PM).

After presenting these problems, Coquin-Viennot and Moreau asked the children two questions. The first could be answered using the SM (i.e., was there an increase or decrease?) while the second required the PM (i.e., by how much?). There were more errors on the first question from younger students on the inconsistent problems, suggesting these students relied more on the SM. There were also more errors found in younger students for the second question; however, the authors did not indicate whether the errors made were a result of children using the SM representation (i.e., the choice of addition or subtraction was consistent with the SM, but the correct numbers were used) or number errors (i.e., the correct numbers were not used). In other words, it is not clear whether children were getting questions wrong because they chose the operation based on the SM rather than the PM, or whether the inconsistent condition confused them to the point of using the wrong numbers.

Martin and Bassok (2005) wanted to know if students (junior high – college) use object relations when solving division word problems. This study, though not designed to represent any one dimension, seems to fit into the causal dimension, as the authors used *symmetrical* (categorically related pairing) and *asymmetrical* (functionally related pairing) object relations to provide semantic cues in the word problem. Asymmetrical sets always supported the correct solution and provided background knowledge supporting the correct size of the compared sets. Take the following example: "At a certain university, there are 3,450 students. There are 6 times as many students as professors. How many professors are there?. The reader should know that universities have fewer professors than students, which should help them to know that their answer should be less than 3,450. Symmetrical sets, on the other hand, did not support the correct mathematical solution and did not provide relevant background knowledge about the relative number of objects in the compared sets. Take the following symmetrical problem: "On a given day, a certain factory produces 3,450 nails. It produces 6 times as many nails as screws. How many screws does it produce?" In this problem, there is no real reason to expect nail to be more numerous than screws, or vice-versa. In other words, students could have used causal or relational information to help them solve problems with the asymmetrical sets, but not with the symmetrical sets. Students performed better on problems that used asymmetrical objects than problems using symmetrical objects. The object relations, specifically the asymmetrical objects, provided situational background knowledge and supported the correct solution [e.g., easy to determine that there should be more cookies than cookie jars (asymmetrical), harder to determine if there should be more apples than oranges (symmetrical)].

Stern and Lehrndorfer (1992) used a story with qualitative comparisons as a preamble to a word problem; this was used to provide a situational context and not as a way to demonstrate causality in the problem. Using a between-subjects design, three types of problems were created, all involving two story characters. The first group received problems where there were no comparisons between the characters. The second and third groups, however, received problems that explicitly stated that one character was superior to the other in various ways (i.e., older, gets more money, better toys, larger bedroom). In the case of the second group, these problems then described a scenario where the superior character would receive more of something. For the third group, it was the inferior character that would end up receiving more of something. The implicit model of causation being tested was that *because* of this imbalanced relationship, the superior character would be more likely to have better/more of the object(s) described in the word problem.

Children in the 1st grade performed better when the story before the word problem provided some kind of comparative situation than the children who received a story with no such comparison (Stern & Lehrndorfer, 1992). Yet, there was no performance difference between the second group (where the superior character received more) and the third group (where the inferior character received more). In other words, children performed better with a comparative situational context regardless of whether the story and the problem were consistent or inconsistent. Perhaps the neutral story is seen as irrelevant (not providing additional cues) and may be comparable to not being provided a story at all. It may also be the case that, despite the authors' assumptions, children do not have an implicit model of causation that implies that a superior character should always get more.

Motivational

The third dimension is that of motivation/intentionality. These terms are used interchangeably in the literature, but, for the sake of this paper, this dimension will be referred to as the motivational dimension. Motivation is a reason for setting and completing a goal, and information is more easily retrievable in memory when it is stated as a goal. Interestingly, goals that have been achieved are not as accessible as goals that have not been achieved. This is because completed goals are converted to long-term memory while incomplete goals remain in short-term memory (Zwann & Radvansky, 1998).

Although there has been very little research on word problem understanding and motivation, one series of studies seems to investigate the motivation dimension by manipulating the place in the text where the goal is stated. Thevenot and colleagues (Thevenot & Oakhill, 2005, 2006; Thevenot et al., 2007) explored whether placing the word problem question prior to the word problem information had an effect on an individual's performance. In other words, some students received problems in a more standard way (e.g., "John has 39 marbles, Tom has 17 marbles, and Paul has 16 marbles. How many marbles do John, Tom and Paul have altogether?") while others received modified questions that had the inquiry first (e.g., "How many marbles do John, Tom and Paul have altogether? John has 39 marbles, Tom has 17 marbles, and Paul has 16 marbles"), Thevenot and her colleagues found undergraduate students did have fewer errors when the question was asked before the problem was presented (Thevenot & Oakhill, 2005), and this had more of a facilitatory effect for students with low memory span (Thevenot & Oakhill, 2006). In another study with fourth grade children (Thevenot et al., 2007), placing the question before the text resulted in shorter mean self-presentation times (i.e., children were quicker to advance to the next page when reading the problem in a series of screens presented on a computer). Furthermore, student performance was best when the question was placed at the beginning of the text for both groups, especially for less-skilled children. This series of studies are categorized here in the motivation dimension because, by placing the question first, there is an increase on the emphasis of what needs to be solved (i.e., the goal). This may help students to be more attentive to the goal of the problem and, therefore, be better able to solve it.

Aside from these studies looking at asking the question first, the motivational dimension is greatly underrepresented in the literature on SMs and how they can be of a benefit in math word problem success. Issues that need to be considered are not only the motivation of the individual attempting the problem but the motivational aspect of the protagonist within the story problem.

Temporal

The fourth dimension is temporal order, as it is necessary for the reader to know when events have taken place both in relation to each other and to the time at which they were narrated (Zwann & Radvansky, 1998). Reusser (1989) claims one of the ways in which situational comprehension is obtained via the problem text is through temporal and functional analysis by searching for the initial and resulting state.

In one of the few studies to explicitly target one of the SM dimensions, Vincente et al. (2008) wanted to determine whether extra-situational information was useful for problem solving when the difficulty of the task had not only a mathematical but also a situational source. To examine this question, they used a temporal manipulation in word problems attempted by 3rd, 4th, and 5th graders. First, problems varied in situational difficulty between easy and hard. Situation difficulty was easy if a natural order of events was described the way they would happen in real life, for example:

> *Two days ago* Peter had 37 meters of cable. *Yesterday* Peter bought 100 more meters of cable than those he already had. *After buying those meters of cable he began a renovation. While making the renovation* he has used some meters of cable, and *when he finishes* there are 11 meters of cable left. How many meters of cable has Peter used? (see Appendix, Vincente et al., 2008).

The situation difficulty level was considered hard if the information in the initial moment was provided at the end of the problem text, for example:

> *Yesterday* Peter bought 100 more meters of cable than those he already had. *After buying those meters of cable he began a renovation. While making the renovation* he has used some meters of cable, and *when he finishes* there are 11 meters of cable left. *Two days ago* Peter had 37 meters of cable. How many meters of cable has Peter used? (see Appendix, Vincente et al., 2008).

Second, problems either contained or did not contain extra information highlighting temporal structure in the situational context (shown in the above examples with italics). This consisted of emphasizing the position of each moment in the temporal sequence (i.e., initial state, first change, second change, and final state), by including phrases indicating sequential actions (e.g.,"*after* buying those meters of cable he began a renovation"). Interestingly, the results not only indicated that there was no effect of situational difficulty, but also that extra situational information did not provide any additional support for students and even hindered the ability of low achievers' performance (Vincente et al., 2008). Although this is only one study, it suggests the possibility that the temporal dimension may not facilitate the solving of math word problems.

Spatial

Spatial location is the fifth and final dimension in SMs. In general, spatial characteristics have been related to recall in text comprehension tasks. Objects in a story that are spatially closer to each other or to the protagonist are more easily remembered when an individual is asked to recall the story (Rinck & Bower, 2000; Zwann & Radvansky, 1998). Spatial location can be difficult to explain in text because two objects can be close in space, but their descriptions can be far apart in the text (and vice-versa). Yet, it is the spatial closeness rather than the closeness of the object descriptions in the text that predicts how well objects are remembered together (Zwann & Radvansky, 1998).

Coquin-Viennot and Moreau (2003), investigating one particular use of the spatial dimension, wanted to explore if using a structuring word would facilitate a factorization strategy for solving word problems. Using a structuring concept, the authors created a spatial dimension by bringing objects spatially closer together. For example, a structuring word such as *bouquet* was used to group flowers together in a word problem (i.e., "For a prize-giving, for each of the 14 candidates, the florist prepares a bouquet made up of 5 roses and 7 tulips. How many flowers does the florist use in total?") This was compared to a question that did not have the structuring word (i.e., the same problem without the word "bouquet"). If the factorization strategy were used, then children would solve this problem using the formula "14(5 + 7)". That is, they would group the flowers in the bouquet together first (i.e. 5 + 7 = 12) before multiplying by 14. If not, children might use a distributive strategy, which involves multiplying each of the groups of flowers by 14 and then adding them [i.e., (14 x 5) + (14 x 7)]. Results demonstrated that the presence of a structuring word did elicit a factorizing strategy, and this effect was stronger for 5th graders than for 3rd graders. Using a factorization strategy supersedes the use of a distributive strategy in the way of reducing the amount of work to complete the problem as well as limiting the opportunities for creating mathematical errors. It appears that information that emphasizes the spatial grouping of objects can affect the approach taken to solve a math word problem.

Mattarella-Micke and Beilock (2010) studied the effect of spatial information in relation to the protagonist. This effect, however, was investigated indirectly by testing whether or not spatial closeness can cause more interference in problem solving, using Siegler's *Distribution of Associations Model* (DOA) (1998, as cited in Mattarella-Micke & Beilock, 2010). According to the DOA, effectiveness of retrieval depends on the associative strength of the correct answer, relative to the incorrect answer. So, when someone is asked a word problem that is solved by the equation "3 x 2", then the number 6 should be highly activated. If, however, the number "5" is highlighted somewhere else in the problem, this may cause interference, because 5 is also highly associated with the numbers 3 and 2 (i.e., because 3 + 2 = 5) and may be almost as highly activated as "6". Mattarella-Micke and Beilock tested whether this interference would be evident in word problems and whether or not this interference would be different if the interfering number (i.e., the "5" in the above example) was spatially associated with the protagonist or spatially disassociated with the protagonist. A sample of 38 undergraduate students were given 72 word problems from a computer program; half were multiplication while the other half were filler division problems. Problems were divided into numerical high/low interference types, and also divided into spatially associated/dissociated types. In the high interference condition, the number was highly interfering with the addition answer (i.e., 5 for 3 x 2 = 6). In the low interference condition, the number was not interfering and was randomly selected from non-interfering numbers between 4 and 18. The association factor varied by changing a few words in the problem to foreground either a spatial association of the object with the protagonist (i.e., picked up, or carried) or a spatial dissociation of the object from the protagonist (i.e., put down, or left behind).

The results showed an interference by association interaction. Performance for the high interference number problems were worse than the low interference problems in the associative condition, while significance was not reached for the dissociative condition. This means that students' performance was only affected by the numerical content of a particular number set if the interfering number of objects was spatially associated with the protagonist. More specifically, the number "5" provided interference in a 3 x 2 word problem only when that number was spatially associated with the protagonist, and not when it was spatially disassociated. Mattarella-Micke and Beilock also found that the extent of this interference was related to working memory.

Spatial information in the literature on general text comprehension has been the most studied dimension and considered to be the most closely associated with constructing a SM (Zwann & Radvansky, 1998). Unfortunately, this does not hold true for the literature in math word problem comprehension. There is definitely a need to expand on the research in this area to determine how spatial information can be used in word problems to help construct SMs that will allow individuals to clearly comprehend the problem text.

The Current Study

As the research summarized above demonstrates, SMs have been applied fruitfully to the study of word problem understanding. Throughout the literature, results have indicated those students considered to be successful problem solvers were much better at formulating/representing a SM (Dark & Benbow, 1990; Kaizer & Shore, 1995; Thevenot & Oakhill, 2006; Thevenot et al., 2007; Vincente et al., 2008). At the same time, the above research has used SMs in a general sense and has rarely specifically examined the roles of the five dimensions of SMs. If SMs are to explain how children solve math word problems, then the role of these dimensions should be investigated.

The study described below aims to fill this gap by examining the role of two of these dimensions in grade 3 and grade 5 children. First, the study considered whether motivational information would make word problems easier to solve. The motivational dimension was chosen because there is a particular lack of research in the area of math word problem performance that could be classified under this particular dimension. Furthermore, the studies that can be considered motivational (Thevenot & colleagues, 2005, 2006, 2007) do not consider the motivational aspects of the protagonist within the story problem. Students were presented with some problems in which the protagonist is described as motivated to find the answer and some problems that are missing this motivational information.

Second, the study investigated whether performance would improve on spatially dissociated questions, and if the spatial association of interfering information could affect the effectiveness of interference. Modeled after the study by Mattarella-Micke and Beilock (2010), who demonstrated this effect in adults, the spatial dimension was examined by giving children problems that include extraneous numbers that may interfere with the solution of the problem. In some of these problems, the interfering information were spatially associated with the protagonist, while in the other problems, the interfering information was spatially disassociated from the protagonist. As noted above, Mattarella-Micke and Beilock have found interesting results in an adult population. The current study investigated whether a similar effect can be found in children.

Furthermore, given the multitude of research that has found differences in performance based on an individual's level of skill in the construction of situation models (Dark & Benbow, 1990; Hegarty et al., 1995; Kaizer & Shore, 1995; Moreau & Coquin-Viennot, 2003; Thevenot & Oakhill, 2006; Thevenot et al., 2007; Vincente et al., 2008), students were assessed on their general word problem solving ability as well as their general cognitive ability to explore whether these abilities interact with the influence of these SM dimensions. The inclusion of these measures will make it possible to test whether the effect of SMs might vary according to ability level (e.g., less-able children may be more helped by SM information than more-able children, or vice-versa).

Finally, this study also considered gender differences. Despite the stereotype of boys doing better in math, overall gender differences in math are (for the most part) non-existent in the general population (Ben-Zeev et al., 2005; Brannon, 2011; Delgado & Prieto, 2004; Hyde, Fennema, & Lamon, 1990). When math word problems are specifically examined, however, gender differences do appear. Despite girls getting better grades in math courses at all grade levels (Byrnes, 2005), boys are consistently found to perform better than girls on math word problems in studies of both children (Delgado & Prieto, 2005; Geary, 1996; Lummis & Stevenson, 1990; Marshall & Smith, 1987;

Stevenson et al., 1990) and adults (Hyde et al., 1990; Johnson, 1984). For this reason, the current study will examine possible gender differences in the effect of the spatial and motivational dimensions on word problem performance.

Method

Participants

A total of 47 grade 3 (26 boys and 21 girls) and 58 grade 5 (32 boys and 26 girls) students were recruited from five elementary schools in St. John's, Newfoundland. Twenty-one children were dropped from the analyses for various reasons [e.g., language barriers (2), learning difficulties (1), reading difficulty (1), procedural error (12), and other (5)]. Procedural errors were due to booklet copy errors where either a combination of motivational problem sets were missing, questions had been doubled up, or questions were missing (see Materials and Procedure for details). The final sample contained 40 grade 3 (21 boys and 19 girls, mean age = 8.62, SD = .426) and 44 grade 5 students (20 boys and 24 girls, mean age = 10.75, SD = .283). Because the existing word problem literature has most often studied Grade 3 and 5 students (e.g., Coquin-Viennot & Moreau, 2003, 2007; Davis-Dorsey et al., 1991; Moreau & Coquin-Viennot, 2003; Vincente et al., 2008), the current study aimed to use the same grade levels.

Materials

Raven's Standard Progressive Matrices (Raven, Raven, & Court, 1998). The Raven's Matrices is a standardized non-verbal, multiple-choice reasoning task used to assess problem solving ability, general intelligence and cognitive ability. Raven and colleagues report a split-half internal consistency modal value of .91. It is designed to be useful for persons of all ages. The measure is made up of 60 problems – five series of 12 problems each – of diagrammatic puzzles with a part missing, which the test taker must find among the options provided. The problems begin as nearly obvious and increase in difficulty. Participants received a smaller selection (28 items) chosen based on age norms deemed appropriate for grade 3 and 5 children. Age norms for this measure range from children five and a half years-of-age to adults 85 year-of-age, which makes it appropriate for the children sampled in this study (Raven, Raven, & Court, 1998).

General Math Word Problems. Word problems consisted of the four basic operations (e.g., addition, subtraction, multiplication, and division), and were taken and modified from Math Makes Sense 2 and 3 (Ball et al., 2008; Appel et al., 2009), and Mathfocus 4, 5, and 6 (Hope, Klassen, Small, Tam Tseng, & Tossell, 2008; Brydon et al., 2008; Canavan-McGrath et al., 2010). These are the approved textbooks for the relevant grades for the provincial curriculum. Word problems were chosen to reflect varying operations and difficulty levels to ensure that all students could answer some of the problems while avoiding ceiling effects. This measure was used to provide an assessment of each child's skill level regarding general math word problems. It also served as a proxy for reading ability, at least in the specific context of solving math word problems (see Appendix A for sample items).

Spatial Math Word Problems (Mattarella-Micke & Beilock, 2010). Spatial word problems consisted of interfering information being associated or dissociated with the protagonist. These word problems were taken and modified from Mattarella-Micke and Beilock (2010). All of the word problems were multiplication and were necessary if an interference effect was going to be found by including interfering addend numerical information. There were six spatial word problems, each having two versions; one version indicated an association of an object to the protagonist, while the other version was the same question modified slightly to include a dissociation of an object from the protagonist.

High interference in the spatial condition was established by having the numeral in the introductory sentence of the word problem be equal to the addition solution to the multiplication math fact (i.e., 7, for 3 x 4). For low interference problems, the numeral in the introductory sentence was a randomly selected non-interfering number from 4 to 18. The *association factor* varied by changing one to three words in a sentence, either associating or dissociating an object with the protagonist (see Appendix B for sample items). For example, consider the problem of Earl. Earl, we are told, has to figure out how many seats are available in the library, but also has to deal with some assignments. The spatially dissociated version of this problem reads, "Earl **dropped off** 9 assignments he had just finished and left for the library." The spatially associated version, however, reads, "Earl **picked up** 9 assignments he had just finished and left for the library." Apart from these small wording changes, the spatially associated and the spatially dissociated versions of these word problems were identical.

Motivational Math Word Problems. Motivational word problems were those using varying phrases indicating a desire for the protagonist to know the solution to the problem or neutral statements indicating no desire for a solution to the problem. Six motivational word problems were used, each problem having two versions; one version included the motivational phrase, while the other version replaced the motivational phrase with a neutral phrase of equal word length. The *motivational factor* varied by changing a statement in the word problem; the motivational condition included a statement indicating success being significant to the protagonist, while the neutral condition had a statement of the same word length with no indication of desire to succeed.

As with the spatial problems, each motivational problem also began with a couple of sentences. The first sentence introduced the scenario, (i.e., "Joe has just moved to a new school and it is his birthday this weekend."). The next sentence included either a motivational statement (i.e., "Joe is going to invite everyone from his new class**, but he also wants to invite friends from his old school**."), or a neutral statement (i.e., "Joe is going to invite everyone from his new class, **and he will be passing out all the invitations by hand**.").

Because the spatial questions were all multiplication problems, the motivation questions were be designed to include the other three operations (i.e., addition, subtraction, and division). This would prevent students from mechanistically thinking that all of these questions should be solved by multiplication (see Appendix C for sample motivational items).

Procedure

Data collection was done in a group setting and was divided into two separate phases. At both visits, instructions were read aloud to students. During the first phase, students completed the Raven's Matrices task (Raven et al., 2000) and the general math word problems assessment. The experimenter conducted two practice problems from the Raven's Matrices task with the class, students then had 25 minutes to complete the task. Upon the second visit students completed the spatial and motivational word problem booklets. Students were instructed to provide an equation as well as a numerical solution on provided media; there were three designated spaces, one for showing their work, another for the numerical equation, and a final separate space for the numerical answer. Students were told to answer all problems to the best of their abilities, that they were not required to show their work, but that it may be useful for them to do so. Students were told they could draw pictures, use tally bars, or any other type of strategy that helped them to solve the problem. Once the instructions were understood and pending questions answered, students began the task, if students had questions during the task they were able to ask for help. If students required help, they were first asked what specifically they needed help with. If the student did not understand the problem they were asked to re-read the text and focus on what the question was asking, then take the information from the text that they felt was necessary to answer the problem. If students required help with reading the text they were given assistance, however, the experimenter did not read the problem to the student, the experimenter helped with sounding out words. At no point was the student given any indication as to whether their answer was correct or incorrect. When an answer was achieved, the student was told they had done a good job and to move onto the next problem. In some instances students were given examples of much easier problems, the experimenter then asked the student how they got to the answer, then told them to apply the same strategy to the current problem. This procedure was adopted to simulate normal classroom support, without giving them the answer.

Both the six spatial problems and the six motivation problems were divided into two groups of three problems, with each group designed to be equivalent in difficulty. The two groups of spatial problems were called S1 and S2, while the two groups of motivational problems were called M1 and M2. Furthermore, each of the problem sets existed in two versions depending on which version of the problems were used. For example, S1-A indicates the S1 group of problems in their spatially associated versions, and S1-D is the same problems in their spatially disassociated versions. Likewise, M2-M indicated the motivated versions of the M2 problem set, while M2-N were the same problems in their neutral version.

Each participant completed 12 problems, but the composition of these problems varied. There were four sets of three problems (i.e., S1, S2, M1 and M2) and each of these problem sets could be one of two-dimensional orientations (e.g., S1 could be either S1-A or S1-D; M1 could be M1-M or M1-N). Given these parameters, there are four possible combinations designed so that students complete one problem set from each of the dimensional possibilities. For the spatial questions, one problem set of associated problems and one problem set of disassociated problems were included. The same process occurred for the motivational problems, one problem set of motivational problems and one problem set of neutral problems. The four different combinations are described in Table 1. For each of these combinations, the order of the questions was also varied. Using a Latin Square, the problem sets with each combination were put into four different orders. Within each problem set, there were six different orders, again generated using a Latin Square. The result of these two Latin Squares was a total of 24 different orders for each of the combinations listed in Table 1. Therefore, in total, there were 96 different possible question booklets.

Table 1.

Problem Set Combinations.

	Spatial		Motivational	
Combinations				
1	S1-A	S2-D	M1-M	M2-N
2	S1-D	S2-A	M1-M	M2-N
3	S1-A	S2-D	M1-N	M2-N
4	S1-D	S2-A	M1-N	M2-M

Note. Spatial versions; (A) – Associated, (D) – Dissociated.
Motivational versions; (M) – Motivational, (N) – Neutral.

An unanticipated problem was encountered with the differing versions of spatial and motivational word problem booklets. Although 96 booklet versions were created, not all booklet versions were used. In the current study all varying combinations from the first two problem set combinations were used (see Table 1), but not those from the third and fourth problem set combinations. Consequently the motivational questions were not seen in all of their possible formats; specifically, the first motivational question set was never encountered in its neutral format and the second motivational question set was not seen in its motivational format. In the current study only one version of motivational and neutral questions were used per problem set. In light of this procedural error, caution must be taken when drawing conclusions regarding the motivational word problem data. There is no way to determine if the differences found in the current study reflect the manipulated portions of the problem, or if they are due to differing difficulty levels of the two problem sets.

Coding

Three pieces of information from the word problem booklets were coded: answer, number sentence/equation, and evidence of using a pattern in the area for showing their work. Answers were coded as either correct or incorrect. Number sentence/equation refers to the same area on the problem sheet, however the terminology varied depending on the grade level; number sentence is the terminology used in grade 3 and equation is the terminology used in grade 5. Patterning was another means that children could use to solve the problem at hand. For example, a student may draw 4 dogs each having 3 spots; the student would then count the number of spots to solve the problem, however, they may not have been able to provide an equation for the problem. Both the number sentence/equation and the patterning information were coded as 0 – incorrect, 1 – improperly executed, and 2 – correct. The coding for "improperly executed" referred to any response that appeared to be correct, however the student used the wrong portion of the equation to answer the problem. For example, in the following word problem, 'Brianna's CD player uses 2 batteries. She has a pack of 8 batteries. How many times can Brianna change the batteries?' an example of an equation provided for this problem is 8-2-2-2-2=0. The equation presented is correct, however the answer was provided as 0, instead of 4, this was considered to be an improperly executed response. Answers were coded as 0 – incorrect and 2 – correct (see Table 2 for the frequency of response types for both spatial and motivational questions). The coding, improperly executed, was rare enough that it was not used in analyses and was instead counted as incorrect; only incorrect and correct coding was used in our analyses. The end result is that, for each problem, children had a separate score indicating if they had the correct response, used a valid equation to solve the problem, and/or used a valid pattern to solve the problem. Although all problems were coded as correct or incorrect (missing answers were considered incorrect), equation and pattern were only coded if there was an equation or a patterning attempt present. For this reason, it was possible on any given problem to have both an equation and a patterning code, one of these codes, or neither of them.

Results

Data were analyzed using ANOVAs to compare students' performance (for both answers, and formula/pattern use) on the different SM dimensions (i.e., comparing a student's performance on spatially associated word problems to their performance on spatially dissociated problems). Because each student completed problems of each dimensional type, the SM dimensions included in the analyses were within-subjects factors. Gender and grade were entered as between-subjects factors, and the spatial and motivation dimensions were each tested in separate analyses. Furthermore, interactions between ability and the spatial and motivational dimensions were tested in separate analyses by including the Raven's Matrices and general word problem test as covariates.

Table 2.

Frequency of response types for formula, pattern, and answers in spatial and motivational questions for grade 3 (n = 40) and 5 students (n = 44).

	Associated			Dissociated		
	Incorrect	Improperly Executed	Correct	Incorrect	Improperly Executed	Correct
Grade 3						
Formula	32	4	71	37	1	75
Pattern	7	6	32	5	5	39
Answer	49		71	46		74
Grade 5						
Formula	10	0	105	5	1	112
Pattern	0	8	29	3	5	34
Answer	18		114	20		112
	Motivational			**Neutral**		
	Incorrect	Improperly Executed	Correct	Incorrect	Improperly Executed	Correct
Grade 3						
Formula	24	7	72	29	0	79
Pattern	4	1	17	4	1	27
Answer	42		78	43		77
Grade 5						
Formula	12	1	92	7	3	100
Pattern	0	5	17	2	2	25
Answer	27		105	34		98

The aim of these additional analyses was to determine if having supplementary situation material (i.e., spatially associating an object to the protagonist, or providing motivating information to solve the problem) affected students differently based on their general problem solving abilities.

Considering the spatial dimension first, a 2 (Grade) x 2 (Gender) x 2 (Spatial Dimension) between-within ANOVA found a main effect for Grade (Grade 5s outperformed Grade 3s, see Figure 1), but also found a three-way interaction $F\ (1, 80) = 7.871$, $p = .006$, $\eta^2 = .090$. Upon further inspection, by Grade, it was determined that the three-way interaction was attributed to the Grade 5 students (Figure 2). Although the mean differences between the spatially associated and spatially dissociated questions ($M = 2.650$ and $M = 2.250$, respectively) for the boys were fairly large, a test of simple main effects determined them to be just outside the range of significance $t(19) = 1.798$, $p = .08$. Breaking down this interaction in another way – comparing the Grade 5 boys and girls on the spatial questions – it was found that girls were performing better than the boys on the spatially dissociated questions, but not the spatially associated questions, $F\ (1, 42) = 5.67$, $p = .02$.

When the Raven's Matrices total score was used in the analysis as a covariate, the three-way interaction remained, $F\ (1, 78) = 6.527$, $p = .013$. A similar three-way interaction was found when using the general word problem total score as a covariate, $F\ (1, 70) = 6.588$, $p = .012$. The interpretation of both of these ANCOVAs paralleled that of the main analysis – Grade 5 girls outperformed boys on the spatially dissociated questions, but no other differences are significant.

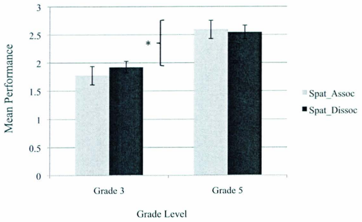

Figure 1. Mean Performance on spatial questions by grade. Error bars represent the standard error, * $p < .001$.

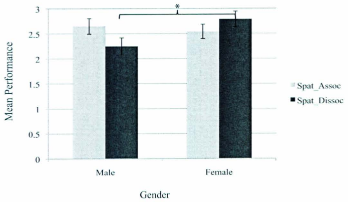

Figure 2. Mean performance on spatial questions by grade 5 students. Error bars represent standard error, * $p < .05$.

To focus on the motivational word problems, a 2 (Grade) x 2 (Gender) x 2 (Motivational Dimension) repeated measures ANOVA found no effect in performance on the motivational questions when compared to their neutral counterparts for both Grade 3 and 5 students, F (1, 82) = 1.598, p = .21, η^2 = .019 (see Figure 3). There was, however, a Gender x Motivational Dimension interaction F (1, 80) = 4.342, p = .04, η^2 = .051 (Figure 4). A test of simple main effects revealed the boys did worse on the neutral problems (M = 2.2195) in comparison to the motivational problems (M = 1.9268), t(40) = 2.395, p = .021, while girls did not show such a decrease in performance.

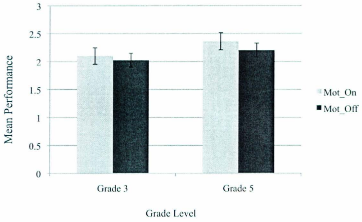

Figure 3. Mean performance on motivational questions by grade. Error bars represent standard error.

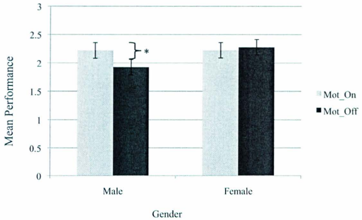

Figure 4. Mean performance on motivational questions by gender. Error bars represent standard error, * $p < .05$.

Using the Raven's Matrices total score as a covariate resulted in a comparable Gender x Motivational Dimension interaction $F\,(1, 78) = 4.692, p = .033$. Similarly, including the general word problem measure total score as a covariate also produced a Gender x Motivational Dimension interaction $F\,(1, 79) = 4.418, p = .039$. Both of these ANCOVAs resulted in the same pattern of results as the first analysis.

The previous analyses examined children's ability to answer word problems correctly. These problems, however, were not only coded as correct or incorrect, but they were also coded on whether or not a correct formula was used and/or whether a correct pattern was used to solve the problem. With this information, it is possible to consider if the motivational or spatial dimension had any effects on the use of formula or patterns, regardless of whether the problem was answered correctly. To include as many people as possible in the analysis, the data concerning formula use and pattern use were combined. Every problem was coded as either using a correct formula and/or pattern or not doing so. Used as the dependent variable, analyses that paralleled those detailed above were conducted to test if the spatial or motivational dimension affected the correct use of formulas of patterns.

In regards to the spatial dimension and the use of patterning and formulas, a 2 (Grade) x 2 (Gender) x 2 (Spatial Dimension) repeated measures ANOVA was used. This analysis did not find any effects, nor were there any effects when the Raven's Matrices were included as a covariate. When the general word problem measure was used as a covariate, however, there were significant differences in performance between the grades $F(1, 64) = 5.315$, $p = .024$ (see Figure 5). When controlling for general word problem ability, Grade 5 children performed better than Grade 3 children, indicating that using patterning and formulas for the spatial word problems were differentially affected by prior ability to solve general word problems. Grade 5 students may have been performing better as a result of having more exposure to solving textbook style word problems in comparison to the Grade 3 students. This experience may have led them to use equations and patterning more when trying to solve word problems.

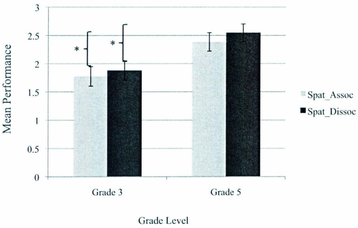

Figure 5. Mean performance on spatial question formula/pattern use by grade when the general word problem measure is used as a covariate. Error bars represent standard error, * $p < .05$.

Using the same analysis as above for the motivational question formula/pattern use, there were no effects for the first analysis, and no effects when the Raven's Matrices was included as a covariate. When the general word problem measure was included as a covariate, there was no effect for the motivation factor, but a Gender x Grade between subjects interaction was found, $F(1, 49) = 4.697$, $p = .035$. Grade 3 boys used more correct formulas or patterns than Grade 3 girls on the motivational questions ($M = 2.921$ and $M = 1.884$, respectively), but only when controlling for general word problem ability.

In motivational dimension, the word problems consisted of three types of operations: addition, subtraction, and multiplication. Although this was not a goal of this study, an exploratory analysis was conducted to see if the effects of the motivation dimension varied across operation. A 2 (Type of Motivation) x 3 (Operation) repeated measures ANOVA with Grade as a between subjects factor revealed a three-way interaction, $F(2, 160) = 6.162$, $p = .003$, $\eta^2 = .072$. Further analyses showed the interaction as being attributed to the Grade 5 students, while a test of simple main effects determined a significant difference in performance on the type of motivational questions in the addition and division problems [$t(41) = 4.109$, $p < .001$ and $t(41) = 3.814$, $p < .001$ respectively], but not in the subtraction problems. Grade 5 students' performance was better on the motivational addition problems and the neutral division problems. It should be noted, however, that each cell in this analysis only contained one question (e.g., there was only one question that was a motivation on addition question), so these results may be due to the vagaries of the particular questions involved rather than the operation.

The final question to be answered pertained to the level of interference experienced by students. This was measured by counting the number of wrong answers for the spatial questions that were answered with the interfering addition fact (e.g., when the solution to a problem was found by 3 x 2 = 6, how many times did children answer 3 + 2 = 5?). A total of 13 wrong answers had the interference response listed as the answer, and these errors were made by 8 students, all of whom were Grade 3's. The 13 interference responses were split between eight interference responses to the spatially associated questions and five interference responses to the spatially dissociated questions.

A binomial test revealed no significant difference in the probability of responding with the interfering numerical information to the spatially associated and dissociated word problems ($p = .291$).

Discussion

The data reported above do support the primary hypothesis. Minor word changes in math word problems designed to reflect the spatial and motivational dimensions of SMs did affect children's performance, although this change in performance depended on both grade level and gender. Differing performance based on grade level was seen for the spatial problems but not the motivational problems. Gender differences were seen in both spatial and motivational problems.

Altering the spatial association of an object to a protagonist in a math word problem did affect students' performance. The results, however, did not suggest that including spatially dissociating information enhanced students' overall performance in comparison to the spatially associated information, as would be expected based on Mattarella-Micke and Beilock's (2010) findings. Nonetheless, there were differences to be found. Grade 5 girls were performing better on the spatially dissociated questions in comparison to the Grade 5 boys. As Mattarella-Micke and Beilock did not find any differences in the way males and females performed in their associated and dissociated versions of the spatial word problems, conclusions on the differences found in this sample cannot be easily explained by previous findings. Interestingly, the Grade 5 boys appear to have performed better on the spatially associated questions in comparison to their spatially dissociated questions, which would be an effect in the opposite direction than predicted. This difference, however, was just outside of significance, so the prudent course would be to presume that Grade 5 boys did not differ in their performance between the spatially associated and dissociated questions.

If, however, we were to assume that there was a difference between Grade 5 boys on these questions, there is one possible explanation that could explain both this effect and the difference between the Grade 5 boys and girls on the spatially dissociated questions. Zwann and Radvansky (1998) have suggested that incomplete goals are maintained in short-term memory, and that once a goal has been complete there may be no need to keep that information stored. The spatially associated questions place an object with the protagonist, but this information is irrelevant to the problem and is not referenced again. For example, when "Debbie picked up 7 chew toys and went to the dog kennel" (see Appendix B for word problem), the reader never knows what happened to the chew toys that Debbie picked up. This incomplete action is arguably similar to an incomplete goal. In the dissociated version of the problem, however, the object is removed from the protagonist: "Debbie put away 7 chew toys and went to the dog kennel." In this version of the word problem, the reader would likely not expect hear more about the chew toys that Debbie has put away, because the action is thought to be complete. In other words, even though these problems were meant to manipulate the spatial dimensions, perhaps they inadvertently manipulated the motivational dimension instead by providing some problems with incomplete goals that were more motivating to solve (i.e., the spatially associated questions) and other problems with complete goals that were less motivating to solve (i.e., the spatially disassociated questions). As such, Grade 5 boys in this sample, may be trending towards a statistical advantage in the spatially associated versions because the protagonist is not doing anything with the associated object, ergo, an incomplete goal is described. The boys may have been giving the question more thought because of the information about the chew toy staying in their short-term memory, and this resulted in more correct answers in the associated versions of the spatial questions.

Although the motivational difference may explain this effect for the Grade 5 boys, it does not explain, by itself, why there was no effect for the Grade 5 girls or for any of the Grade 3s. One possibility, however, is that Grade 5 girls do not require the extra motivation factor to do well on a problem. If we speculate the motivational manipulation described does not provide any more motivation than the girls already feel for the problem (but that it does provide boys with more motivation), then this would also explain why the Grade 5 girls outperformed the Grade 5 boys on the disassociated problems. With the associated problems, boys have extra motivation, so they perform just as well as the girls. With the disassociated problems, however, that extra motivation is not present, which decreases the Grade 5 boys' performance, but not that of the Grade 5 girls. As for the Grade 3s, it may be that they are not yet old enough to be sensitive to the motivational information in these problems.

Turning to the questions that were actually designed to manipulate the motivational dimension, performance was affected differently by gender. In this case boys performed better on the motivational questions in comparison to the neutral questions. This result supports the second hypothesis, but only for the boys; that performance would increase on motivational questions. On further inspection, however, boys are not outperforming girls when they have motivational information. Looking back at Figure 4, boys and girls have a similar performance on the motivationally worded questions. The differences lie in the neutral word problems, as the boys were performing much worse on the neutral problems in comparison to the girls. It may be that the motivational information equally helped both the boys and girls; however, without it (e.g., in the neutral versions) the boys' performance was hindered, suggesting the boys most benefitted from the motivational information. The girls, as was speculated above for the spatial questions, may not have needed the extra motivational information as they performed equally well on both the motivational and neutral problems. It remains unclear as to why this difference is seen in the genders; perhaps the girls maintain a level of motivation for solving the problem that the boys lack when the motivational information is not salient. The motivational dimension questions were comprised of three types of arithmetic operations. Geary (1994) suggests that children's performance differs depending on arithmetic operations primarily because of the order they learn these operations in formal education, their exposure, and experience performing them. The results of performance on the different arithmetic operations used in current study were puzzling and did not comply with Geary's take on children's performance of differing arithmetic operations. As previously mentioned, there were procedural errors within the motivational word problem sets (see Materials), because of this, the difference in performance of the various arithmetic operations could reflect of the word problem itself (e.g., one set of the division problems were only seen in their motivational forms) and not the operation used to solve the problem.

The gender differences found in the current study are only slightly supported by the existing literature on gender differences in math. The Grade 5 girls performing better than the boys on the spatially dissociated version of the word problems is not what would be expected, given that boys in this age group tend to perform better than girls on word problems (Delgado & Prieto, 2005; Geary, 1996; Lummis & Stevenson, 1990; Marshall & Smith, 1987; Stevenson et al., 1990). The boys did have an increased performance in the motivational questions, but only in comparison to their neutral counterparts and not in comparison to the girls' performance. Further investigation of these gender differences are warranted; perhaps further enhancing the motivational features – placing the motivational statement prior to the problem text (e.g., Thevenot et al., 2007) – may result in girls performing better on the motivational problems in comparison to the neutral problems.

The Grade 5 students performed better than the Grade 3 students on both types of spatial questions (associated and dissociated). These grade differences are to be anticipated, based on both previous literature (Coquin-Viennot & Moreau, 2003, 2007; Davis-Dorsey et al., 1991) and the fact that Grade 5 students are further along in their education and would be expected to perform better than younger students. This difference would certainly be evident if the word problems were the exact same, but the word problems were not the exact same in the current study. Both the spatial and motivational word problems were different only in terms of the size of the numbers used. The numbers were designed to be challenging for each grade; however, creating numbers that were equivalent in difficulty level for both grades cannot be assured. In light of this, grade differences could be due to either problems being too difficult for the Grade 3 students, too easy for the Grade 5 students, or both.

It is, however, interesting that grade differences were only found for the spatial word problems and not for the motivational word problems. This suggests that the grade difference may not only be related to the size of the numbers used, but may be due to the nature of the spatial and motivational problems. The spatial questions were adapted from Mattarella-Micke and Beilock (2010); all of their spatial questions were designed to be multiplication problems with additional irrelevant numerical information. As Grade 5 students are more familiar with multiplication problems and are taught strategies for identifying extraneous information, they may have performed better on the spatial questions due to prior exposure during their educational experience. In the motivational questions, the lack of performance differences supports the prior argument regarding the nature of the problems. The motivational problems did not have any multiplication and consisted of only addition, subtraction, and division. There were more opportunities in the motivational problem sets for the Grade 3 students to answer correctly as a result of their comfort with those types of 'easier' mathematical operations. Davis-Dorsey et al. (1991) have attributed grade differences in their study to older students having better-developed schemata of textbook word problems. This may also explain grade differences found in the current study.

The current study used the Raven's Progressive Matrices task (Raven et al., 2000) (as a proxy for IQ) and a measure of general word problem ability as a means to examine differences in performance on the spatial and motivational word problems based on prior ability. There have been studies examining differences in performance on math word problems based on variables such as math skills (Thevenot et al., 2007), level of achievement (Vincente et al., 2008), and problem solving ability (Coquin-Viennot & Moreau, 2003; Hegarty et al., 1995) in both young children and young adults (undergraduate students). In all these cases, the authors found performance in math word problem tasks varied depending on the abovementioned factors and that children with lower math abilities performed better on SM problems than would be expected based on their general ability. Those students with higher levels of math skill, achievement, and problem solving ability also performed better on the word problems, but the improvement was greater for lower-ability students. In the current study it was hypothesized that students with lower abilities may have benefited from the additional information in the spatial and motivational problems. That was not the case; in this sample, level of performance on word problems targeting the spatial and motivational dimensions of SMs did not vary by general ability. Including motivational information did not help the students with lower general abilities any more than it helped those with higher general abilities, nor did providing spatial cues in the spatial dimension.

When used as a covariate, the general word problem measure did demonstrate interesting differences in correct formula and pattern use. Grade 5 students were using more correct formulas and/or patterns than the Grade 3 students on the spatial problems when general word problem ability was covaried out of the relation. As with the grade effect in correct answers on the spatial problems, this difference may be explained by the additional experience of Grade 5 students in working with multiplication. Similar results have been found in terms of formula formation among the grades; Coquin-Viennot and Moreau (2003) found that older students were more likely to use a factorizing strategy to solve the word problems with a structuring term. In the case of the motivation problems, Grade 3 boys also had more correct formula/pattern use than the Grade 3 girls. Although previous studies have found gender differences in word problems (Delgado & Prieto, 2005; Geary, 1996; Lummis & Stevenson, 1990; Marshall & Smith, 1987; Stevenson et al., 1990), the gender differences found in these studies do not address prior word problem ability and cannot account for how general ability may interact with word problem performance differences.

The final question this study aimed to address was if children would experience interference in spatial word problems when highly interfering numerical information was introduced on a pencil and paper task. Of the students who reported the numerical answer to the word problem as the interfering numerical information (which was a very small number), they did so more often on the spatially associated problems. There were a small number of students (all Grade 3) who actually displayed interference. According to the DOA model (see Mattralle-Micke & Beilock, 2010) the effectiveness of retrieval depends on the associative strength of the correct answer. In the case of the Grade 3 students the interfering numerical information is more likely to hold increased associative strength with the problem because they are more familiar with common addition equations (i.e., 7, for 3 x 4). The DOA model can also explain this lack of an interference response in the Grade 5 students. Grade 5 students have become more familiar with common multiplication equations and therefore would not hold the associative strength of the addend as the Grade 3 students would. Nevertheless, it should be kept in mind that there actually was very little interference evident in the responses of these Grade 3 and 5 children. It is just that the small amount of interference that was present was found exclusively within the sample of Grade 3 children.

There are limitations to the current study. One is a lack of measure for baseline reading ability of the students in the sample. This is problematic as story problems are harder to solve than their corresponding numerical format, suggesting that factors other than math skills play a role in successfully solving a mathematical word problem (Reusser, 1990). During data collection it was noted that some students were having difficulty, or were unable to read the problems; one student who was noted for not being able to read the problems was eliminated from analyses. Other students with similar reading problems may have been missed, so it is unknown as to whether some students' performance is a reflection of their math ability or their reading ability. Nevertheless, the effect of missing those students with reading difficulties should not take away from the current results regarding the spatial and motivational dimensions. These analyses were conducted using the SM dimensions as within-subjects factors, which means that poor readers would not bias the results in favour of one dimension or another. Furthermore, many of the differences in reading word problems were probably also controlled for by the inclusion of the general word problem measure. Given that this measure did not interact with either of the SM dimensions, we can have some confidence that the result reported here can be generalized across a range of reading levels. Nevertheless, future studies should still obtain some measure of students' reading comprehension, as it may identify an interaction of reading ability with these SM factors that is over and above the influence of general IQ and general word problem ability. It is important to note that mathematical English is not the same as ordinary English. Regular reading measures may not be appropriate for measuring the comprehension of mathematical English texts (Kane, 1978), as many of the commonly used reading measures do not always measure the same kinds of comprehension and vary with developmental level (Keenan, Betjemann, & Olson, 2008). Nevertheless, Vilenius-Tuohimaa and colleagues argue that technical reading skill level is related to math word problem solving and reading comprehension, and those students with poor decoding skills may have increased difficulty with the text itself and as a secondary result struggle with the solving of the math problem (Vilenius-Tuohimaa, Aunola, & Nurmi, 2008).

Conclusion

Performance on math word problems can be differentially affected by both grade and gender through enhancing varying dimensions of SMs. Future research should continue exploring the use of SM dimensions to add complexity to word problems. Educators could then create sequenced teaching agendas so that children are able to progress in solving increasingly complex word problems. Given the pre-existing and current research on how SM dimensions affect students' performance in math word problems, teachers may gain a better understanding of how and where children experience difficulty when solving math word problems. Teachers will be better able to look at the word problems in textbooks and determine if any particular problem is lacking helpful information or incorporating confusing extraneous information that may hinder a certain grade levels performance. With this increased understanding, teachers may feel better equipped to modify or create their own word problems to give to their students.

Furthermore, researchers in the field of mathematical cognition should continue focusing efforts on how performance of math word problems are affected by each SM dimension (e.g., positive or negative), and ways to enhance math word problems using various SM dimensions. Another point of interest would be to examine how performance is affected when individual SM dimensions are amalgamated, and would combining individual dimensions help or hinder performance. For example, in the current literature, the temporal dimension has been shown to have no affect on performance in math word problems, and in some cases using the temporal dimension can hinder performance; however, it is possible that the temporal dimension can affect performance in a positive manner if it is paired with another SM dimension.

Gender differences also deserve further investigation, especially given the current study's findings of girls performing better than boys in the spatial dimension (dissociated versions), which goes against the pre-existing findings of boys performing better at word problems than girls. Regarding the motivational dimension, forthcoming research should work to ensure a procedure that allows for all combinations of motivational word problems to be used as well as creating problems that are not too easy/hard for the grade level. Grade level teachers could be asked to review the problems and provide feedback based on reasonable difficulty of particular items. Finally, regarding the interference effect, subsequent examination of interference in children this age should explore Mattarella-Micke and Beilock's (2010) computer task, as measuring reaction time may serve as a better means for testing the interference effect.

References

Aiken, L. R. –Jr. (1972). Language factors in learning mathematics. *Review of Educational Research, 42*(3), 359-385.

Appel, R., Brown, T., Galvin, D., Gibeau, L., Jeroski, S., Morrow, P., Weight, W. … Wortzman, R. (2009). *Math Makes Sense 3*. Canada: Pearson.

Austin, J. L., & Howson, A. G. (1979). Language and mathematical education. *Education Studies in Mathematics, 10*, 161-197.

Ball, S., Martin Connell, M., Hantelmann, L. J., Jeroski, S., Morrow, P., Saundry, C., & Wood, M. (2008). *Math Makes Sense 2.* Canada: Pearson.

Ballew, H., & Cunningham, J. W. (1982). Diagnosing strengths and weaknesses of sixth-grade students in solving word problems. Journal for Research in Mathematics Education, 13, 3, 202-210.

Ben-Zeev, T., Duncan, S., & Frobes, C. (2005). Stereotypes and math performance. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp.235-249). NY, US: Psychology Press.

Brannon, L. (2011). *Gender: Psychological perspectives* (6th ed., pp. 157-161). Boston, MA: Pearson.

Brydon, C., Grill-Donovan, E., Hope, J., Klassen, W., Small, M., Stuart, S., & Tseng Tam, R. (2008). *Mathfocus 5*. Toronto, ON: Nelson.

Byrnes, J. P. (2005). Gender differences in math: Cognitive processes in an expanded framework. In A. M. Gallagher, & J. C. Kaufman (Eds.), *Gender differences in math: An integrative psychological approach* (pp. 73-98). NY, US: Cambridge University Press.

Canavan-McGrath, C., Hope, J., Small, M., Sterenberg, G., Stuart, S., & Tseng Tam, R. (2010). *Mathfocus 6*. Toronto, ON: Nelson.

Chase, C. I. (1960). The position of certain variables in the prediction of problem-solving in arithmetic. *Journal of Educational Research, 54*(1), 9-14.

Clements, D. H. (2010, April). *The building blocks or early mathematics: Learning trajectories for young children*. Paper presented at the biennial meeting of the Society for Research in Child Development, Montreal, Quebec.

Cohen Levine, S. (2010, April). *The role of early input in children's numerical and spatial development.* Paper presented at the biennial meeting of the Society for Research in Child Development, Montreal, Quebec.

Coquin-Veinnot, D., & Moreau, S. (2007). Arithmetic problems at school: When there is an apparent contradiction between the situation model and the problem model. *British Journal of Educational Psychology, 77*, 69-80. doi:10.1348/000709905X79121

Coquin-Viennot, D., & Moreau, S. (2003). Highlighting the role of the episodic situation model in the solving of arithmetical problems. *European Journal of Psychology of Education, 18*(3), 267-279.

Dark, V. J., & Benbow, C. P. (1990). Enhanced problem translation and short-term memory: Components of mathematical talent. *Journal of Educational Psychology, 82*(3), 420-429.

Davis-Dorsey, J., Ross, S. M., & Morrison, G. R. (1991). The role of rewording and context personalization in the solving of mathematical word problems. *Journal of Educational Psychology, 83*(1), 61-68.

Delgado, A. R., & Prieto, G. (2004). Cognitive mediators and sex-related differences in mathematics. *Intelligence, 32*, 25-32. doi:10.1016/S0160-2896(03)00061-8

Dunlap, W. P., & McKnight, M. B. (1978). Vocabulary translations for conceptualizing math word problems. *The Reading Teacher, 32*(2), 183-189.

Geary, D. C. (1996). Sexual selection and sex differences in mathematical abilities. *Behavioral and Brain Sciences, 19*, 229-284.

Geary, D. C. (1994). *Children's mathematical development: Research and practical applications*. Washington, DC: American Psychological Association.

Gernsbacher, M. A., Verner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning. Memory, and Cognition, 16*(3). 430-445.

Ginsberg, H. P. (2010, April). *Young children's mathematical competence.* Paper presented at the biennial meeting of the Society for Research in Child Development, Montreal, Quebec.

Hall, J. W., & Fuson, K. C. (1986). Presentation rates in experiments on mnemonics: A methodological note. *Journal of Educational Psychology, 78*, 3, 233-234.

Hegarty, M., Mayer, R. E., & Monk, C. A. (1995). Comprehension of arithmetic word problems: A comparison of successful and unsuccessful problem solvers. *Journal of Educational Psychology, 87*(1), 18-32.

Hope, J., Klassen, W., Small, M., Tam Tseng, R., & Tossell, S. (2008). *Mathfocus 4.* Toronto, ON: Nelson.

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, *107*, 139-155. doi:10.1037/0033-2909.107.2.139

Johnson, E. S. (1984). Sex differences in problem solving. *Journal of Educational Psychology, 76*(6), 1359-1371.

Kaizer, C., & Shore, B. M. (1995). Strategy flexibility in more and less competent students on mathematical word problems. *Creativity Research Journal, 8*(1), 77-82.

Kane, R. B. (1968). The readability of mathematical English. *Journal of Research in Science Teaching, 5*, 296-298.

Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. Scientific Studies of Reading, 12(3), 281-400. doi:10.1080/1088430802132279

Lummis, M., & Stevenson, H. W. (1990). Gender differences in beliefs and achievement: A cross-cultural study. *Developmental Psychology, 26*(2), 254-263.

Marshall, S. P., & Smith, J. D. (1987). Sex differences in learning mathematics: A longitudinal study with item and error analyses. *Journal of Educational Psychology, 79*(4), 372-383.

Martin, S. A., & Bassok, M. (2005). Effects of semantic cues on mathematical modeling: Evidence from word-problem solving and equation construction tasks. *Memory & Cognition, 33*(3), 471-478.

Mattarella-Micke, A., & Beilock, S. L. (2010). Situating math word problems: The story matters. *Psychonomic Bulletin & Review, 17*(1), 106-111. doi: 10.3758/PBR.17.1.106

Mayer, R. E. (1982). Different problem-solving strategies for algebra word and equation problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*(5), 448-462.

Moreau, S., & Coquin-Viennot, D. (2003). Comprehension of arithmetic word problems by fifth-grade pupils: Representations and selection of information. *British Journal of Educational Psychology, 73*, 109-121.

Muth, K. D., & Glynn, S. M. (1985). Integrating reading and computational skills: The key to solving arithmetic word problems. *Journal of Instructional Psychology, 12*(1), 34-38.

O'Mara, D. A. (1981). The process of reading mathematics. *Journal of Reading, 25*(1), 22-30.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales: Section 3: Standard Progressive Matrices.* San Antonio, TX: Harcourt Assessment.

Reusser, K. (1990). From text to situation to equation: Cognitive simulation of understanding and solving mathematical word problems. In H. Mandl, E. DeCorte, N. Bennett, & F. Helmut Felix (Eds.), *Learning and Instruction: European Research in an International Context Vol. 2.1* (pp. 477-498). Elmsford, NY, US: Pergamon Press.

Rinck, M., & Bower, G. H. (2000). Temporal and spatial distance in situation models. *Memory & Cognition, 28*(8), 1310-1320.

Rothman, R. W., & Cohen, J. (1989). The language of math needs to be taught. *Academic Therapy, 25*(2), 133-142.

Son, J. Y., Thai, K., Burke, T., & Kellman, P. (2010, April). *Perceiving structure in word Problems: Applying perceptual learning to elementary math pedagogy.* Paper presented at the biennial meeting of the Society for Research in Chlid Development, Montreal, Quebec.

Stern, E., & Lehrndorfer, A. (1992). The role of situational context in solving word problems. *Cognitive Development, 7*, 259-268.

Stevenson, H. W., Lee, S., Chen, C., Lummis, M., Stigler, J., Fan, L., & Ge, F. (1990). Mathematics achievement of children in China and the United States. *Child Development, 61*, 1053-1066.

Thevenot, C. (2010). Arithmetic word problem solving: Evidence for the construction of a mental model. *Acta Psychologica, 133*, 90-95. doi:10.1016/j.actpsy.2009.10.004

Thevenot, C., Devidal, M., Barrouillet, P., & Fayol, M. (2007). Why does placing the question before an arithmetic word problem improve performance? A situation model account. *The Quarterly Journal of Experimental Psychology, 60*(1), 43-56. doi:10.1080/17470210600587927

Thevenot, C., & Oakhill, J. (2006). Representations and strategies for solving dynamic and static arithmetic word problems: The role of working memory capacities. *European Journal of Cognitive Psychology, 18*(5), 756-775. doi:10.1080/09541440500412270

Thevenot, C., & Oakhill, J. (2005). The strategic use of alternative representations in arithmetic word problem solving. *The Quarterly Journal of Experimental Psychology, 58A*(7), 1311-1323. doi:10.1080/02724980443000593

U.S. Department of Education. (2008). *Foundations for success: The final report of the national mathematics advisory panel* (ED04CO0082/0001).

Verschaffel, L., & DeCorte, D. (1997). Word problems: A vehicle for promoting authentic mathematical understanding and problem solving in the primary school? In T. Nunes, & P. Bryant (Eds.), *Learning and teaching mathematics* (pp. 69-97). East Sussex, UK: Psychology Press Ltd.

Vilenius-Tuohimaa, P. M., Aunola, K., & Nurmi, J-E. (2008). The association between mathematical word problems and reading comprehension. *Educational Psychology, 28*(4), 409-426. doi:10.1080/01443410701708228

Vincente, S., Orrantia, J., & Verschaffell, L. (2008). Influence of situational and mathematical information on situationally difficult word problems. *Studia Psycholigica, 50*(4), 337-356.

Zwann, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*(2), 162-185.

Appendices

Appendix A

Sample of General Math Word Problems

Grade	**Sample Problem**
Grade 3	The library had 12 books about the moon. Penny borrowed some of them. There are 4 books left. How many books did Penny borrow?
Grade 5	Grace walks 9 km each day. There are 28 days in February. How many kilometers does Grace walk in February?

Appendix B

Sample of Spatial Word Problems

	Sample Problem	
Grade	Associated	Dissociated
Grade 3	Debbie owned a kennel and was tending to the puppies in her care. She **picked up** 7 chew toys and went to the dog kennel. Each puppy has 3 spots. If there are 4 puppies, how many spots do the puppies have altogether?	Debbie owned a kennel and was tending to the puppies in her care. She **put away** 7 chew toys and went to the dog kennel. Each puppy has 3 spots. If there are 4 puppies, how many spots do the puppies have altogether?
Grade 5	Earl was working hard in his room, already late for school. Earl **picked up** 15 notebooks and left for the library. At the library there were 7 tables with 8 empty chairs at each table. How many seats is Earl able to choose from?	Earl was working hard in his room, already late for school. Earl **put down** 15 notebooks and left for the library. At the library there were 7 tables with 8 empty chairs at each table. How many seats is Earl able to choose from?

Appendix C

Sample of Motivational Word Problems

	Sample Problem	
Grade	Motivational	Neutral
Grade 3	Amy just started baby-sitting for her neighbour. She is saving money to buy a new bike **because she wants to join the bicycle club**. If the bike is $60 and Amy makes $10 every time she baby-sits. How many times will Amy have to baby-sit before she can buy the bike?	Amy just started baby-sitting for her neighbour. She is saving money to buy a new bike, **but she can borrow her older sister's bike**. If the bike is $60 and Amy makes $10 every time she baby-sits. How many times will Amy have to baby-sit before she can buy the bike?
Grade 5	Ian and his mother are flying to Toronto for his grandmother's birthday. They have a birthday gift for grandma that weighs 11.5 pounds, and the airline only allows 55-pound bags. Ian wants to bring as much of his toys as possible, **but he needs to bring his grandmother's gift**. If he packs his grandmother's gift, how many pounds of his toys can he bring?	Ian and his mother are flying to Toronto for his grandmother's birthday. They have a birthday gift for grandma that weighs 11.5 pounds, and the airline only allows 55-pound bags. Ian wants to bring as much of his toys as possible, **but he can send the gift through the mail**. If he packs his grandmother's gift, how many pounds of his toys can he bring?