THE ROLE OF WORKING MEMORY IN DEDUCTIVE REASONING: A DUAL TASK AND INDIVIDUAL DIFFERENCES APPROACH

by

Stephanie Solcz

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Arts

in

Psychology

Waterloo, Ontario, Canada, 2008

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The ‘belief bias’ effect is one of the most pervasive findings in the study of syllogistic reasoning. Here, participants respond “valid” to more believable than unbelievable conclusions, regardless of the actual validity of the conclusion. There is also an interaction characteristic of the belief bias effect, in that conclusion believability plays a greater role when conclusions are invalid than when they are valid. The experiments reported in this thesis had two goals: first, to determine how individual differences in working memory (WM) capacity influence belief bias in reasoning; and second, to indentify which WM systems are involved in syllogistic deductive reasoning. To this end, both experiments employed a dual task paradigm.

In Experiment 1, participants remembered spatial arrays whilst reasoning through syllogisms in order to load the visuospatial sketchpad. Results demonstrated that performance on the secondary spatial memory task suffered when participants reasoned through syllogisms of which the validity and believability of conclusions were incongruent (i.e., “conflict” problems), indicating that reasoning through conflict problems utilized limited visuospatial WM resources. Also, only participants with high WM capacities showed the typical belief-bias effect, with greater effects of conclusion believability on invalid than on valid conclusions. This interaction was not present for low WM span participants, because they made greater errors on problems with invalid, unbelievable conclusions.

In Experiment 2, participants remembered digit sequences whilst reasoning in order to load the phonological loop. Both of the major results from Experiment 1 were replicated. Accuracy on the secondary digit recall task was impaired when participants reasoned through conflict problems, demonstrating that limited verbal WM resources were directed toward reasoning. Again, only high WM span participant showed the interaction between conclusion validity and believability characteristic of the belief bias effect. Effects were additive for low WM span participants because they made more errors on invalid, unbelievable syllogisms.

Results from both experiments demonstrate first, that both visuospatial and verbal WM resources are involved in syllogistic reasoning, and second, that individuals with different amounts of available WM resources demonstrate differential belief bias. These results are discussed in terms of the mental models and mental logic theories of reasoning and in terms of dual process accounts of reasoning.

Acknowledgements

I would like to thank my advisor, Jonathan Fugelsang. Nothing I could write would capture how appreciative I am of his endless academic and emotional support.

Thank you also to the research assistants who volunteered their time to help collect data.

Finally, thank you to my family and friends, who never failed to provide much needed support and encouragement.

Table of Contents

List of Tables

List of Illustrations

Introduction

It is a well-accepted finding that humans are far from rational when reasoning and problem-solving (e.g., Cohen, 1981). Rather than reasoning in a purely logical or normative fashion, people's decision-making is often influenced by past experiences and knowledge. A prime example of this is the belief-bias effect, first systematically examined by Evans, Barston, and Pollard (1983) and replicated in numerous studies (see Klauer, Musch, & Naumer, 2000 for a review). Evans et al. discovered that when people reason through syllogisms in studies of deductive reasoning, in which they must determine whether a conclusion logically follows from a set of premises, they are often influenced by the believability of the conclusions. Specifically, people are more likely to endorse the conclusion as valid when it is believable than when it is unbelievable. For example, people often respond that the conclusion, "Therefore, some cigarettes are not addictive" is invalid because it is unbelievable, regardless of whether it logically follows from a set of premises. Further, there is an interaction between the effects of conclusion validity and believability characteristic of the belief bias effect, in that conclusion believability plays a greater role for invalid than for valid conclusions.

Biases in reasoning may vary among people with differing cognitive abilities. For example, those with greater cognitive abilities may be more likely to reason in accordance with logic and less prone to bias when reasoning. This thesis provides a direct examination of the role of one such cognitive ability, working memory (WM) capacity, in belief bias and examines how individual differences in WM capacity influence people's susceptibility to bias in syllogistic reasoning. For the purposes of this thesis, WM is conceptualized in terms of Baddeley and Hitch's (1974) model of working memory, which proposes that working memory is composed of three separate, yet interdependent components, the phonological loop, visuospatial sketchpad, and central executive.

The Role of WM in Deductive Reasoning

Theories of deductive reasoning are hotly debated, with each specifying a significant role of WM. The mental models account of deductive reasoning, pioneered by Johnson-Laird (e.g., Johnson-Laird & Bara, 1984), proposes that when confronted with a syllogism, people construct models of information contained in the premises. Mental models can be thought of as iconic, diagram-like spatial representations of information contained in premises (Johnson-Laird, 2001; Knauff, Mulack, Kassubek, Salih, & Greenlee, 2002). For simple, "one model" syllogisms (e.g., "All poodles are dogs; All dogs are mammals; Therefore, all poodles are mammals"), only one model is required to sufficiently represent the premises; whereas for other, "multiple model" syllogisms (e.g., "Some poodles are dogs; No mammals are dogs; Therefore, some mammals are not poodles"), reasoners must create several models and counter models (which represent information contrary to premises) to represent the information contained in the premises. Reasoners then test the problem's conclusion against their mental models. If conclusions are consistent with the model and if no counter models can be found, then it is accepted as valid, and if it is inconsistent with the model or is consistent with a counter model, it is rejected as invalid (Chater & Oaksford, 2001).

A key feature of the mental models account is the search for counter models – without a thorough search for falsifying models, reasoning errors are likely (Newstead, Thompson, & Handley, 2002). For example, an invalid conclusion may be accepted as valid when the search for falsifying models is prematurely halted, and evidence from numerous studies suggests that reasoners often accept or reject a conclusion based on a perfunctory search of alternative models (e.g., Evans, Handley, Harper, & Johnson-Laird, 1999). Further, several studies have found that reasoning accuracy is much higher for syllogisms with only one possible model than for syllogisms with several models that must be considered, presumably because people fail to consider all possible models (Chater & Oaksford, 2001). Naturally, considering several possible models simultaneously requires sufficient WM capacity (Evans et al., 1999). Thus, according to the mental models account, individual differences in people's reasoning abilities may result because limitations in WM capacities mean that some people are unable to conceptualize all possible models. For the mental models account, given that mental models are spatial representations of relationships contained in premises, spatial WM (i.e., the visuospatial sketchpad) is imperative in deductive reasoning.

A competing class of theories, knows as mental logic accounts (e.g., Braine & O'Brien, 1998, Rips, 1994), explain deductive reasoning, particularly conditional reasoning, by proposing that people reason through deductive reasoning problems by utilizing rules of formal logic. Such rules allow reasoners to make inferences about relations among entities within premises. Problem difficulty is dictated by the number of rules or processing steps required to verify or falsify the conclusion. Similar to the mental models account, reasoning errors arise when people fail to use correct formal logic rules due to limitations in processing, or WM, capacity (Schroyens, Schaeken, & d'Ydewalle, 2001). However, contrary to the mental models account, which propose that premises are represented as spatial models, mental logic accounts assume that premises and formal logic rules are represented in a purely verbal, language-like fashion. Hence, mental logic accounts assign an important role to verbal WM (i.e., the phonological loop).

Studies confirming the role of WM in syllogistic deductive reasoning are abundant, although the results of these studies are somewhat inconsistent. Several researchers found correlations between WM capacity and reasoning accuracy. Copeland and Radvansky (2004), for example, found a positive correlation between general WM span and accuracy on a syllogistic reasoning task. Similarly, Kyllonen and Christal (1990) found strong positive correlations between several measures of WM capacity and several reasoning measures, including syllogistic reasoning. Further, Capon, Handley, and Dennis (2003) report positive correlations between performance on both spatial WM and verbal WM measures and a syllogistic reasoning measure.

Other studies have employed dual task paradigms, in which participants must complete a secondary task designed to deplete WM resources while simultaneously reasoning through syllogisms. For example, Gilhooly, Logie, Wetherick, and Wynn (1993) loaded the phonological loop through sequential number verbalization, the visuospatial sketchpad through clockwise finger-tapping, and the central executive through random number generation. They found that loading the phonological loop and central executive impaired accuracy on syllogistic reasoning, although loading the visuospatial sketchpad had no effect on reasoning accuracy. In a later study (Gilhooly, Logie, & Wynn, 2002), the researchers found that concurrent reasoning impaired finger-tapping only when premises were presented sequentially, which presumably places a greater load on WM than when premises are presented concurrently. Thus, the researchers concluded that the phonological loop and central executive play major roles in syllogistic reasoning, whereas the visuospatial sketchpad plays a more minor role. However, another possibility is that finger tapping did not load its targeted WM component to the same extent that number verbalization or random number generation did, and thus its effects were not seen on reasoning performance when premises were presented simultaneously. Also, importantly, these studies only looked at the role of WM in syllogistic reasoning with abstract, non-belief laden content.

The Role of WM in Belief Bias

Fewer studies have been conducted to illuminate the involvement of WM in biased responding in deductive reasoning, such as in the belief bias effect, although the popular dual process theory of information processing can lend some useful predictions. Dual process accounts of reasoning (Evans & Over, 1996; Stanovich & West, 2000; Sloman, 1996) assert that there are two systems of reasoning, referred to here as the heuristic and analytic system. The heuristic system is conceptualized to be automatic and efficient, can operate without awareness, draws on prior beliefs and experiences, and is not limited by strained cognitive resources. Conversely, the analytic system is thought to be conscious and effortful, to be free from bias, and to demand sufficient WM resources (Osman, 2004). For many decisions, the heuristic system, drawing on previous experiences and beliefs, provides a quick response. However, for situations requiring careful analysis, such as in syllogistic reasoning, the analytic system must draw on available WM resources to carefully reason through the problem. In these situations, the automatic heuristic system must be suppressed, which also requires WM resources. When sufficient WM resources are unavailable, the analytic system fails to suppress the heuristic system, and a response based on unconscious, biased processing results. In the case of syllogistic reasoning, it is assumed that the heuristic system initially provides a response based on the believability of the conclusion - if it is believable, it is accepted as valid, whereas it if it is unbelievable, it is rejected as invalid. Subsequently, the analytic system logically reasons through the problem and overrides the heuristic system if necessary. However, when WM resources are strained, possibly due to a limited WM span, divided attention, or to problem difficulty, strategic analytical processing fails. The reasoner then decides upon the response suggested by the heuristic system and provides a response based on the believability of the conclusion.

The use of the heuristic versus the analytic systems is evident on problems in which the validity and believability of the conclusions conflict – that is, for valid, unbelievable conclusions and for invalid, believable conclusions. For these “conflict” problems, analytic and heuristic processes elicit different responses (Evans, 2007). For instance, consider the following syllogism: No healthy people are athletic; Some astronauts are athletic; Therefore, some healthy people are not athletic. The conclusion is invalid (i.e., does not logically follow from the premises), yet believable. In this case, because the conclusion is consistent with what the reasoner knows about the world, the heuristic system automatically indicates that the conclusion is valid. Subsequently, if the analytic system successfully reasons through the problem, it will correctly identify the conclusion to be invalid.

Clearly, dual processing accounts of reasoning assign a primary role to WM capacity in successful reasoning. Given that sufficient WM resources are necessary to suppress the biased responses provided by the heuristic system, these theories predict that individuals with a higher WM span should be *less* susceptible to belief bias. Previous research supports this prediction: Quayle and Ball (2000), for example, found that participants with high spatial WM spans showed a decreased belief bias effect (i.e., were more accurate) relative to participants with low spatial WM spans. Similarly, De Neys (2006) found that participants with high WM span were more accurate when responding to problems in which the validity and believability of the conclusions conflicted than were participants with low WM spans. Thus, research supports a pivotal role of WM in deductive reasoning, and indicates that when insufficient WM resources exist, belief bias results.

Further, there is evidence to suggest that those with high and low WM capacities use different strategies when reasoning through syllogisms, with individuals with more WM resources using more complex strategies. For example, Copeland and Radvansky (2004) report that individuals with large WM spans responded to syllogisms in ways that were consistent with complex reasoning strategies, such as considering several mental models as predicted by the mental models account, whereas responses made by individuals with small WM spans were consistent with simple, heuristic based strategies. If people with larger WM spans use more complex strategies, a concurrent WM task should yield greater performance decrements for them than for people with smaller WM spans, whose heuristic-based strategies do not draw heavily on WM resources. Although this phenomenon has not been reported in deductive reasoning, it has been found in other domains: High WM participants were more impaired by a concurrent load task during a memory retrieval task than were low WM participants (Rosen & Engle, 1997). Further, mathematical problem-solving accuracy decreases more for high WM capacity individuals than for low WM capacity individuals under conditions of high pressure, which are presumed to consume WM resources (Beilock & Carr, 2005; Beilock & DeCaro, 2007). Thus, taxing cognitive resources should have larger effects on reasoners with high WM spans, presuming that they are more inclined to use complex reasoning strategies that are adaptive under normal conditions. When their strategies, which are heavily dependent on limited WM resources, are disrupted, they may be more susceptible to heuristic-based responses than are individuals with low WM spans, thus demonstrating a heightened belief bias effect under conditions of WM load.

The Current Research

Unlike past studies examining the role of WM in syllogistic reasoning, which used correlational or between subjects approaches, the current studies used a dual task approach to manipulate the amount of WM resources available during syllogistic reasoning to directly investigate the role of WM in belief bias. This approach will not only address inconsistencies in the results of studies examining the role of WM in deductive reasoning, it will also illuminate which components of WM are crucial for unbiased analytic reasoning. The current studies will also compare the belief bias effect in high WM and low WM participants as a function of a high and low WM load, to determine whether high and low WM span participants are differentially impaired by a load task due to different strategy use.

Additionally, because participants may choose to assign limited cognitive resources to the primary reasoning task, this thesis takes a somewhat novel approach by emphasizing impairments on the secondary load task along with the primary reasoning task. Reduced WM resources may cause decrements on the reasoning task: Specifically, limited WM resources may result in failure of the analytic system, causing the heuristic system to provide a response consistent with the believability of the conclusion. Conversely, effects of depleted WM resources may manifest on the load task: Accuracy may decrease when WM resources are required to suppress the heuristic system when the heuristic and analytic system provide competing responses (that is, when the validity and believability of the conclusion conflict). Either finding would indicate that the primary reasoning task and the secondary load task compete for the same pool of limited WM resources.

Experiment 1

The purpose of Experiment 1 was to determine whether concurrent syllogistic reasoning and a visuospatial memory task would cause interference in either task, thus supporting the assertion of mental models accounts that visuospatial WM is pivotal in deductive reasoning. Also, Experiment 1 directly compares belief-bias effects (both main effects of conclusion believability and validity, and the believability x validity interaction) across individuals of different WM capacities to directly examine whether reasoning processes vary according to individuals' WM spans. It was predicted that high WM span participants would be more impaired when under high WM load than would low WM span participants, thereby relying more heavily on reasoning according to beliefs than when not under load. This is in line with the somewhat counterintuitive finding indicating that individuals with low WM spans are less impaired than those with high WM spans when WM resources are taxed, possibly because they already use simple, heuristic-based strategies that do not draw heavily on WM resources (Rosen & Engle, 1997).

Method

Participants

Sixty undergraduate students in psychology courses at the University of Waterloo participated in the experiment in exchange for course credit.

Materials

Reasoning task. Participants were required to evaluate the logical validity of 16 categorical syllogisms, eight of which were valid and eight of which were invalid. All syllogisms used in the current experiment took the following forms:

No A are B; Some C are B; Therefore, some C are not A (Valid)

No A are B; Some C are B; Therefore, some A are not C (Invalid)

The syllogisms used in this thesis were adapted from Evans et al. (1983) and from Klauer et al., 2000 and appear in Appendix A. The syllogisms were constructed such that the conclusions of half were believable and the conclusions of the other half were unbelievable. In the stimuli set, there were four syllogisms in each of the validity/believability cells [i.e., valid/believable (VB), valid/unbelievable (VU), invalid/believable (IB), and invalid/unbelievable(IU)]. To counterbalance stimuli, after 30 participants were tested the validity of each syllogism was changed, while retaining the content and conclusion believability, by altering the order of items within the premises. For example, the valid syllogism, “No cigarettes are inexpensive; Some addictive things are inexpensive, Therefore, some addictive things are not cigarettes,” becomes invalid when the premises are altered: “No addictive things are inexpensive; Some cigarettes are inexpensive; Therefore, some addictive things are not cigarettes.”

Memory Task. A visuospatial memory task served as the secondary task in this experiment. Patterns of shapes appeared on the screen prior to each syllogism. Participants were instructed to remember the pattern for a recall test following each syllogism. In Low load trials, a single five-sided shape with an incomplete right or left side appeared on the screen briefly. In the recall task following each syllogism, a shape appeared on the screen that was either identical to or different from the shape prior to the syllogism (i.e., the incomplete side was on the same side or was on the opposite side). Half of the arrays matched the initial arrays, and half were different. Participants indicated, by key press, whether the shape was the same as or different from the shape they were asked to remember. High load trials proceeded as low load trials, except that displays consisted of four incomplete shapes instead of one in order to place a greater load on cognitive resources. In the recall test, the display was either identical to the studied array or differed in one or more shape.

WM Measure. In order to obtain an estimate of working memory, participants completed a version of the Computation Span task, adapted from Salthouse and Babcock (1991). Participants saw a series of simple addition or subtraction problems on the computer screen (e.g., 4 + 2 = ?). They were instructed to say the answer to each problem aloud while remembering the final digit in the equation (e.g., "2" in the previous example). When participants had recited the answers to all equations in the trial, the equations were removed from the screen and participants were prompted to write down the final digit from each equation, in order, on a sheet of paper.

The task began with only one equation in each trial. After three such trials, an additional equation was added to each trial. The task continued in this fashion until participants failed to correctly recall, in order, the numbers in at least two of the three trials at a given difficulty level. The WM span of each participant was thus defined as the number of digits in a set for which the participant successfully recalled digits in two of the three trials.

Procedure

Participants were tested individually and completed the deductive reasoning task first. For this task, participants read the following instructions: "This is an experiment to test people's reasoning ability. You will be given 16 problems. On each screen, you will be shown two statements and you are asked if certain conclusions (given below the statements) may be logically deduced from them. You should answer this question on the assumption that the two statements are, in fact, true. If you judge that the conclusion necessarily follows from the statements, you should answer "yes", otherwise "no". Answer "yes" by pressing the '/' key, and answer “no” by pressing the ‘Z’ key. Please take your time and be sure that you have the right answer before moving on.”

Participants completed two practice trials with the experimenter before beginning the task, and were given the opportunity to ask any questions before the experiment began. Of 16 syllogisms, eight were randomly assigned to High Load or Low Load conditions for each participant, and within each block, problems appeared in a random order. High load and low load trials were blocked within participants, with their order being randomly determined. The sequence of each trial is depicted in Figure 1. At the beginning of each trial, a fixation cross appeared on the screen for 1000 ms, followed by a display of one shape (in low load trials) or four shapes (in high load trials) for 3000 ms. Participants then evaluated a syllogism by indicating whether the conclusion was valid or invalid by pressing the appropriate key. An array of shapes followed, and participants indicated whether the array was the same as (‘/’) or different from (‘Z’) the array preceding the syllogism. Following the computerized reasoning task, participants completed the Computation Span task with the experimenter.

Results and Discussion

Sixty participants were tested. Data from one were removed due to failure to understand instructions. Prior to analyses, subjects were divided into two groups based on WM span. WM Span, according to the Computation Span task, ranged from 2 to 7, with a mean of 4.90 and a median of 5. Participants were divided into high and low span groups according to their position relative to the median. Because 17 participants scored the median span, and a median split would yield unbalanced groups, only the data from participants falling on either side of the median were analyzed. Participants scoring 4 or lower on the Computation Span task were categorized as “low span" (*n* = 20), and those scoring 6 or higher were categorized as "high span" (*n* = 22). Mean endorsement proportions, categorized by load and by WM span, are displayed in Table 1.

In order to determine effects of load and WM span on belief bias, a 2 (Load condition: High, Low) x 2 (Conclusion validity: Valid, Invalid) x 2 (Conclusion believability: Believable, Unbelievable) x 2 (WM Span: Low, High) repeated measures ANOVA was carried out on endorsement proportions, using WM Span as a between-subjects variable. Results from this ANOVA will be discussed first, in terms of the general belief bias effect, second, in terms of load effects, and third, in terms of span differences. Finally, accuracy on the load task will be discussed.

Belief Bias

The typical belief bias effect was observed in that participants were more likely to endorse valid than invalid conclusions, $F(1, 40) = 34.2$, $MSE = .663$, $p < .001$, and believable than unbelievable conclusions, $F(1, 40) = 11.8$, $MSE = .746$, $p = .001$. An interaction between the effects of validity and belief was also found, such that the effect of believability was more pronounced on invalid conclusions than valid conclusions, $F(1, 40) = 5.01$, $MSE = .349$, $p = .031$.

WM Span Differences

Mean endorsement rates of high span and low span participants, collapsed across load, are shown in Figure 2. The four-way ANOVA using endorsement proportions as the dependent variable revealed an interaction between WM span and conclusion validity, $F(1, 40) = 4.072$, $MSE = .663$, $p = .050$ (WM span did not interact with any other factors, nor was there a main effect of WM span, all *F*'s < 1). To explore this interaction and to directly explore belief bias in the two WM groups, two-way ANOVAS were performed on each WM group individually.

High span participants endorsed more valid than invalid conclusions, *F*(1, 21) = 27.28, *MSE* = 1.437, *p* > .001 and more believable than unbelievable conclusions, *F*(1, 21) = 8.47, *MSE* = 1.329, *p* = .009. The typical interaction between the effects of validity and believability was marginally significant, *F*(1, 21) = 3.13, *MSE* = .739, *p* = .085. Similarly, for low span participants there were significant main effects of validity, *F*(1, 19) = 8.33, *MSE* = 1.23, p = .009, and marginally significant effects of believability, *F*(1, 19) = 3.99, *MSE* = 1.64, *p* = .059. Interestingly, however, these effects were clearly additive; The typical interaction between validity and believability was *not* found, *F*(1, 19) = 1.71, *MSE* = .66, *p* = .20. Thus, only participants with high WM spans demonstrated the interaction characteristic of the robust belief bias effect.

For participants with low WM spans, the typical belief bias interaction was not found because they quite frequently responded “Valid” to invalid, unbelievable conclusions. Participants with high WM spans, however, were more likely to solve these problems correctly.

Load Effects

A crucial question addressed in this thesis is whether participants demonstrated a higher belief bias effect under depleted WM conditions. That is, if the visual load task successfully depleted WM resources required for deductive reasoning, participants would be expected to reason in accordance with beliefs more in the High Load condition than in the Low Load condition. Similarly, participants should use logic more successfully in the Low Load condition than in the High Load condition. Finally, the entire belief bias effect may be exacerbated under depleted WM conditions. That is, the interaction between belief and logic may be more pronounced on High Load trials than on Low Load trials.

The four-way repeated measures ANOVA indicated that WM load did not interact with validity or believability, nor with the typical Validity x Belief interaction, for either high or low WM span participants (for all, $F < 1$). Thus, loading WM had no impact on endorsement rates on the reasoning task and did not exacerbate the belief bias effect. Two interpretations are possible: Either visuospatial WM is not required for deductive reasoning, or any deficits in performance due to the load task manifested on the load task itself, rather than on the deductive reasoning task. For example, when the believability and validity of the syllogism elicit conflicting responses (e.g., a valid, unbelievable syllogism or an invalid, believable syllogism), WM resources may be directed toward deducing the correct answer to the syllogism rather than rehearsing the visuospatial pattern from the load task. This possibility was examined in the following analysis.

Mean accuracy proportions for high and low WM span participants are found in Table 2. Load accuracy was collapsed across load difficulty (i.e., High vs. Low Load), because load difficulty did not affect responding in the previous analysis discussed. To examine whether participants sacrificed accuracy on the load task to reason through syllogisms with conflicting validity and believability, a 2 (Validity) x 2 (Believability) x 2 (WM span) ANOVA was run using accuracy on the load task as the dependent variable and WM span as a between subjects variable.

There were no main effects of conclusion validity ($F = 1.17$), or believability ($F = 2.27$) on load accuracy. Interestingly, the analysis revealed a significant interaction between the effects of validity and believability, $F(1,40) = 4.84$, $MSE = .136$, $p = .034$. To explore this interaction, load accuracy corresponding to conflict and no-conflict problems was directly compared. Data confirm that participants sacrificed accuracy on the load task in order to reason about conflict problems: Memory for the arrays was less accurate for trials involving conflict problems (i.e., when beliefs and logic provided incongruent responses) than for those involving no-conflict problems, $t(41) = 2.24$, *S.E.* = .159, $p = .03$, indicating that reasoning and visuospatial rehearsal compete for the same cognitive resources – thus, visuospatial WM is in fact involved in deductive reasoning. Moreover, participants with high WM spans were more accurate on the load task than were participants with low WM spans, $F(1, 40) = 4.373$, $p = .043$. An interaction between WM Span and Believability was also found, $F(1, 40) = 4.09$, $MSE = .196$, $p = .050$, whereby individuals with low WM spans showed larger costs on the secondary task when reasoning about unbelievable, as opposed to believable, syllogisms. Other than the significant findings reported above, there were no interactions among WM span, validity, and believability (all *F*'s < 1).

To summarize the results of Experiment 1, only participants with high WM spans demonstrated the typical belief bias effect, whereas participants with low WM spans did not demonstrate the interaction typically found between validity and believability. Further, although there were no effects of the visuospatial WM load task on the reasoning task, effects of the reasoning task were found on the load task. Specifically, accuracy was lowest on the load task when participants reasoned through syllogisms in which the validity and believability of conclusions conflicted. This finding suggests that limited WM resources were particularly taxed when participants reasoned through conflict problems, and is in line with dual process theories. When the validity and believability of conclusions are not congruent, the analytic and heuristic systems provide opposite responses. Thus, participants may have utilized WM resources to suppress the biased response of the heuristic system, resulting in fewer resources available to dedicate to the load task.

Experiment 2

Experiment 2 was conducted, first, to replicate the novel finding from Experiment 1 that only participants with high WM spans show the typical belief bias effect. Second, results from Experiment 1 indicate that the visuospatial load task drew from the same WM resources as did the syllogistic reasoning task. Experiment 2 also sought to determine whether syllogistic reasoning relies on verbal WM resources, as predicted by the mental logic account of reasoning, by loading the phonological loop with a digit span task, assumed to elicit verbal rehearsal. Due to the finding that the effects of the concurrent reasoning and memory task manifested on the load task in Experiment 1, it was expected that syllogisms conflicting in validity and believability would draw WM resources away from digit rehearsal, thereby reducing accuracy on the load task in this experiment. In addition, a verbal load task may be a more domain specific WM load manipulation, given that the syllogisms are presented verbally on the computer screen. As such, the phonological load task used here may disrupt reasoning on the primary task as well as impairing performance on the secondary load task.

Method

Participants

Eighty-two undergraduate students enrolled in psychology courses at the University of Waterloo participated in exchange for course credit or for $5.

Method and Materials

In order to deplete the verbal resources likely utilized in deductive reasoning, Experiment 2 employed a verbal recall secondary task as opposed to the visuospatial recognition task used in Experiment 1. The sequence of events in one trial is shown in Figure 3. In Low Load trials, participants remembered two digits that appeared on the screen for 3000 ms, and following the syllogism recalled the numbers by keying them in, in the order given, using the keyboard and then pressing ‘ENTER.’ In High Load trials, participants remembered five digits, also shown on the screen for 3000 ms. The rest of the materials and procedure were identical to those used in Experiment 1.

Results and Discussion

Of the 82 subjects run, data from 2 were excluded from analyses due to failure to understand instructions. Again, participants were divided into two groups based on WM span. WM span ranged from 1 to 7, with a mean of 4.78 and a median of 5. A median split categorized participants into High Span (span of 5 or greater, $n = 47$) and Low Span (span of 4 or lower, $n = 33$) groups. Mean endorsement proportions, categorized by load and by WM span, are displayed in Table 3.

Again, endorsement proportions were submitted to a 2 (Load condition: High, Low) x 2 (Conclusion validity: Valid, Invalid) x 2 (Conclusion believability: Believable, Unbelievable) x 2 (WM Span: Low, High) repeated measures ANOVA, using WM Span as a between-subjects variable. Referring to this ANOVA, belief bias will be examined first, followed by span differences. Finally, load effects will be examined.

Belief Bias

As expected, the typical belief bias effect was found: Participants were more likely to endorse valid than invalid conclusions, $F(1, 78) = 37.63$, $MSE = .147$, $p < .001$, and believable than unbelievable conclusions, $F(1, 78) = 23.04$, $MSE = .164$, $p < .001$. An interaction between validity and belief was marginally significant: Effects of believability were more pronounced on invalid conclusions than valid conclusions, $F(1, 78) = 3.650$, $MSE = .096$, $p = .060$.

WM Span Differences

Endorsement proportions for High Span and Low Span participants are shown in Figure 4. In the four-way ANOVA described above, WM Span did not interact with Conclusion Validity (F = 2.14) nor with Conclusion Believability (F < 1). However, a significant three-way interaction between WM Span, Conclusion Validity, and Conclusion Believability was found, F(1, 78) = 11.78, MSE = .096 p = .001.

To explore this interaction, 2 (Validity) x 2 (Believability) ANOVAs were run on High Span and Low Span groups independently. Results corroborate the findings of Experiment 1. The typical belief bias effect was found in high span participants, with main effects of both validity, F(1, 46) = 31.160, MSE = 1.322, p < .001, and believability, F(1, 46) = 10.738, MSE = 1.340, p = .002, and a validity x believability interaction, F(1, 46) = 22.081, MSE = .602, p < .001. For low span participants, however, only main effects of validity, (1, 32) = 11.27, MSE = .971, p = .002, and believability, F(1, 32) = 16.03, MSE = 1.208, p = .001, were found. These effects were additive rather than interactive (F < 1). Thus, as in Experiment 1, only participants with high WM spans showed the interaction between validity and believability characteristic of the typical belief bias finding. Again, the interesting difference between participants with high and low WM spans is found when conclusions are invalid and unbelievable. Low span participants make more errors for these problems by endorsing the conclusions as valid, whereas high span participant are relatively accurate on these problems.

Load Effects

Like Experiment 1, in Experiment 2 I examined whether depleting WM resources by introducing a cognitive load would magnify belief bias. Again, there was no three-way interaction between load condition, conclusion validity, and conclusion believability, $F < 1$, indicating that load did not influence belief bias on the reasoning task.

Load accuracy was analyzed to determine whether participants were more accurate on the secondary load task when the validity and believability of conclusions conflicted, as was found in Experiment 1. Proportions of correct trials on the load task are shown in Table 4. Again, because load did not influence responding on the reasoning task, load accuracy was collapsed across load condition when analyzing load accuracy. Load accuracy was submitted to a 2 (Validity) x 2 (Believability) x 2 (WM Span) repeated measures ANOVA, using WM Span as a between-subjects factor. An interaction between validity and believability was found, $F(1, 78) = 4.330$, $MSE = .225$, $p = .041$. As in Experiment 1, participants were less accurate on the load task when the validity and believability of the syllogism elicited conflicting responses than when validity and believability were congruent, $t(79) = 1.996$, $SE = .150$, $p = .049$. Thus, participants again dedicated limited WM toward the reasoning task at the expense of the load task. This finding indicates that a digit span task, presumed to elicit phonological rehearsal, and syllogistic reasoning draw from the same pool of verbal WM resources, and supports the position of mental logic theorists that reasoning involves verbal representation of the problem.

This ANOVA also revealed an interaction between WM Span and Believability, $F(1, 78) = 11.426$, $MSE = .042$, $p = .001$), unlike Experiment 1, however, the locus of this interaction was such that individuals with high WM spans showed larger costs on the secondary task when reasoning with unbelievable, as opposed to believable syllogisms. No other main effects or interactions were found (all *F*'s < 1.33).

Experiment 2 replicated and extended the results of Experiment 1. First, the novel finding that only individuals with high WM capacities demonstrate the characteristic interaction between conclusion validity and believability was replicated. In particular, participants with low WM spans erroneously responded “Valid” to invalid, unbelievable conclusions more often than did participants with high WM spans. Second, syllogistic reasoning interfered with a secondary cognitive load task such that when participants were required to suppress a heuristic response that conflicted with an analytic or logical response, accuracy on the load task suffered. Taken together, the results from Experiments 1 and 2 may indicate that because syllogistic reasoning interfered with both a visuospatial and verbal memory load task, reasoning requires both visual and verbal WM resources. Conversely, it could be that both load tasks taxed resources in the central executive, and it is these overlapping resources that are crucial for reasoning through conflict problems.

General Discussion

The current studies investigated the role of WM in syllogistic reasoning using two strategies: A dual task paradigm independently loaded components of WM, and belief bias was compared across participants of different WM spans. The results and implications derived from each of these strategies will be discussed individually.

The Role of Visuospatial and Verbal WM in Reasoning

First, involvement of WM in syllogistic reasoning was directly investigated by reducing limited WM resources with a load task and examining the resulting deficits on both the reasoning task and on the load task. Although loading WM did not directly influence responding on the reasoning task, when reasoners were faced with conflict problems (i.e., for which logic and beliefs were incongruous), accuracy on the load task suffered. This finding, which was found in both Experiments 1 and 2, can be interpreted in terms of dual process theories of reasoning. For conflict problems, the analytic system and heuristic system provide different answers. For example, consider a problem with a valid but unbelievable conclusion. The heuristic system, which is thought to proceed relatively automatically and thus demands few cognitive resources, provides a response based on past knowledge or bias – In this case, “Invalid,” because the conclusion contradicts the reasoner’s knowledge about the world. Because this answer is incorrect, as is the case for heuristic responses to all conflict problems, the heuristic system must be suppressed, which requires cognitive resources. The slow and effortful analytic system then reasons logically through the problem. Because WM resources are recruited to suppress the heuristic system, fewer resources are available to dedicate to rehearsing the WM load. As a result, accuracy on the load task decreases.

Reasoning about conflict problems was associated with more errors on both the visuospatial load task used in Experiment 1 and on the phonological load task used in Experiment 2. Thus, it seems as though syllogistic reasoning (specifically, the suppression of responses based on heuristics) relies on both visuospatial and phonological WM. Another possibility, however, is that the verbal and visuospatial load tasks both taxed the central executive component of WM, which is thought to be heavily involved in reasoning (Gilhooly et al., 1993). Disentangling independent roles of the central executive, visuospatial sketchpad, and phonological loop to syllogistic reasoning poses a challenge to researchers because scores on tasks designed to tap into the three components are frequently highly correlated (Capon et al., 2003).

The findings of this study do not directly oppose either the mental models account of reasoning, which asserts that deductive reasoning problems are represented as spatial, diagram-like entities, or the mental logic account, which asserts that problems are reasoned through using verbal formal logic rules. It is possible that problems are represented both spatially and verbally. Although proponents of the mental models account clearly ascribe a large role to spatial representation in reasoning, they do not rule out concurrent verbal representation. That is, although mental models are often referred to as being iconic or spatial in nature, in their basic form they are simply representations of relations among entities. It is conceivable that verbal strategies assist in the conceptualization of these relations, especially given the assertion that models "underlie visual images, although many components of models are not visualizable" (Johnson-Laird, 2001, p. 434).

Whereas the finding that both verbal and visuospatial WM plays a role in deductive reasoning can be unified with the mental models account, it is not consistent with the mental logic account. This account staunchly rejects involvement of visual imagery in reasoning while proposing that reasoning is accomplished solely through verbally represented formal logic rules (Rips, 1994). Thus, this account would not predict the finding of Experiment 1 that visuospatial rehearsal and syllogistic reasoning competed for cognitive resources, assuming that participants did in fact rehearse the arrays visually, as opposed to using a verbal rule to represent the arrays.

Recently, fMRI studies have examined neural mechanisms involved in syllogistic reasoning. Studies by Goel and colleagues (see Goel, 2003 for a summary) corroborate the finding of the current studies that reasoning is accomplished through both visuospatial and linguistic processes. Participants in his studies reasoned through syllogisms containing statements with concrete content (e.g., “All poodles are dogs”) or abstract content (e.g., “All N are P”). He found activation primarily in a left frontal-temporal network, implicated in language and semantic processing, for problems with concrete content, and in a bilateral occipito-parietal network, implicated in the processing of spatial information, for problems with abstract content. However, *both* neural networks were necessary for reasoning about both types of problems. Goel later discovered that it is the familiar, belief-laden content in concrete problems that activates the left frontal-temporal network. He posits that the two neural networks activated in syllogistic reasoning can be mapped onto the two reasoning systems outlined by dual process accounts of reasoning. The frontal-temporal network automatically and effortlessly responds to syllogisms with belief-laden content. If no semantic or belief-laden content is available, or when a conflict is detected between the believability and validity of the conclusion, the slower and effortful parietal network is activated. Thus, according to Goel, the heuristic system maps onto the frontal-temporal network and is language-based, whereas the analytic system maps onto the parietal network and is spatially-based. This theory is not entirely consistent with the results of the current study: If the analytic system is spatially-based, we would expect reasoning about conflict problems, which should recruit the analytic system, to impair performance on the secondary load task only when the load task is also spatially-based. Conversely, Experiment 2 demonstrated that reasoning through conflict problems also impaired performance on a verbal load task, indicating that the analytic system is at least partially language-based. Nonetheless, taken together, the results from the current studies along with Goel's results suggest that neither verbal nor visuospatial reasoning strategies are utilized in isolation.

Individual Differences in Working Memory and Reasoning

A second approach used in the current studies compared the reasoning of participants with low WM spans to participants with high WM spans. Both Experiments 1 and 2 demonstrated that Low Span participants made significantly more errors on Invalid, Unbelievable syllogisms than did High Span participants (see Figures 2 and 4). This novel finding was unexpected, given that dual process theories predict that individuals with limited WM resources should rely more on heuristic processes when faced with a challenging logic problem. Thus, these theories would predict that participants with low WM spans would respond "Invalid" more frequently to unbelievable conclusions. Rather, it seems as though participants with more cognitive resources were particularly astute at indentifying when the believability and validity of conclusions conflicted and when they did not. When there was no conflict, such as when conclusions were invalid and unbelievable, these participants were able to utilize belief information, along with validity information, to their advantage when reasoning through syllogisms. In contrast, participants with fewer cognitive resources ignored belief information.

This finding is consistent with past studies indicating that when belief information is available, people are at times more accurate when reasoning. Newstead, Pollard, Evans, and

Allen (1992; and later replicated by Evans, Newstead, Allen, & Pollard, 1994), found relatively high error rates for invalid conclusions free of belief information. People were more accurate when reasoning about invalid conclusions when the conclusions were unbelievable. Interestingly, however, the converse was not found when conclusions were believable: People did not endorse more invalid conclusions when conclusions were believable compared to neutral. In general, reasoners in Newstead's studies seemed to use belief information to aid their reasoning only when it was useful; otherwise, it was ignored.

Differences between high WM and low WM reasoners found in the current studies can be explained in terms of dual process theories. In particular, high WM reasoners were able to recognize when the believability and validity of conclusions did not conflict. Thus, when conclusions were invalid and unbelievable, they likely responded in accordance with heuristic processing, without appealing to analytic processing. Conversely, low WM participants seemed unable to recognize that when conclusions were unbelievable, belief information could aid their reasoning. They likely used analytic processing to reason through all problems with unbelievable conclusions, and because analytic processing is sensitive to WM limitations, made more errors. Response time analyses in future studies may support the hypotheses set forth by the current studies. Because high WM reasoners are thought to use quick, heuristic processes for problems in which the validity and believability of the conclusion conflicts, their responding on no-conflict problems should be much quicker than their responding on no-conflict problems, for which they presumably use slow, effortful analytic reasoning. Conversely, response latencies of low WM reasoners should be equally long for conflict and no-conflict problems (particularly for invalid, unbelievable problems) if they resort to analytic reasoning for conflict and no-conflict problems alike.

Conclusions

Utilizing belief information when it is consistent with the logical validity of the conclusion is naturally a very useful strategy, particularly given that in everyday life, we do not often have to reason about scenarios that are contrary to what we know about the world. By using belief-based heuristics, reasoners can avoid time-consuming, effortful analytic reasoning. Thus, perhaps a hallmark of astute reasoners is the ability to utilize information consistent with their knowledge about the world, along with information about the objective validity of conclusions, when reasoning. Indeed, good decision-making involves integrating several sources of useful information while ignoring information that may be misleading. In the current studies, only individuals with sufficient WM resources were able to do this. Examining whether reasoning according to biases is a useful strategy in syllogistic reasoning, and in decision-making in general, should be the focus of future studies.

References

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 47-90). New York: Academic Press.

Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail: Working memory and "choking under pressure" in math. *Psychological Science, 16*(2), 101-105.

Beilock, S. L., & DeCaro, M. S. (2007). From poor performance to success under stress: Working memory, strategy selection, and mathematical problem solving under pressure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(6), 983-998.

Braine, M. D. S. & O'Brien, D. P. (Eds.). (1998). *Mental logic* Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Capon, A., Handley, S., & Dennis, I. (2003). Working memory and reasoning: An individual differences perspective. *Thinking & Reasoning, 9*(3), 203-244.

Chater, N., & Oaksford, M. (2001). Human rationality and the psychology of reasoning: Where do we go from here? *British Journal of Psychology, 92*(1), 193-216.

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences, 4*(3), 317-370.

Copeland, D. E., & Radvansky, G. A. (2004). Working memory and syllogistic reasoning. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 57A*(8), 1437-1457.

De Neys, W. (2006). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science, 17*(5), 428-433.

Evans, J. St. B. T. (2007). On the resolution of conflict in dual process theories of reasoning. *Thinking & Reasoning, 13*(4), 321-339.

Evans, J. St. B. T., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition, 11*(3), 295-306.

Evans, J. St. B. T., Handley, S. J., Harper, C. N. J., & Johnson-Laird, P. N. (1999). Reasoning about necessity and possibility: A test of the mental model theory of deduction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(6), 1495-1513.

Evans, J. St. B. T., Newstead, S. E., Allen, J. L., & Pollard, P. (1994). Debiasing by instruction: The case of belief bias. *European Journal of Cognitive Psychology, 6*(3), 263-285.

Evans, J. St. B. T. & Over, D. E. (1996). *Rationality and reasoning*. Oxford, England: Psychology/Erlbaum (UK) Taylor & Fr.

Gilhooly, K. J., Logie, R. H., Wetherick, N. E., & Wynn, V. (1993). Working memory and strategies in syllogistic-reasoning tasks. *Memory & Cognition, 21*(1), 115-124.

Gilhooly, K. J., Logie, R. H., & Wynn, V. E. (2002). Syllogistic reasoning tasks and working memory: Evidence from sequential presentation of premises. *Current Psychology, 21*(2), 111-120.

Goel, V. (2003). Evidence for Dual Neural Pathways for Syllogistic Reasoning. *Psychologica,* Vol. 32, pp. 301-309.

Johnson-Laird, P. N., & Bara, B. G. (1984). Syllogistic inference. *Cognition, 16*(1), 1-61.

Johnson-Laird, P. N. (2001). Mental models and deduction. *Trends in Cognitive Sciences, 5*(10), 434-442.

Klauer, K. C., Musch, J., & Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychological Review, 107*(4), 852-884.

Knauff, M., Mulack, T., Kassubek, J., Salih, H. R., & Greenlee, M. W. (2002). Spatial imagery in deductive reasoning: A functional MRI study. *Cognitive Brain Research, 13*(2), 203-212.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?p. *Intelligence, 14*(4), 389-433.

Newstead, S. E., Pollard, P., & Evans, J. S. (1992). The source of belief bias effects in syllogistic reasoning. *Cognition, 45*(3), 257-284.

Newstead, S. E., Thompson, V. A., & Handley, S. J. (2002). Generating alternatives: A key component in human reasoning? *Memory & Cognition, 30*(1), 129-137.

Osman, M. (2004). An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin & Review, 11*(6), 988-1010.

Quayle, J. D., & Ball, L. J. (2000). Working memory, metacognitive uncertainty, and belief bias in syllogistic reasoning. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 53A*(4), 1202-1223.

Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking*. Cambridge, MA, US: The MIT Press.

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General, 126*(3), 211-227.

Salthouse, T. A., & Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology, 27*(5), 763-776.

Schroyens, W. J., Schaeken, W., & d'Ydewalle, G. (2001). The processing of negations in conditional reasoning: A meta-analytic case study in mental model and/or mental logic theory. *Thinking & Reasoning, 7*(2), 121-172.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*(1), 3-22.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23*(5), 645-726.

Appendix A: Syllogisms used in Experiments 1 and 2

Valid, Believable	Invalid, Believable
No video recorders are metal things Some electric appliances are metal things Therefore, some electric appliances are not video recorders	No electric appliances are metal things Some video recorders are metal things Therefore, some electric appliances are not video recorders
No astronauts are athletic Some healthy people are athletic Therefore, some healthy people are not astronauts	No healthy people are athletic Some astronauts are athletic Therefore, some healthy people are not astronauts
No skyscrapers are wooden things Some buildings are wooden things Therefore, some buildings are not skyscrapers	No buildings are wooden things Some skyscrapers are wooden things Therefore, some buildings are not skyscrapers
No cigarettes are inexpensive Some addictive things are inexpensive Therefore, some addictive things are not cigarettes	No addictive things are inexpensive Some cigarettes are inexpensive Therefore, some addictive things are not cigarettes
No police dogs are vicious Some highly trained dogs are vicious Therefore, some highly trained dogs are not police dogs	No highly trained dogs are vicious Some police dogs are vicious Therefore, some highly trained dogs are not police dogs
No liquors are sweet things Some drinks are sweet things Therefore, some drinks are not liquors	No drinks are sweet things Some liquors are sweet things Therefore, some drinks are not liquors
No trout are colourful things Some fish are colourful things Therefore, some fish are not trout	No fish are colourful things Some trout are colourful things Therefore, some fish are not trout
No millionaires are scientists Some rich people are scientists Therefore, some rich people are not millionaires	No rich people are scientists Some millionaires are scientists Therefore, some rich people are not millionaires

Valid, Unbelievable	Invalid, Unbelievable
No pieces of furniture are attractive things Some tables are attractive things Therefore, some tables are not pieces of furniture	No tables are attractive things Some pieces of furniture are attractive things Therefore, some tables are not pieces of furniture
No musical instruments are shiny Some violins are shiny Therefore, some violins are not musical instruments	No violins are shiny Some musical instruments are shiny Therefore, some violins are not musical instruments
No well-educated people are impartial Some judges are impartial Therefore, some judges are not well-educated people	Some judges are impartial Some well-educated people are impartial Therefore, some judges are not well-educated people
No plants are living things Some grasses are living things Therefore, some grasses are not plants	No grasses are living things Some plants are living things Therefore, some grasses are not plants
No religious people are married Some priests are married Therefore, some priests are not religious people	No priests are married Some religious people are married Therefore, some priests are not religious people
No good swimmers are vegetarians Some deep sea divers are vegetarians Therefore, some deep sea divers are not good swimmers	No deep sea divers are vegetarians Some good swimmers are vegetarians Therefore, some deep sea divers are not good swimmers
No tools are heavy Some screwdrivers are heavy Therefore, some screwdrivers are not tools	No screwdrivers are heavy Some tools are heavy Therefore, some screwdrivers are not tools
No nutritional things are expensive Some vitamin tablets are expensive Therefore, some vitamin tablets are not nutritional things	No vitamin tablets are expensive Some nutritional things are expensive Therefore, some vitamin tablets are not nutritional things

Table 1

Experiment 1: Mean Endorsement Proportions for Low and High WM Span Participants under Low and High Load

	Low Load	High Load
Low Span (*n* = 22)		
Valid/Believable	.71 (.37)	.80 (.30)
Valid/Unbelievable	.61 (.37)	.73 (.37)
Invalid/Believable	.66 (.32)	.61 (.38)
Invalid/Unbelievable	.43 (.36)	.45 (.43)
High Span (*n* = 20)		
Valid/Believable	.82 (.24)	.78 (.34)
Valid/Unbelievable	.70 (.34)	.70 (.38)
Invalid/Believable	.52 (.41)	.55 (.39)
Invalid/Unbelievable	.20 (.34)	.32 (.37)

Note: Values in parentheses are standard deviations.

Table 2

Experiment 1: Mean Accuracy Proportions on Load Task for Low and High WM Span Participants

	Low Span (*n* = 22)	High Span (*n* = 20)
Valid/Believable	.94 (.11)	.95 (.13)
Valid/Unbelievable	.81 (.23)	.93 (.12)
Invalid/Believable	.93 (.18)	.93 (.12)
Invalid/Unbelievable	.90 (.15)	.97 (.11)

Note: Values in parentheses are standard deviations.

Table 3

Experiment 2: Mean Endorsement Proportions for Low and High WM Span Participants under Low and High Load

	Low Load	High Load
Low Span (*n* = 33)		
Valid/Believable	.76 (.33)	.82 (.24)
Valid/Unbelievable	.53 (.37)	.62 (.38)
Invalid/Believable	.59 (.40)	.62 (.40)
Invalid/Unbelievable	.54 (.34)	.39 (.35)
High Span (*n* = 47)		
Valid/Believable	.65 (.33)	.71 (.33)
Valid/Unbelievable	.67 (.32)	.68 (.35)
Invalid/Believable	.57 (.39)	.59 (.38)
Invalid/Unbelievable	.29 (.36)	.32 (.37)

Note: Values in parentheses are standard deviations.

Table 4

Experiment 2: Mean Accuracy Proportions on Load Task for Low and High WM Span Participants

	Low Span (n = 22)	High Span (n = 20)
Valid/Believable	.61 (.24)	.71 (.28)
Valid/Unbelievable	.62 (.31)	.61 (.25)
Invalid/Believable	.55 (.21)	.73 (.25)
Invalid/Unbelievable	.69 (.24)	.66 (.20)

Note: Values in parentheses are standard deviations.

Figure Captions

Figure 1. Experiment 1 high load task sequence.

Figure 2. Experiment 1: Mean endorsement proportions for Low Span and High Span participants collapsed across load condition.

Figure 3. Experiment 2 low load task sequence.

Figure 4. Experiment 2: Mean endorsement proportions for Low Span and High Span participants collapsed across load condition

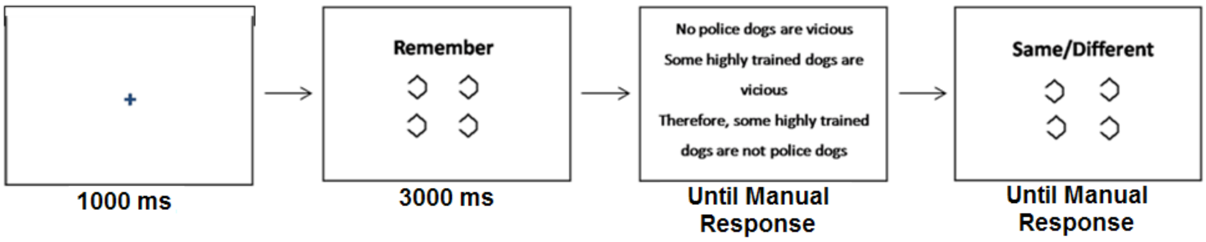
Remember
No police dogs are vicious
Some highly trained dogs are
vicious
Therefore, some highly trained
dogs are not police dogs
Same/Different
1000 ms
3000 ms
Until Manual
Response
Until Manual
Response

Figure 1

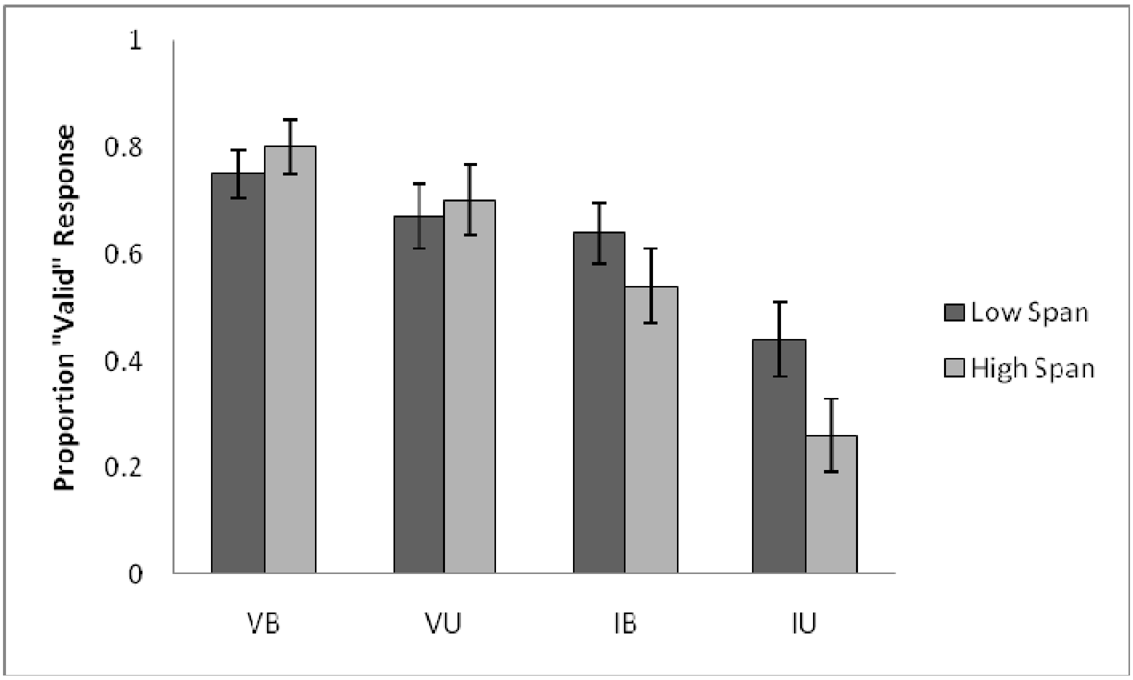
1
0.8
0.6
0.4
0.2
0
Proportion "Valid" Response
VB
VU
IB
IU
Low Span
High Span

Figure 2

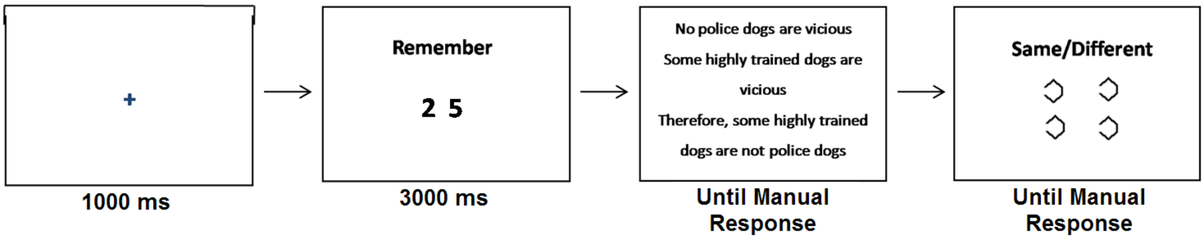
+
1000 ms
Remember
2 5
3000 ms
No police dogs are vicious
Some highly trained dogs are vicious
Therefore, some highly trained dogs are not police dogs
Until Manual Response
Same/Different
Until Manual Response

Figure 3

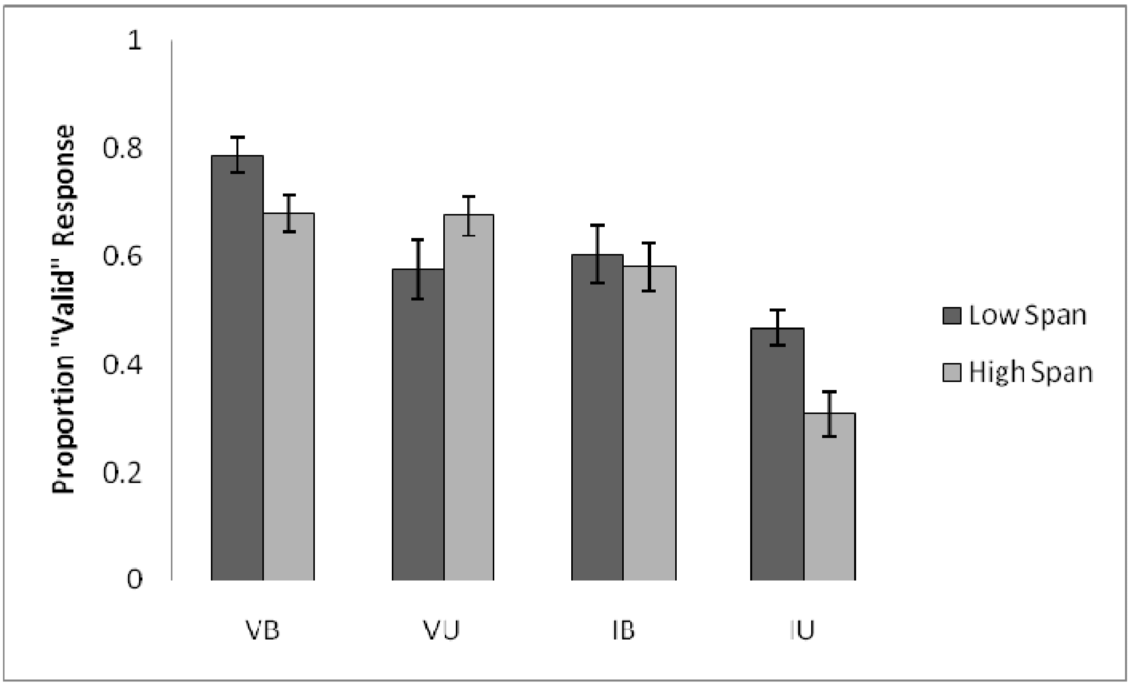
Proportion "Valid" Response
1
0.8
0.6
0.4
0.2
0
VB
VU
IB
IU
Low Span
High Span

Figure 4