ABSTRACT

PREDICTORS OF REASONING ABILITY: WORKING MEMORY CAPACITY AND FUZZY PROCESSING PREFERENCE INDEX

by Audrey Margaret Weil

Working memory capacity and fuzzy processing preference are two constructs that predict reasoning ability. To date, no studies have examined the relationship between these two constructs. This study aimed to test fuzzy-trace theory's prediction of a gist processing preference in light of working memory capacity, and measure the extent to which these two constructs interact to affect reasoning ability. I assessed this relationship using an individual differences measure and a dual-task paradigm. Reasoning ability was examined in the context of base rate neglect and conjunction fallacies. Most participants preferred to reason with gist. Working memory capacity and fuzzy processing preference were independent of each other, but did not interact to produce sound reasoning. Surprisingly, fuzzy processing preference did not predict reasoning ability, and for base rate questions, predicted higher errors. However, there is evidence that data may have been confounded by various issues, leading to counterintuitive findings.

PREDICTORS OF REASONING ABILITY: WORKING MEMORY CAPACITY AND FUZZY PROCESSING PREFERENCE

A Thesis

Submitted to the

Faculty of Miami University

In partial fulfillment of

The requirements for the degree of

Master of Arts

Department of Psychology

by

Audrey M. Weil

Miami University

Oxford, Ohio

2014

Dr. Christopher Wolfe

Dr. David Waller

Dr. Robin Thomas

Table of Contents

List of Figures

Introduction

The ability to reason and to navigate our way through problems is one of the crowning achievements of intelligence. However, when presented with a formal logic problem, people often fail to give the correct normative response (Tversky & Kahneman, 1973; Wolfe, Fisher & Reyns, 2012). This phenomenon has been observed in a wide variety of reasoning tasks such as syllogisms (Gilhooly, Logie, & Wynn, 1999; Meiser, Klauer, & Naumer, 2001), selection tasks (De Neys, 2006a), joint probability problems (Wolfe & Reyna, 2010b; Tversky & Kahneman, 1983) and base-rate problems (Wolfe, 1995). As such, the study of rationality and reasoning has dominated much of the cognitive literature throughout the past several decades (for review see: Evans, 2002; Johnson-Laird, 1999).

Although there are several theories to explain why our reasoning ability is limited, one of the most well-established explanations is that reasoning is directly constrained by our working memory capacity (Capon Handley, & Dennis, 2003; Kyllonen & Christal, 1990; Gilhooly, Logie, & Wynn, 2002; Markovitz, Doyon & Simoneau, 2002). Working memory is conceptualized as a temporary storage and maintenance system that allows people to manipulate mental representations in the pursuit of a goal (Baddely, 1992, 2007; Cowan, 1995, 2009). This system is hierarchically structured such that slave systems like the phonological loop and the visuospatial sketchpad serve a central component responsible for control of information processing (Baddeley and Hitch, 1974; Cowan, 1995). The central component, also known as the central executive, has a limited capacity through which it can regulate attentional resources. Engle, Tuholski, Laughlin, and Conway (1999) demonstrated a strong link between working memory and general fluid intelligence. In addition, several other studies have found support for a clear relationship between working memory capacity and performance on a wide variety of cognitive tasks such as reading comprehension (Daneman & Carpenter, 1980), problem solving (La Pointe & Engle, 1990; Unsworth & Engle, 2007), and IQ tests (Kane, Bleckley, Conway, & Engle, 2001). The link between intelligence, cognitive ability and working memory capacity has led many to believe that working memory capacity is the driving force behind our ability to reason (Kyllonen & Christal, 1990; Capon et al., 2003; Markovits et al., 2002; Süß, Oberauer, Wittmann, Wilhelm, & Schulze, 2002).

Research on the relationship between working memory capacity and reasoning ability has historically taken one of two routes: an individual differences or a dual-task approach. The individual differences approach to studying working memory allows researchers to assess the relationship between working memory and other cognitive abilities, and further understand what differentiates working memory from highly related processes such as IQ or reasoning ability. They accomplish this by measuring variation across participants in both working memory and another factor of interest to assess the relationship between the factors. For instance Daneman and Carpenter's (1980) study on individual differences in reading comprehension and working memory capacity found that the participants who had higher working memory capacity also had a higher reading comprehension, indicating that they are highly related processes. Research using this style of assessment has similarly found that participants with high working memory capacity tend to have a high reasoning ability when compared to participants with lower working memory capacity (Capon et al., 2003; Kane, Hambrick, Tuholski, Wilhelm, Payne, & Engle, 2004).

Dual-tasks paradigms, on the other hand, take an experimental approach by burdening executive resources (for a review see: Pashler, 1994). They accomplish this by requiring participants to perform a secondary task that requires working memory capacity such as remembering number sequences (Sternberg, 1966) or tapping patterns on a keyboard (Kane &

Engle, 2000; Moskovitch, 1994) while they also complete a primary task. Such an approach enables researchers to infer causal relationships between working memory capacity and performance on various tasks following the logic of experimentation. Research investigating the relationship between reasoning and working memory through dual-task paradigms has found that when people's working memory capacity is taken up by a dual-task their reasoning ability suffers (Halford, Maybery, & Bain, 1986; De Neys, 2006a).

From these two bodies of literature we can see that working memory capacity is necessary for reasoning ability, and that higher working memory capacity is associated with higher reasoning ability. However, despite the undoubtedly strong relationship between working memory capacity and reasoning, there is still some variation in reasoning that is not accounted for by working memory capacity (Kane et al., 2004). Dual-processing theories offer another promising explanation in the ongoing investigation into reasoning errors. Dual-processing theories posit that people use two different systems to process information (Evans, 2003). Although there are a wide variety of dual-processing models (for review see: Evans, 2008), most agree on a distinction between a heuristic system and an analytic system. Heuristic processing is unconscious, automatic, based on personal experiences, contextualized, and does not require executive resources. Analytic processing, on the other hand, is more conscious, controlled, detail-oriented, decontextualized, and demands executive resources (Kahneman, 2011).

According to most dual-processing theories, the systematic errors we typically see in reasoning are due to the cursory, associative reasoning that is typical of heuristic processing. In this way heuristic processing is characterized as evolutionarily primitive, stereotypical, impulsive, and is often the cause of errors in reasoning. Analytic processing, on the other hand, is thought to be responsible for thoughtful and deliberate reasoning which often leads to normatively correct performance (Stanovich & West, 1998).

The simplicity of the heuristic-analytic dichotomy provides an appealing view of cognition. However, in this simplicity it categorizes affect- and personal-experience-based heuristic processing as inherently inferior to detached, quantitative analytical processing. Fuzzy-trace theory is a dual-processing theory that gives heurist processing a relatively more respectable place in mature reasoning, and more accurately reflects the type of reasoning we observe in every-day scenarios (Reyna & Brainerd, 1995, 2011; Reyna, 2008, 2012; Brainerd & Reyna, 2001). Fuzzy-trace theory's two processing systems, gist and verbatim, are similar to those of other dual-processing systems in that they represent two systems of mental representation and processing. However, fuzzy-trace theory makes several novel predictions that are not made by standard dual-processing theories, such as a developmental increase in gist processing (Reyna & Casillas, 2009), an increased reliance on gist processing as expertise develops (Reyna & Lloyd, 2006), and a preference for verbatim processing in people with autism (Reyna & Brainerd, 2011). Much like other dual-processing theories, gist processing is automatic, does not require as much working memory capacity, and is based on affect and personal experiences. Similarly, verbatim, much like analytic processing, is controlled, requires working memory capacity, and is decontextualized.

One of the key differences between gist and heuristic processing is that gist processing is really focused on extracting the essence of the information – the decision-relevant meaning. For example, in the sentence "Anna went to Farm Mart to buy strawberries, mangoes, and bananas for her smoothie," for most purposes, the gist of the information is that Anna went to the store to buy fruit. The details aren't important, but rather what the sentence is trying to convey. For a friend listening with an allergy to mangos the gist of that same sentence might be, Anna bought

mangos at the store. On the opposite end of the spectrum, the main difference in verbatim processing is that it deals primarily in surface details, and not necessarily with meaning. So in our example sentence the verbatim information would be the specific fruit she bought, the girl's name, or the store she went to. These pieces of information by themselves do not convey meaning. When these two types of representations, gist and verbatim, are combined together in a meaningful way, we can often see a more complete and informative picture than either individually.

According to fuzzy-trace theory, we encode, store, and retrieve both gist and verbatim information separately and in parallel This notion that both types of representations are encoded in parallel as opposed to the traditional idea that meaning is derived from the processing of surface information is unique to fuzzy-trace theory. These representations are stored along a continuum, with fuzzy gist representations on one end, and precise verbatim representations on the other. Thus information is encoded such that global gist representations are stored alongside precise verbatim representations (Reyna & Casillas, 2009). In addition, fuzzy-trace theory posits that people prefer to reason with representations closer to the gist end of the continuum rather than the more precise verbatim information. (Brainerd & Reyna, 2001; Reyna 2012). People want to understand what the information really means, i.e. the bottom line, rather than know the precise or superficial details. The underlying meaning that people seek often reflects personal experience, knowledge, and affect, all of which are often integral to decision making in everyday life.

In much of the dual-task literature analytical processing is seen as the more correct or evolved processing, and heuristic processing is seen as primitive (Chiesi et al., 2011, Stanovich & West, 1998; Stanovich, Toplak, & West, 2008; Evans, 2002; Evans, 2003; Kokis, Macpherson, Toplak, West, & Stanovich, 2002). Fuzzy-trace theory, on the other hand, claims that not only do the majority of people prefer to process with gist information, but that gist processing is in many circumstances superior to verbatim processing (Reyna & Lloyd, 2006; Reyna & Brainerd, 1995). For example, despite the fact that both gist and verbatim information are encoded in parallel, memory of verbatim information fades more rapidly than gist traces (Brainerd, Reyna, & Brandse, 1995). These verbatim traces are also more sensitive to interference, and require a greater degree of working memory capacity (Reyna & Lloyd, 2006). Gist, on the other hand, results in fewer errors compared to the fragile and cumbersome verbatim traces (Reyna & Lloyd, 2006). In the face of information overload, gist processing enables people to organize and focus on the meaningful information, leading to better decision making (Fukukura, Ferguson, & Fujita, 2012). Along with fewer errors and higher stability, gist reasoning is also associated with higher rates of positive decisions, such as less unhealthy risk taking in adolescents (Reyna & Farley, 2006). Furthermore, when people think using gist representations they often process information unconsciously, allowing for parallel rather than serial processing (Reyna & Brainerd, 1992). For these reasons, as well as its simplicity and ease of use, gist processing tends to result in more coherent thinking (Reyna & Brainerd, 1995).

Reliance on the adaptive tendency to process information using gist increases both with age and development (Reyna & Casillas, 2009). Studies have shown that young children tend to rely on verbatim representations when reasoning, but later switch to gist sometime around middle childhood (Reyna & Casillas, 2009; Reyna & Farley, 2006; Reyna, 1996; Reyna & Ellis). Studies have also found that experts typically rely on gist processing more than verbatim when making important decisions in their field (Reyna & Lloyd, 2006; Reyna & Brainerd, 1994).

Thus, gist processing is not just important for casual reasoning, but also for advanced and mature reasoning.

Though in many ways adaptive, gist by itself is often not enough for normative reasoning tasks. In these tasks the detailed information is crucial to understanding the problem, and often the essence of the problem can lead to fallacies like belief bias or conjunction fallacies (Reyna, 2008). According to fuzzy-trace theory, the "trick" to successfully navigating through these reasoning problems is in forming an appropriate gist representation and understanding when more precise information is required. Thus the ability to combine the gist of the problem with the appropriate verbatim information leads to better reasoning (Wolfe & Fisher, 2013).

Wolfe and Fisher (2013) created an instrument consisting of a set of logic problems that assesses individual differences in ability to incorporate verbatim information into their probability estimates as a predictor for gist or "fuzzy" processing preference. In these problems participants are presented with several base rate problems in which the quantitative base rate pulls in the opposite direction as the qualitative text in the problem. An example of one item is "at Cloverdale High School 10% of the seniors go on to college. Bob is a senior at Cloverdale High. He gets mostly As and Bs in school and is well liked by his teachers. What is the probability that Bob will go to college?" Here the gist of the problem is that Bob has a good chance of going to college because he is a good student. The verbatim information is that quantitative base rate which indicates that students from Cloverdale High have a low probability of going to college. By measuring which type of information participants weigh more heavily, the text or the number, we can measure their gist or verbatim processing preference.

This fuzzy processing preference index (FPPI) has been successfully used to measure individual differences in people's tendency to engage in gist processing or integrating quantitative verbatim information with qualitative gist. Wolfe and Fisher combined the results from five of their studies using the FPPI, and found that, consistent with fuzzy-trace theory's predictions, the majority of participants tended to use gist processing over verbatim processing, and that people who were able to incorporate verbatim information into their gist representations (i.e. had a higher FPPI score) were better reasoners.

Fuzzy-trace theory is largely silent with respect to how working memory capacity may interact with processing style. However, studies have shown that people who have higher cognitive ability (e.g. people with high working memory capacity) are less likely to commit logical fallacies than those with lower cognitive ability (Evans, Handley, Neilens, Bacon, & Over, 2010; Stanovich & West, 1997). We also know that verbatim processing requires more working memory capacity than gist (Reyna & Casillas, 2009). It is reasonable to assume that those with high working memory capacity are aided in their reasoning attempts because they would be more inclined or able to incorporate verbatim and gist information in comparison to those with low working memory capacity. This, along with the ability to hold and manipulate more information, would lead to fewer fallacies for individuals with high working memory capacity. If this is indeed the case it is reasonable to assume there may be some interaction between these two constructs. To date, however, no studies have directly tested this hypothesis. Therefore, the first aim of this study is to outline the heretofore undefined relationship between fuzzy processing preference and working memory capacity in regards to reasoning ability.

In order to assess reasoning ability I chose to use conjunction fallacies and base rate problems, both of which are well-established reasoning problems in the dual processing literature (Tversky & Kahneman, 1983; Kahneman, Slovic, & Tversky, 1982; Kahneman & Tversky, 1973; Reyna & Brainerd, 2008; Wolfe, 1995; Bar-Hillel, 1980; Hamm, 1993, 1994). Both of

these problems require integration of the gist of the problem and the detailed verbatim information.

In a typical conjunction question people are presented with a short description of a person or scenario, and are asked to rate several probabilities based on this description. Kahneman and Tversky's (1983) classic Linda problem is an apt example: Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in antinuclear demonstrations. What is the probability that A) Linda is a feminist, B) Linda is a bank teller, C) Linda is both a bank teller and a feminist? When presented with this problem, people often erroneously rate the probability that Linda is both a bank teller and a feminist as more likely than the probability that she is a bank teller (Tversky & Kahneman, 1983). One of the most basic rules of probability is that the joint probability of two events cannot be greater than the probability of its constituents. This formula of P(A and B) > P(A) is known as the conjunction fallacy.

Kahneman and Tversky (1983) attributed this fallacy to the use of the representative heuristic, which refers to people's tendency to assess the probability of an event by how similar it is to our preexisting prototype of that event. The description of Linda is more representative of a feminist than a bank teller, thus prompting people to rate the probability of her being both a bank teller and a feminist as more likely than the probability that she is a bank teller. A more recent approach taken by fuzzy-trace theory (Reyna, 2008, 2012) ascribes the conjunction fallacy to denominator neglect. The basic idea is that when we are prompted to recall information about the two classes or objects being compared in conjunction questions we can be easily confused by overlapping or nested classes. In light of this confusion, we tend to reason based on the salient gist of the problem which is often a comparison of numerators (e.g. bank tellers who are feminists). This often leads individuals to overemphasize the numerator and ignore the denominator (e.g. bank tellers who are feminists plus bank tellers who aren't feminists). Thus we can see that emphasizing the numerator at the expense of the denominator leads people to commit the conjunction fallacy.

In addition to conjunction fallacies, I also chose to include base-rate fallacies in my measure of reasoning ability. A base-rate problem contains an overall probability for a given outcome, known as a base-rate, and a description of the scenario that makes that outcome more or less likely. For example, "on the advanced deductive reasoning task, 10% answer at least 9 out of 10 questions correctly. Michelle enjoys solving puzzles and brainteasers. What is the probability that Michelle will correctly answer at least 9 out of 10 questions on the advanced deductive reasoning task?" (Wolfe & Fisher, 2013). Here we have two types of information given to us. The first is the background data on the distribution of people's answers – specifically that 10% of the population answer at least 9 out of 10 questions correctly. This is known as the base-rate. The other type of information given relates specifically to the person in question. In order to reason correctly through this problem both types of information must be taken into consideration. However, the majority of people will often overemphasize the particular information about the event, and discount the base-rate. This leads many to respond with a high probability when a more realistic probability is actually much lower. This type of response in called base-rate neglect (Reyna & Brainerd, 2008).

In order to answer the types of reasoning problems used in this study correctly participants will need to incorporate both gist and verbatim information into their answer. Understanding the meaning, or gist, of the events described in the problem is essential in

reasoning problems, as is the ability to process specific verbatim information such as the base-rate or the probability that Linda is a bank teller. Effectively combining these two types of information into a whole understanding of the problem requires working memory capacity.

This study will employ both an individual differences and dual-task approach to understand more holistically what the relationship is between fuzzy processing preference and working memory capacity in regards to reasoning ability. First participants will take two individual differences tests to measure their fuzzy processing preference and their working memory capacity. In addition, half of the participants will randomly be assigned to a dual-task condition while they perform the reasoning problems, and half will be assigned to a control task. This dual-task will occupy working memory resources and enable me to directly show the degree to which working memory capacity is necessary for accurate analytical reasoning.

With evidence to suggest that both high working memory capacity and fuzzy processing preference are positively correlated with reasoning ability, I predict that high working memory capacity is necessary but not sufficient to reason effectively through analytical problems. Both the processing of verbatim information as well as the incorporation of this information and the bottom-line meaning of the problem will require working memory capacity. Therefore, in the absence of any additional cognitive load those with both high working memory capacity and a verbatim processing preference will achieve the highest level of proficiency in formal reasoning.

In sum, I tested the levels of contribution working memory capacity and fuzzy processing preference had on reasoning ability through the use of a dual-task paradigm and an individual difference measures. In addition, I tested fuzzy-trace theory's prediction that the majority of the population prefer to use gist processing. I predicted that working memory capacity and fuzzy processing preference will be independent of one another, but that both working memory capacity and a high FPPI score are needed for optimal reasoning ability. I also predicted that different combinations of working memory capacity and fuzzy processing preference would be differentially affected by the dual-task.

Method

Participants

A total of 174 undergraduate students from the psychology participant pool at Miami University in Oxford, Ohio were included in the final sample. Of the original 222 participants, 15 were excluded because they failed to meet the previously established cut-off of 85% accuracy score on the automated operation-span task (Unsworth, et al., 2005). Twenty-four were excluded for getting the dual-task manipulation wrong at least 75% of the time, and 9 were excluded for answering each probability in a set of questions as equally likely.

Measure of Fuzzy Processing Preference

The fuzzy processing preferences index (Wolfe & Fisher, 2013) is an instrument used to measure individual differences in fuzzy processing preference. FPPI is reliable, consistently producing Cronbach's Alpha > 0.90 in 5 studies. Content validity was supported by a solid correlation with the Process Dissociation Procedure, and the fact that the FPPI scores reliably predicted conjunction fallacies in joint probability estimates. It consists of 19 base rate problems (see Appendix A) and 4 M-scale problems (see Appendix B). The M-scale problems consisted of descriptions of scenarios that could have a reasonably high or low outcome, but the specific information provided in the text of the problem makes the likelihood of that outcome either extremely low or extremely high. For example one problem from the M-scale reads: "In Little Rock, Arkansas only 10% of the High School soccer referees are women. Sam has been a High School soccer referee for three years. Sam will not be refereeing this year because Sam is

pregnant. What is the probability that Sam is a woman?" Given that Sam is pregnant, it is safe to assume she is a woman. However, participants who estimate that there is a 10% chance Sam is a women are simply copying the provided base-rate instead of reading through the problems and integrating gist and verbatim information. By eliminating participants who responded using a simple matching strategy I was able to collect data that more accurately measures participants' processing preference.

Measure of Working Memory Capacity

Participants' individual differences in working memory capacity were measured with an automated o-span task (Unsworth, Heitz, Shrock & Engle, 2005) an adaptation of the operation-span task (La Pointe & Engle, 1990) designed for group testing. The instrument produced a Cronbach's Alpha of .78 indicating high reliability (Unsworth et al., 2005). The correlation between the Automated o-span and the Turner and Engle's (1989) o-span was .57, suggesting acceptable validity. The task involves presenting a math problem to the participant followed by a number. The participant judges whether the presented number is the correct solution to the problem or not. A letter is then presented for 800ms. This sequence of events repeats itself 3 to 7 times. The participants are then instructed to recall the letters in the correct order, and are then given appropriate feedback. The measure of working memory is based on the total number of correctly recalled letters.

Experimental Tasks

Participants were presented with a total of 20 reasoning problems consisting of 12 conjunction questions adapted from Fisher and Wolfe (in press) (see Appendix C) and 8 base rate questions (see Appendix D). All problems were randomly presented to the participants through E-prime 2.0.8

For the joint probability questions, participants read a short description, and then rated the probability of 4 items P(A), P(B), P(A and B), and P(A or B). If the participants said that P(A and B) > P(A) or P(A and B) > P(B) then that problem was scored as having a conjunction fallacy. Participants could have had up to 12 conjunction fallacies.

For the base rate questions, participants read a short description of a scenario with a provided base rate. Participants then assigned a probability to the outcome of the scenario described. Participants' accuracy for the base rate problems was assessed by measuring the absolute difference between their answer and the correct answer. Participants with a lower absolute difference were labeled as having a higher reasoning ability.

Dual-task

For the dual-task, I used a spatial dot memory task (Law, Morrin, & Pellegrino, 1995). This task is a spatial memory task that has been shown to load the central executive (Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001). In this task a 3x3 matrix was presented for 850 ms. In this matrix four of the squares were filled with a dot. Participants were instructed to memorize the dot configuration, maintain it in their mind while answering a logic problem, and reproduce it afterwards (see figure 1). One advantage of this task is that it appears to tax executive functions without interfering with participants' ability to read the problems, form representations, or perform other necessary sub-tasks (De Neys, 2006a; De Neys 2006b).

The ability to load the central executive through maintaining spatial information has been supported by various other studies (Suess, Oberauer, Wittman, Wilhelm, & Schultze, 2002; Klauer, Stegmaier, & Meiser, 1997). Previous work has found that reasoning is most disrupted by tasks that affect central executive functioning (Gilhooly, Logie, Wetherick, & Wynn, 1993; Toms, Morris, & Ward, 1993). Furthermore, this method of spatial loading has been successfully

used in the literature to tap executive resources in reasoning tasks (De Neys, 2006a; De Neys 2006b; & Verschueren, Schaeken, & d'Ydewalle, 2004).

Procedure

Participants were tested in groups of 1-3 in a lab on individual computers. The experiment began by administering the FPPI questionnaire, followed by the automated o-span task. Once they completed both individual differences measures participants were randomly assigned to either the experimental or control condition, and advanced to the reasoning tasks. Participants in the experimental condition saw a 3x3 matrix with dots placed in 4 of the squares for 850 ms before each reasoning problem. After answering the reasoning problem, the participants were shown another 3x3 matrix with each box filled with a number from one to nine. Participants were prompted to recreate the dot configuration previously presented to them by typing the four numbers that represent the location of the dots. In order to become comfortable with the dot matrix task, all participants completed two practice trials prior to testing. Participants in the control condition were only required to answer the 20 problems with no other task impeding their performance.

Results

Relationship between working memory and FPPI

As expected the distribution of working memory capacity scores was relatively normal in shape (see figure 2), with a mean of 44.98 (*SD* = 16.29). Looking at fuzzy processing preference, participants had a mean FPPI score of .43 (*SD* = .26), with the distribution of FPPI scores skewed to the right (see figure 3). These results are comparable to those found by Wolfe and Fisher (2013). Low scores in this measure represent a gist processing preference, and high scores represent a preference to integrate gist and verbatim information. In accordance to fuzzy-trace theory, these results support the idea that the majority of people prefer to reason with gist information.

For conceptual clarity I divided working memory capacity into equal numbers of low, medium and high capacity, and split FPPI scores using a previously established cut-off of .45 (Wolfe & Fisher, 2013). Comparing the individual scores for FPPI with scores from the automated o-span task produced a correlation of r = -.0213, indicating that participants' working memory capacity was independent from their fuzzy processing preference. Both of the measures had high reliability, with the FPPI producing a Cronbach's alpha of .93, and the automated o-span producing a Cronbach's alpha of .78.

Dual-task manipulation

Participants in the dual-task condition had a lower than expected accuracy of 66% when recalling the dot matrixes. Additionally, the dual-task manipulation did not significantly impact the number of conjunction or base rate fallacies committed by the participants ($p > .05$). Likewise, being assigned to either the dual-task or control condition did not significantly interact with level of working memory capacity or FPPI when completing either conjunction, $F(1,162) = 1.095$, $p = .30$ or base rate questions, $F(1,162) = .008$, $p = .93$.

Conjunction Fallacies

As predicted, working memory capacity significantly impacted the number of conjunction fallacies participants made, $F(2,162) = 5.603$, $p = .004$. Subsequent pairwise analyses revealed that participants with high working memory capacity ($M = 3.58$ $SD = 2.16$) had significantly fewer conjunction fallacies than both low ($M = 4.93$ $SD = 2.52$), ($p = .001$) and medium ($M = 4.52$ $SD = 2.52$) ($p = .32$) working memory capacity. Contrary to predictions, fuzzy processing preference was not related to how well participants answered conjunction questions,

$F(1,162) = .894$, $p = .346$. Working memory capacity and fuzzy processing preference did not interact to affect the number of conjunction fallacies, $F(2,162) = 1.418$, $p = .245$.

Base Rate Fallacies

Contrary to expectations, participants with high working memory capacity did significantly worse on estimating base rate probabilities, $F(2,162) = 3.60$, $p = .030$. Further pairwise analyses revealed a significant difference in accuracy on base rate problems between participants with high working memory capacity ($M = 40.90$ $SD = 9.55$) and both low ($M = 36.11$ $SD = 11.20$) (p=.017) and medium ($M = 36.21$ $SD = 10.37$) (p=.035) working memory capacity.

Much like with the conjunction fallacies, FPPI failed to show significant differences in base rate fallacies between high and low FPPI scores, $F(1,162) = .149$, $p = .70$. However, these results were also analyzed with a multiple regression, which found that higher scores on FPPI predicted more base rate fallacies, $\beta = .173$, $t(171) = 2.336$, $p = .021$. In addition, there was no interaction between working memory capacity and fuzzy processing preference, $F(2,162) = .133$, $p = .876$. These results show the opposite pattern than what was expected.

Upon further analysis it was discovered that participants with low working memory capacity tended to employ a strategy in which they reported a probability estimate of 50% more often than participants with high working memory, X^2 (1, $N = 73$) =3.96, p =.047. Similarly, participants with a low FPPI score reported descriptively more 50s than participants with a high FPPI score, X^2 (1, $N = 115$) = 1.47, p =.23. Due to the nature of the base rate problems, many of the answers were smaller than the participants expected, making a guess of a 50% probability a more accurate answer than the mean answer on many questions. Upon excluding participants with estimates of 50 in their answer, the main effect of working memory capacity in base rates disappeared, $F(2,103) = .845$, $p = .433$. However, the means for the different groups of working memory capacity did not reverse in direction, and instead high ($M = 40.64$) and low ($M = 40.14$) working memory capacity were highly comparably while medium working memory capacity differed ($M = 37.11$). This suggests that the high number of 50s reported does address the entire issue with the base rate problems.

Discussion

The present study sought to clarify two questions which had not been previously addressed in the literature. First, what is the relationship between working memory capacity and fuzzy processing preference. Second, how do these two constructs interact, if at all, to produce sound reasoning? I predicted that fuzzy processing preference and working memory were independent, and that normatively correct levels of reasoning ability would emerge from an interaction between high working memory capacity and verbatim processing preference. I found a near-zero correlation between working memory capacity and fuzzy processing preference. In addition, the dual-task manipulation did not impact the number of fallacies in any condition, and there were no interactions between working memory capacity, fuzzy processing preference, and dual-task condition. As expected, high working memory capacity was associated with fewer conjunction fallacies, supporting the idea that working memory capacity and reasoning ability are highly related constructs (Kyllonen & Christal, 1990; Capon et al., 2003**)**. However, working memory capacity was also associated with more base rate fallacies, directly contradicting what previous literature has found on the relationship between working memory capacity and reasoning (Markovits et al., 2002; Süß et al., 2002). Contrary to expectations, fuzzy processing preference had no effect on conjunction fallacies, and actually predicted higher base rate fallacies.

The first major finding was the independence of working memory capacity and fuzzy processing preference. Both measurers were reliable and showed the expected distributions. This indicates that FPPI isn't simply measuring working memory capacity as a predictor for reasoning ability. This dissociation is further supported by the distinct distributions found for fuzzy processing preference and working memory capacity. As expected, working memory capacity was roughly normal (see figure 3), and fuzzy processing preference was skewed to the right (see figure 2). The distribution and low mean (M = .43) of the FPPI scores indicate that most people had a gist processing preference, meaning that the majority of participants relied almost exclusively on the verbal description and gave little weight to the quantitative base rate much like in Wolfe and Fisher (2013). I predicted that high working memory capacity was necessary, but not sufficient for sound reasoning, and that the combination of high working memory capacity and verbatim processing would lead to the fewest fallacies. Surprisingly, there were no interactions between working memory capacity and fuzzy processing preferences.

Results for the dual-task manipulation indicated that participants tended to ignore the dot matrix memory task. After excluding participants who failed to accurately recall the spatial dot configuration at least 75% of the time, participants still had an average accuracy of 66% when recalling the dot configuration. In addition, the added task of maintaining, and later recalling, dot matrixes did not seem to burden executive resources as evidenced by the participants in the dual-task doing equally well on the reasoning task as the participants in the control condition.

A possible explanation for the failed dual-task manipulation can be found in the manner with which participants rehearsed and later recalled the dot configuration. Although this method has been successfully used to load executive resources in previous work (De Neys, 2006a), it is possible that participants who were not just ignoring the dual-task altogether were able to make the task much easier on themselves by encoding the information into 3 separate modalities. They could have encoded the spatial configuration spatially, as intended, verbally by converting the location of the dots into the numbers representing the location (see figure 3), and finally kinesthetically by practicing the hand movement to type out the correct sequence of numbers. By rehearsing information in this way participants could have made the dual-task much easier on themselves, leading to the ineffectiveness of the dual-task (Kane et al., 2004; personal communication with pilot subject).

Results from the conjunction questions replicated previous literature on the strong relationship between working memory capacity and reasoning ability. As we would expect, participants with high levels of working memory capacity committed fewer conjunction fallacies, indicating that working memory capacity supports stronger reasoning. Somewhat surprisingly, though, fuzzy processing preference failed to predict differences in the number of fallacies committed. This may be due to the fact that, in comparison with previous studies that used the same conjunction questions, this study had fewer than expected conjunction fallacies. It is possible that this contributed to a floor effect, making it harder for FPPI to predict differences in the number of fallacies. Another possible source of error for the conjunction fallacies could be the frequency with which participants put the same probability for an answer in the same set of questions. For example in the Linda problem, participants might have rated Linda's likelihood of being a feminist, a bank teller, and being both a feminist and a bank teller all at 40%. While this isn't necessarily a conjunction fallacy, it is not a logically reasonable response and is likely indicative of a low-effort guessing strategy.

The base rate questions proved to be more problematic than the conjunction fallacies. Surprisingly, participants with high working memory capacity did significantly worse on base rate problems. The results from FPPI matched this counterintuitive finding by showing that participants with a higher FPPI score also did worse on the base rate questions. These counterintuitive results may be due a high number of 50% reported for the probability of each base rate question. Although it was not significant in either case, both participants with low working memory capacity and participants with gist processing preference had higher numbers of 50s reported. By putting a 50 as a response to a probability question, participants are simply using a guessing strategy instead of trying to reason through the problem (Wolfe & Reyna, 2010a and personal communication). This would support the finding that both participants with low working memory capacity and participants with a gist processing preference had more 50s than their counterparts. The "guess 50% strategy" was observed by Wolfe & Reyna (2010a) on joint probability problems with nonsense syllables, and they found that interventions designed to improve judgment actually increased fallacies by decreasing use of the suboptimal guess 50% strategy.

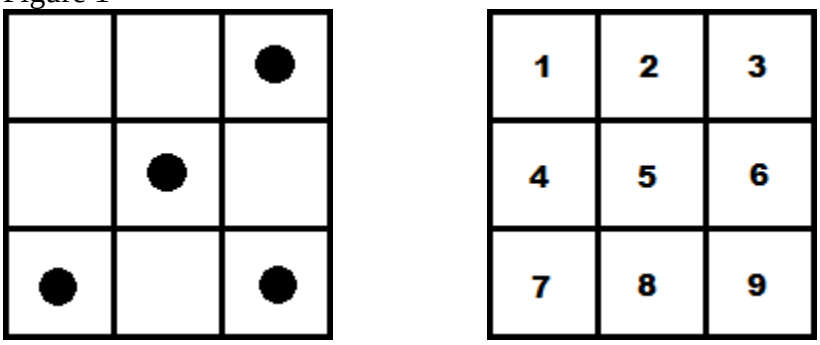
In sum, we found that working memory capacity and fuzzy processing preference were independent, and did not interact to produce higher reasoning. The results from the conjunction fallacies supported previous findings in the working memory capacity literature in that participants with high working memory capacity made fewer conjunction fallacies. Fuzzy processing preference did not predict fewer conjunction fallacies; however I believe this was due in part to a possible floor effect in addition to an "all-the-same" strategy. The results from the base-rate questions largely conflicted with previous literature, showing an unexpected pattern of high working memory and high FPPI score actually increasing the level of fallacies. These surprising results were most likely due to the high numbers of 50s reported. Finally, the dual-task manipulation was unsuccessful in that there were no differences between the dual-task and control group in any of the conditions. This was most likely a result of both low motivation, and ability encode the information spatially, verbally, and procedurally.

In future studies I plan to analyze the relationship between working memory capacity and fuzzy processing preference further with a more tightly controlled study. In order to reduce error that may result from the difficult nature of the base-rate questions, I will only use conjunction fallacies to assess reasoning ability. The guessing strategy seen in the conjunction fallacies could be reduced by simply telling the participants that each probability in a set should be a different number. The possible floor effect in this study will be addressed in the follow-up study by switching the between subjects design of the dual-task to a within subjects design. In this way, each participant will experience a control and a dual-task condition, thereby increasing the number of fallacies in the dual-task for all participants. Additionally, I plan to use a different, more difficult dual-task such as the Sternberg task (Sternberg, 1966) that will restrict the modality with which participants can encode information while still presenting stimuli to participants without needing experimenter intervention. Participant motivation may have also been an issue in the present study. To help motivate participants the follow-up study will include a contract with the participants, making sure they are aware of the importance of their responses. In sum, the follow-up study will have many of the same components of the present study, with a few key differences meant to increase the rate of fallacies, decrease random error from the all-the-same strategy and low motivation, and effectively load executive resources.

The results discussed above paint a surprising, and somewhat conflicted, picture of the relationship between working memory capacity and fuzzy processing process. However, it's

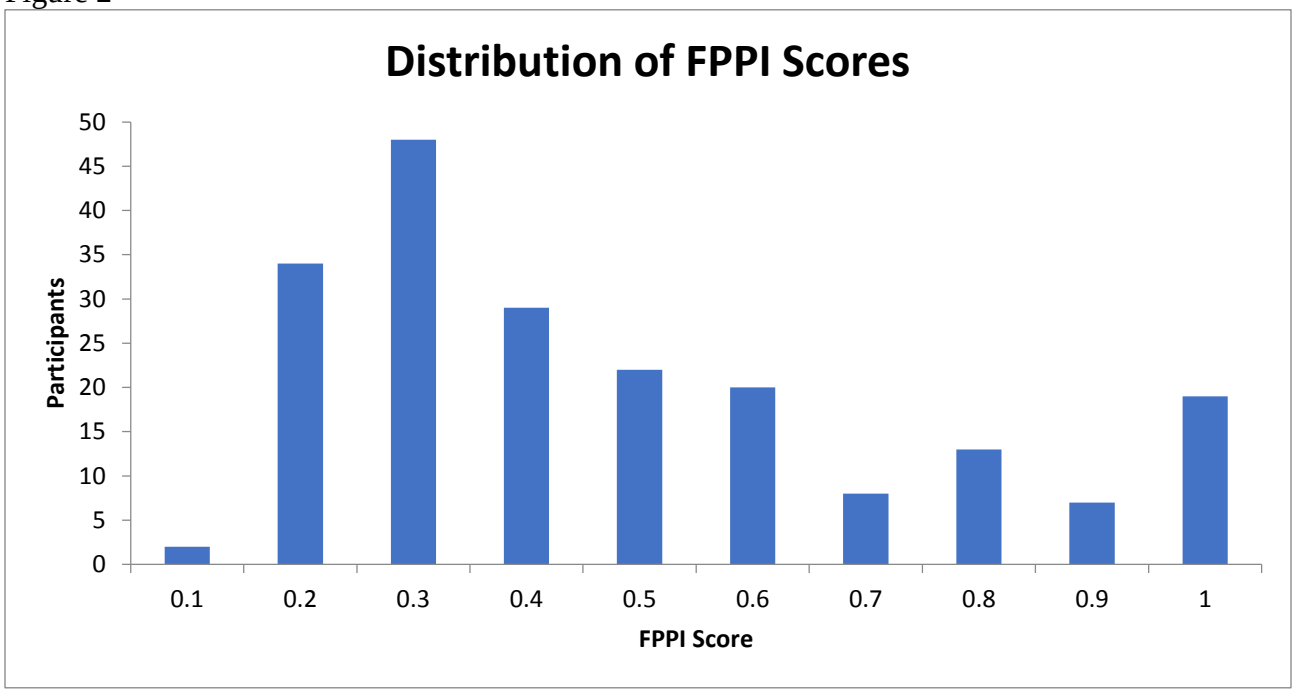
possible that the true relationship between these two constructs was not made clear in this study due to several confounds. The previously described follow-up study will hopefully shed light on the true nature of these predictors of reasoning ability by controlling for the possible confounds discussed. Although not entirely supported here, it is clear from previous research that both working memory capacity and fuzzy processing preference are reliable predictors of reasoning ability. The relationship between working memory capacity, fuzzy processing preference and reasoning ability will require further study.

Figure 1



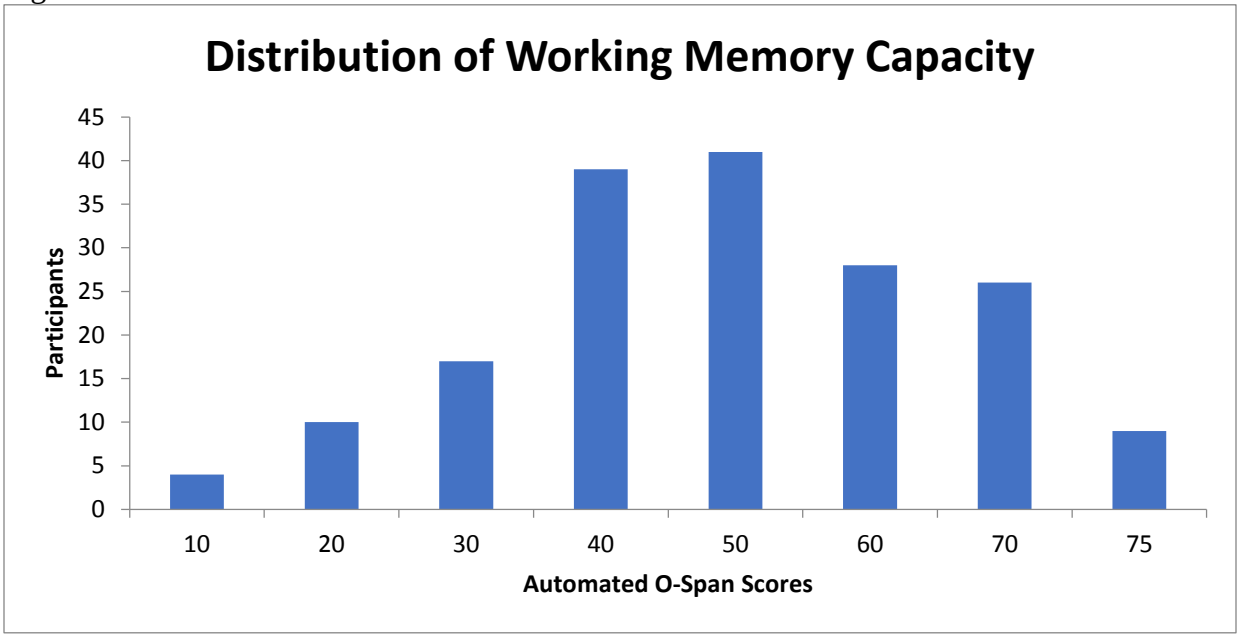
Example dot pattern and response screen for the dual-task.

Figure 2



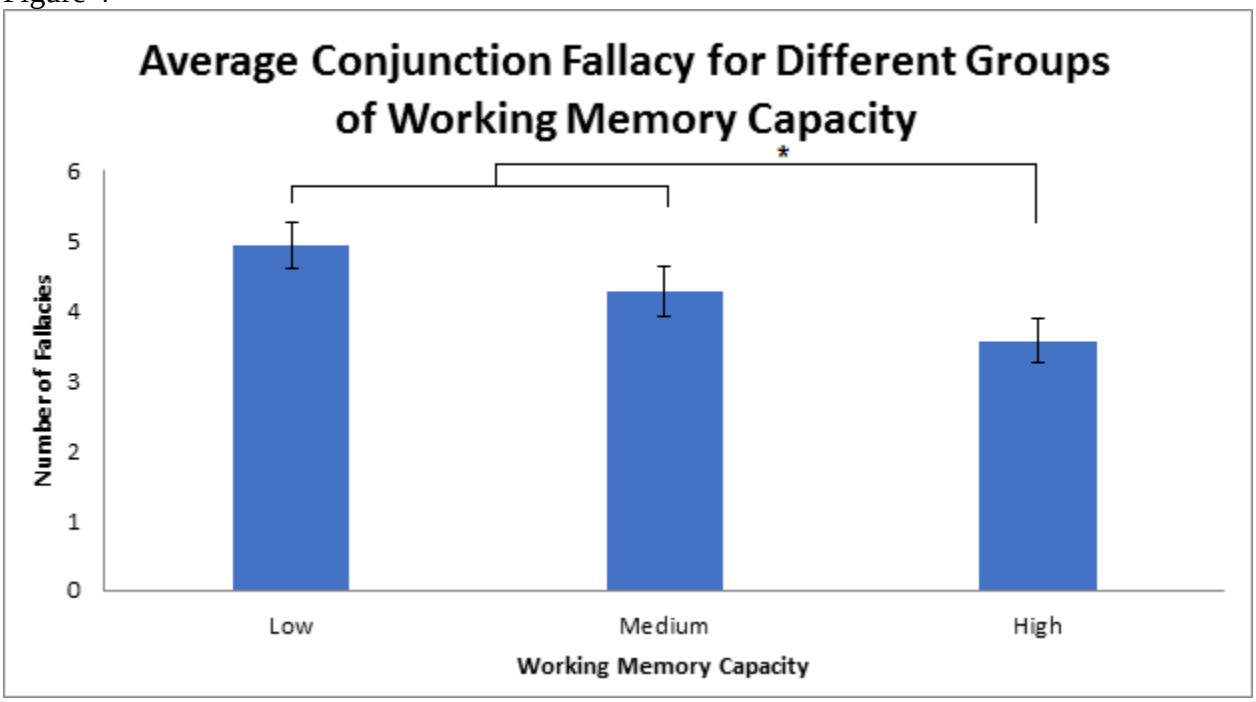
Distribution of FPPI scores with lows scores indicating a gist preference and high scores indicating a preference to incorporate verbatim and gist representations.

Figure 3



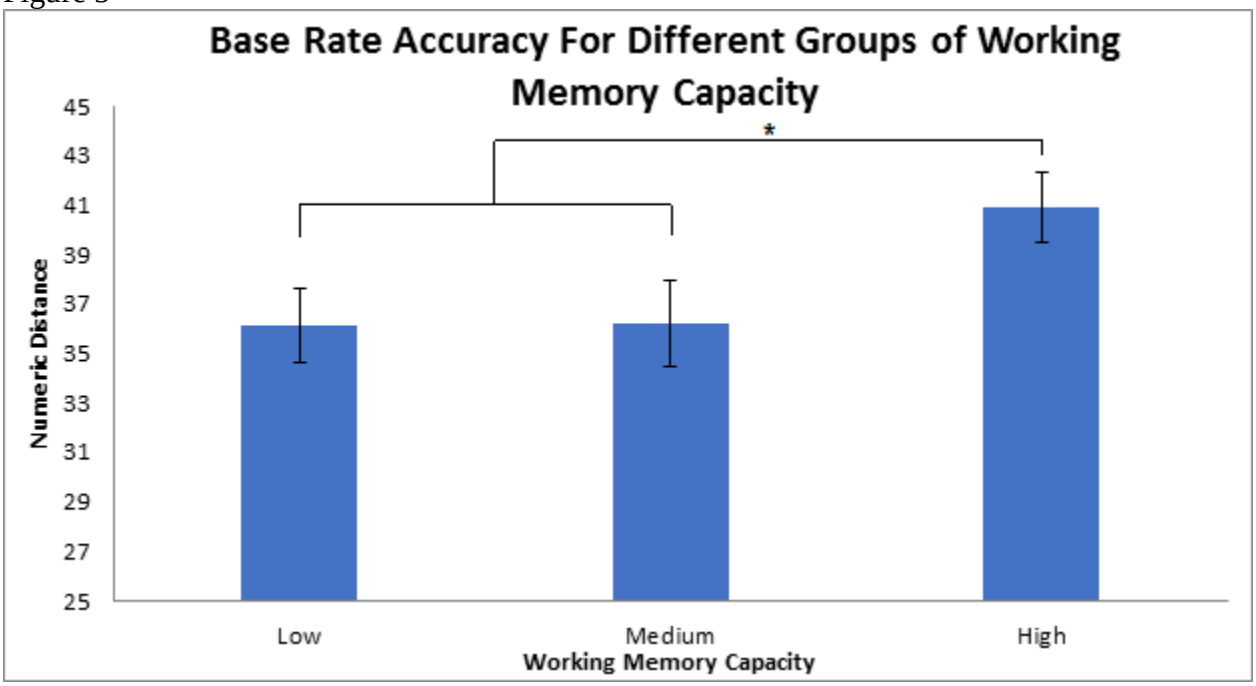
Distribution of working memory capacity as measured by the automated o-span task.

Figure 4



Main effect of working memory capacity for working memory capacity for Conjunction Fallacies. Participants with high working memory capacity committed significantly fewer conjunction fallacies than participants with either low or medium working memory capacity.

Figure 5



Main effect of working memory capacity for base rate questions. Participants with high working memory did significantly worse than participants with both low and medium working memory.

References

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47–90). New York: Academic Press.

Baddeley, A. (1992). Working memory, *255*(5044), 556–559.

Baddely, A. (2007). *Working Memory, Thought and Action*. Oxford, UK: Oxford University Press

Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta psychologica*, *44*(3052), 211–233.

Brainerd, C. J., & Kingma, J. (1984). Do children have to remember to reason? A fuzzy-trace theory of transitivity development. *Developmental Review, 4*(4), 311-377.

Brainerd, C. J., & Reyna, V. F. (2001). Fuzzy-trace theory: dual processes in memory reasoning and Cognitive Neuroscience. *Advances in Child Development and Behavior, 28*, 41–100.

Brainerd, C. J., Reyna, V. F., & Brandse, E. (1995). Are children's false memories more persistent than their true memories? *Psychological Science, 6*, 359–364. doi:10.1111/j.1467–9280.1995.tb00526.x

Capon, A., Handley, S., & Dennis, I. (2003). Working memory and reasoning: An individual differences perspective. *Thinking & Reasoning, 9*(3), 203–244. doi:10.1080/13546781343000222

Chiesi, F., Primi, C., & Morsanyi, K. (2011). Developmental changes in probabilistic reasoning: The role of cognitive capacity, instructions, thinking styles, and relevant knowledge. *Thinking & Reasoning*, *17*(3), 315–350. doi:10.1080/13546783.2011.598401

Cowan, N. (1995). *Attention and memory: An integrated framework.* New York: Oxford University Press.

Cowan, N. (2009). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research, 168*, 323–338. doi:10.1016/S0079-6123(07)00020-9

Conway, A. R., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in cognitive sciences, 7*(12), 547-552.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466.

De Neys, W. (2006a). Automatic-heuristic and executive-analytic processing during reasoning: Chronometric and dual-task considerations. *Quarterly journal of experimental psychology*, *59*(6), 1070–1100. doi:10.1080/02724980543000123

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

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, 128, 309–331

Evans, J. S. B. T. (2002). Logic and human reasoning: An assessment of the deduction paradigm. *Psychological Bulletin*, *128*(6), 978–996. doi:10.1037//0033-2909.128.6.978

Evans, J. S. B. T. (2003). In two minds: dual-process accounts of reasoning. *Trends in Cognitive Sciences, 7*(10), 454–459. doi:10.1016/j.tics.2003.08.012

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual review of psychology*, *59*, 255–78. doi:10.1146/annurev.psych.59.103006.093629

Evans, J. S. B. T., Handley, S. J., Neilens, H., Bacon, A. M., & Over, D. E. (2010). The influence of cognitive ability and instructional set on causal conditional inference. *Quarterly Journal of Experimental Psychology, 63*, 892–909

Fisher, C.R. & Wolfe, C.R. (in press). Are people naive probability theorists? A further examination of the probability theory + variation model. *Journal of Behavioral Decision Making.*

Fukukura, J., Ferguson, M. J., & Fujita, K. (2012). Psychological distance can improve decision making under information overload via gist memory. *Journal of experimental psychology. General*, 1–8. doi:10.1037/a0030730

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

Gilhooly, K. J., Logie, R. H., & Wynn, V. (1999). Syllogistic reasoning tasks, working memory, and skill. *European Journal of Cognitive Psychology*, 11, 473–498.

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.

Halford, G. S., Maybery, M. T., & Bain, J. D. (1986). Capacity limitations in children's reasoning : a dual-task approach. *Child Development*, *57*(3), 616–627.

Hamm. R. M. (1993). Explanations for the common responses to the blue/green cab probabilistic inference word problems. *Psychological Reports, 72,* 219-242.

Hamm, R. M. (1994).Underweighing of base-rate information reflects important difficulties people have with probabilistic inference. *PSYCOLOQUY.* 5

Heuer, R. J. *Psychology of Intelligence Analysis*. Washington, D.C.: Center for the Study of Intelligence, Central Intelligence Agency, 1999. Print.

Johnson-Laird, P. N. (1999). Deductive reasoning. *Annual review of psychology*, *50*(1), 109-135.

Kahneman, D. (2011). *Thinking Fast and Slow*. New York: Macmillan.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, MA: Cambridge University Press.

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working memory capacity. *Journal of Experimental Psychology: General*, *130*, 169–183

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 336–358

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology. General*, *133*(2), 189–217. doi:10.1037/0096-3445.133.2.189

Kardash, C. M., & Scholes, R. J. (1996). Effects of pre-existing beliefs, epistemological beliefs, and need for cognition on interpretation of controversial issues. *Journal of Educational Psychology*, *88*, 260–271.

Klaczynski, P. A. (1997). Bias in adolescents' everyday reasoning and its relationship with intellectual ability, personal theories, and self-serving motivation. *Developmental psychology*, *33*(2), 273–283.

Klaczynski, P. A. (2000). Two-process approach to adolescent cognition motivated scientific reasoning biases, epistemological beliefs, and theory polarization. *Child development*, *71*(5), 1347–1366.

Klaczynski, P. A. (2001). Analytic and heuristic processing influences on adolescent reasoning and decision-making. *Child development*, *72*(3), 844–61.
Klaczynski, P.A. & Gordon, D.H. (1996). Self-serving influences on adolescents' evaluations of belief-relevant evidence. *Journal of Experimental Child Psychology*, *62*(3), 317–39.
Klauer, K. C., Stegmaier, R. & Meiser, T. (1997). Working memory involvement in propositional and spatial reasoning. *Thinking and Reasoning, 3,* 9-47.
Kokis, J. V, Macpherson, R., Toplak, M. E., West, R. F., & Stanovich, K. E. (2002). Heuristic and analytic processing: age trends and associations with cognitive ability and cognitive styles. *Journal of experimental child psychology*, *83*(1), 26–52.
Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, *14*(4), 389–433. doi:10.1016/S0160-2896(05)80012-1
La Pointe, L.B. & Engle, R.W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 1118–1133.
Law, D. J., Morrin, K. A., & Pellegrino, J. W. (1995). Training effects and working memory contributions to skill acquisition in a complex coordination task. *Learning and Individual Differences, 7,* 207-234.
Markovits, H., Doyon, C., & Simoneau, M. (2002). Individual differences in working memory and conditional reasoning with concrete and abstract content. *Thinking & Reasoning*, *8*(2), 97–107. doi:10.1080/13546780143000143
Meiser, T., Klauer, K. C., & Naumer, B. (2001). Propositional reasoning and working memory: The role of prior training and pragmatic content. *Acta Psychologica,* 106, 303–327.
Miyake, A., Friedman, N. P., Rettinger, D. a., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General*, *130*(4), 621–640. doi:10.1037//0096-3445.130.4.621
Moscovitch, M. (1994). Cognitive resources and dual-task interference effects at retrieval in normal people: The role of the frontal lobes and medial temporal cortex. *Neuropsychology, 8,* 524–534.
Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. *Psychological bulletin*, *116*(2), 220.
Reyna, V. E. (1991). Class inclusion, the conjunction fallacy, and other cognitive illusions. *Developmental Review, 11*, 317-336.
Reyna, V. F. (1996). Conceptions of memory development with implications for reasoning and decision making. *Annals of Child Development, 12*, 87−118.
Reyna, V. F. (2008). A theory of medical decision making and health: fuzzy-trace theory. *National Institute of Health, 28*(6), 850–865. doi:10.1177/0272989X08327066.A
Reyna, V. F. (2012). A new intuitionism: Meaning, memory, and development in Fuzzy-Trace Theory. *Judgment and Decision Making*, *7*(3), 332–359.
Reyna, V. F., & Brainerd, C. J. (1990). Gist is the Grist: Fuzzy-trace Theory and the New Intuitionism.
Reyna, V. E. and Brainerd, C. J. (1991). Fuzzy-trace theory and framing effects in choice: Gist extraction, truncation, and conversion. *Journal of Behavioral Decision Making, 4*, 249-262.
Reyna, V. F. and Brainerd, C. J. (1992). A fuzzy-trace theory of reasoning and remembering: paradoxes, patterns, and parallelism, in edited by Healy, A., Kosslyn, S., and Shiffrin, R.

(eds.), *From Learning Processes to Cognitive Processes: Essays in Honor of William K. Estes* (pp. 2: 235–259). Erlbaum, Hillsdale, NJ

Reyna VF, Brainerd CJ. (1995). Fuzzy-trace theory – an interim synthesis. *Learning and Individual Differences*, *7*:1–75.

Reyna, V. F., & Brainerd, C. J. (1994). The origins of probability judgment: A review of data and theories. In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 239−272). New York: Wiley

Reyna, V. F., & Brainerd, C. J. (2008). Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learning and Individual Differences*, *18*(1), 89–107. doi:10.1016/j.lindif.2007.03.011

Reyna, V. F., Brainerd. C. J. and Connolly, T. (1990) Just the bottom line, please: A fuzzy-trace theory of framing effects in choice, paper presented at the 5th International Conference on the Foundation and Application of Utility, Risk, and Decision Theories. Duke University. Durham, NC.

Reyna VF, Casillas W. (2009). Development and dual processes in moral reasoning: A fuzzy-trace theory approach. *Psychology of Learning and Motivation*, *50*, 207–236

Reyna, V. F., & Ellis, S. C. (1994). Fuzzy-trace theory and framing effects in children's risky decision making. *Psychological Science*, *5*(5), 275−279.

Reyna, V. F. and Farley, F. (2006). Risk and rationality in adolescent decision making: implications for theory, practice, and public policy. *Psychological Science in the Public Interest*, 7, 1–44.

Reyna VF, Lloyd FJ. (2006). Physician decision making and cardiac risk: Effects of knowledge, risk perception, risk tolerance, and fuzzy processing. *Journal of Experimental Psychology: Applied*, *12*:179–195.

Stanovich, K. E., Toplak, M. E., & West, R. F. (2008). The development of rational thought: A taxonomy of heuristics and biases. *Advances in Child Development and Behavior*, *36*, 251–285.

Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology*, *89*(2), 342.

Stanovich, K. E., & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General*, *127*(2), 161–188. doi:10.1037//0096-3445.127.2.161

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: implications for the rationality debate? *The Behavioral and brain sciences*, *23*(5), 645–65; discussion 665–726.

Stanovich, K. E., West, R. F., & Toplak, M. E. (2013). Myside bias, rational thinking, and intelligence. *Current Directions in Psychological Science*, *22*(4), 259–264. doi:10.1177/0963721413480174

Sternberg, S. (1966). High speed scanning in human memory. *Science*, *153*, 652-654.

Süß, H. M., Oberauer, K., Wittmann, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability—and a little bit more. *Intelligence*, *30*(3), 261-288.

Suess, H.-M., Oberauer, K., Wittmann, W., Wilhelm, O., & Schultze, R. (2002). Working memory explains reasoning ability – and a little bit more. *Intelligence, 30*, 261-288

Toms, M., Morris, N., & Ward, D. (1993). Working memory and conditional reasoning. *Quarterly Journal of Experimental Psychology, 46*A, 679–699

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? Journal of Memory & Language, 28, 127-154

Tversky, A., & Kahneman, D. (1973). Availability: a heuristic for judging frequency and probability. *Cognitive psychology*, *5*, 207–232.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90, 293–315.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*, 498-505.

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: an examination of simple and complex span and their relation to higher order abilities. *Psychological bulletin*, *133*(6), 1038–66. doi:10.1037/0033-2909.133.6.1038

Verschueren, N., Schaeken, W., & d'Ydewalle, G. (2004). Everyday conditional reasoning with working memory preload. *Proceedings of the Annual Meeting of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Wolfe, C. R. (1995). Information seeking on bayesian conditional probability problems : a fuzzy-trace theory account. *Journal of Behavioral Decision Making*, *8*, 85–109.

Wolfe, C. R. (2012). Individual differences in the "Myside Bias" in reasoning and written argumentation. *Written Communication*, *29*(4), 477–501. doi:10.1177/0741088312457909

Wolfe, C. R., & Fisher, C. R. (2013). Individual differences in base rate neglect: a fuzzy processing preference index. *Learning and Individual Differences*, (513), 1–27.

Wolfe, C. R., & Reyna, V. F. (2010a). Semantic coherence and fallacies in estimating joint probabilities. *Journal of Behavioral Decision Making*, *23*, 203–223. doi:10.1002/bdm

Wolfe, C. R. & Reyna, V. F. (2010b). Assessing semantic coherence and logical fallacies in joint probability estimates. *Behavior Research Methods*, *42*, 366-372.

Wolfe, C. R., Fisher, C. R., & Reyna, V. F. (2012). Semantic coherence and inconsistency in estimating conditional probabilities. *Journal of Behavioral Decision Making*.

Appendix A

Fuzzy Processing Preference Index Items with Gist Base Rates in Parentheses from Wolfe & Fisher (2013)

At Cloverdale High School 10% of the seniors go on to college. Bob is a senior at Cloverdale High. He gets mostly As and Bs in school and is well liked by his teachers. What is the probability that Bob will go to college? (.891)

Researchers in Spain found that 20% of Spanish couples in their early 20s are still together after 3 months. Marco and Racquel, who are both 23 years old, recently initiated a relationship. Both like to go hiking on the weekends. They also have similar preferences in terms music and movies. What is the probability that they will be together after 3 months? (. 685)

At Talia's Restaurant the soup of the day is Cream of Cauliflower. The chef's special recipe calls for plenty of Parmesan cheese. So far, 80% of the customers have described the soup as spicy. Brittany is just about to try the Cream of Cauliflower soup for the first time. What is the probability that Brittany will find the soup very spicy? (.278)

The Heights Tigers football team is playing their last game of the regular season. So far they have won 90% of their games. In their last practice they looked worse than they have all season. The players appear to be unmotivated and the coach sounds pessimistic. What is the probability that the Tigers will win their last game? (.395)

At Fisher Primary School, 15% of the 4th grade boys identify blue as their favorite color. Chris is a 4th grader at Fisher Primary School. He has a blue bicycle, and a blue helmet. What is the probability that Chris' favorite color is blue? (.745)

At the Big Bird Supermarket 20% of the customers have a Big Bird Savings Card. Emma is a smart shopper who is often on the lookout for bargains. What is the probability that Emma has a Big Bird Savings Card? (.727)

At Jay's Auto Sales, 90% of newly hired salespersons are able to meet their sales goal in the first week. During a recent interview at Jay's Auto, Mark was described as nice, but unprepared and lacking in ambition. If Mark is hired, what is the probability that he will meet his weekly sales goal in the first week? (.429)

On the advanced deductive reasoning task, 10% answer at least 9 out of 10 questions correctly. Michelle enjoys solving puzzles and brainteasers. What is the probability that Michelle will correctly answer at least 9 out of 10 questions on the advanced deductive reasoning task? (.695)

Sam attended a speed dating service in which 10% of male participants receive the phone number of at least one female participant. Sam is considered good looking, likes new experiences, and is a natural leader in groups. What is the probability that Sam will receive at least one phone number? (.865)

20% of Jessop University female students spend an average of $100 a month on new clothing. Heather is a Jessop University student who follows fashion trends and has never been seen wearing the same outfit twice. What is the probability that she spends at least $100 a month on new clothing? (.825)

15% of Greek women get married before the age of twenty-five. Matina is a twenty-three years old Greek women and has been in a serious, committed relationship for the past three years. What is the probability that she will be married by the time she turns twenty-five? (. 746)

Frank is an outgoing teen and his family has a phone plan that allows unlimited data and phone calls. 20% of teens on his plan send at least 45 text messages per day. What is the probability that Frank sends at least 45 text messages per day? (.779)

A recent marriage survey found that 20% of newlyweds argue mostly about money. Andrew and Faye have been married for eight months. Faye is the sole breadwinner and they are struggling to make ends meet. What is the probability that their arguments are mostly about money? (.765)

In 2008, 85% of the children in Clayton, KY were considered overweight or obese. Kate is an active 12 years old resident of Clayton, KY. Since she was in first grade she has been on the soccer team and is already talking about colleges where she can get a soccer scholarship. Her favorite class is gym. What is the probability that Kate is overweight? (.197)

During the summer in New Haven, there is a 10% chance that it will rain on any given day. On July 1st in New Haven, it was mostly cloudy in the morning with dark clouds above. What is the probability that it will rain in New Haven at some point during that day? (.710)

WNL collects customer satisfaction information through a mail survey. Based on previous experience, 80% of costumers return their surveys. Holly is a costumer of WNL and is described by her friends as lazy, forgetful, and somewhat disorganized. What is the probability that Holly will complete and return the survey? (.275)

At South Gym, an anonymous survey indicated that 80% of the members have used steroids within the last year. Tony belongs to South Gym and cannot bench press 100 pounds. What is the probability that Toney has used steroids within the last year? (.214)

At the local bar and grill, wine and beer are served. According to the bartender's records, 10% of all drinks sold are beer. Bill is a construction worker that describes himself as a "regular Joe." Bill subscribes to Hotrod magazine. What is the probability that Bill will order a beer? (.779)

A rock band is conducting auditions for a new guitarist using a two-stage process. Based on previous auditions, 80% progress to the second stage. Andrew has been playing guitar for 2 months and has never performed even for friends and family. What is the probability that Andrew will progress to the second stage? (.296)

Appendix B
M-Scale Items from Wolfe & Fisher (2013)

Richard is an avid skier and spends 90% of his vacations skiing. Today he has plane tickets to Aspen, Colorado and has been looking forward to this weekend trip for months. Unfortunately, Richard had a bad accident and both of his legs are broken. What is the probability that Richard will go skiing this weekend?

In Little Rock, Arkansas only 10% of the High School soccer referees are women. Sam has been a High School soccer referee for three years. Sam will not be refereeing this year because Sam is pregnant. What is the probability that Sam is a woman?

The town of Springfield has a nice botanical garden. What makes the garden unique is that it has dozens of wind chimes that make beautiful music in the soft summer breeze. About 70% of visitors to the garden report that they come to listen to the wind chimes. Mrs. Addison, who happens to be deaf, is visiting the botanical garden today. What is the probability that Mrs. Addison is visiting to listen to the wind chimes?

Jessica is on her lunch break. She doesn't have much time so she decided to pop into McDonalds for a quick lunch. About 80% of McDonalds customers order some kind of hamburger. However, Jessica is a vegetarian. What is the probability that Jessica will order a hamburger?

Appendix C
Joint Probability Problems from Fisher and Wolfe (in press)

Steve is 50 years old and has a sedentary lifestyle. He is a movie buff. When he comes home from his job as a computer programmer, he likes to watch movies from his movie collection and eat his favorite ice cream: double fudge, chocolate chip with sprinkles. What is the probability that Steve is overweight? What is the probability that Steve can do 50 pushups? What is the probability that Steve is overweight and can do fifty pushups? What is the probability that Steve is overweight or can do fifty pushups?

Tony is 5'10, 175 pounds and has a membership at Gold's Gym. He works out at least four times a week. What is the probability that Tony can bench press at least 150 pounds? What is the probability that Tony uses steroids? What is the probability that Tony can bench press at least 150 pounds and uses steroids? What is the probability that Tony can bench press at least 150 pounds or uses steroids?

The Texas state Senate consists of 31 senators and each term is four years. Historically, there has been a republican majority in the Texas state senate. What is the probability that the Democrats will take control of the Texas state Senate in the next term? What is the probability that the Texas state Senate will raise taxes during the next term? What is the probability that the Democrats will take control of the Texas state Senate next term and the Texas state Senate will raise taxes during the next term? What is the probability that the Democrats will take control of the Texas state Senate next term or the Texas state Senate will raise taxes during the next term?

Greensburg is a small quaint town in North Carolina with a population of 9,500. People say going to Greensburg is like going back in time. Main Street is lined with cherry blossom trees and many of the buildings are registered historic sites. What is the probability that Wal-Mart will build a store in Greensburg in the next year? What is the probability that Greensburg's Mom and Pop grocery store will go out of business in the next year? What is the probability that Wal-Mart will build a store in Greensburg in the next year and Greensburg's Mom and Pop grocery store will go out of business in the next year? What is the probability that Wal-Mart will build a store in Greensburg in the next year or Greensburg's Mom and Pop grocery store will go out of business in the next year?

Nichole lives in San Diego, California. Its summertime, her favorite time of the year. What is the probability that it will be 85 degrees or hotter in San Diego today? What is the probability that Nichole will turn on her air conditioner today? What is the probability that it will be 85 degrees or hotter today and Nichole will turn on her air conditioner today? What is the probability that it will be 85 degrees or hotter today or Nichole will turn on her air conditioner today?

According to studies, the economy has shown mixed improvement over the past few years. Many economists believe the recession may continue. What is the probability that the US economy will experience above average growth over the next year? What is the probability that the there will be a large-scale flu epidemic in the US during the next year? What is the probability that the US economy will experience above average growth over the next year and there will be a large-scale flu epidemic in the US during the next year? What is the probability that the US economy will experience above average growth over the next year or there will be a large-scale flu epidemic in the US during the next year?

Stephanie and Matt are both 25 years old and have been in a committed relationship together since their sophomore year in college. What is the probability that Stephanie and Matt

will marry in the next two years? What is the probability that Stephanie and Matt are in love? What is the probability that Stephanie and Matt will marry in the next two years and they are in love? What is the probability that Stephanie and Matt will marry in the next two years or they are in love?

Climatologists have become increasingly concerned about climate change. What is the probability that a tsunami will hit the cost of the Far East within the next three years? What is the probability that increased carbon emissions will continue to cause severe global warming? What is the probability that a tsunami will hit the cost of the Far East within the next three years and that increased carbon emissions will continue to cause severe global warming? What is the probability that a tsunami will hit the cost of the Far East within the next three years or that increased carbon emissions will continue to cause severe global warming?

Julie is 45 years old and works in an office. She rarely exercises. What is the probability that Julie can run three miles? What is the probability that Julie is a smoker? What is the probability that Julie can run three miles and is a smoker? What is the probability that Julie can run three miles or is a smoker?

Tom is a simpleton and enjoys the company of animals. He recently got eight small furry pets, who run loose in his apartment. What is the probability that cats are among Tom's pets who run loose in his apartment? What is the probability that mice are among Tom's pets who run loose in his apartment? What is the probability that cats and mice are among Tom's pets who run loose in his apartment? What is the probability that cats or mice are among Tom's pets who run loose in his apartment?

Tim is 41 years old and lives in a rural area. In his spare time, Tim likes to go hunting, fishing and to tractor pulls. What is the probability that Tim likes country music? What is the probability that Tim likes heavy metal? What is the probability that Tim likes country music and likes heavy metal? What is the probability that Tim likes country music or likes heavy metal?

Mary is a psychology student with a very busy schedule. She works part time, is taking 15 credit hours of classes and volunteers regularly. She is trying to find time in her busy schedule to join a discussion group. She is very interested in gender issues and emotion. What is the probability that Mary will join a discussion group about emotion? What is the probability that Mary will join a discussion group about gender issues? What is the probability that Mary will join a discussion group about emotion and a discussion group about gender issues? What is the probability that Mary will join a discussion group about emotion or a discussion group about gender issues?

Appendix D
Base Rate Problems

There has been a hit and run accident at night involving a taxi cab. The only eyewitness identified the taxi cab as being blue. There are two taxi companies operating in the city. One company has green cabs which make up 85% of the total cabs in the city. The other company has blue cabs, comprising the remaining 15% of taxis. The eyewitness was tested by the court under similar night time conditions and was able to accurately distinguish the color of cab 80% of the time. What is the likelihood that the eyewitness correctly identified the cab as blue? (Kahneman and Tversky 1973).

A study was done on causes of suicide among young adults (aged 25 to 35). It was found that the percentage of suicides is three times larger among single people than among married people. In this age group, 80% are married and 20% are single. Of 100 cases of suicide among people aged 25 to 35, how many would you estimate were single? (Bar-Hillel, 1980)

Studies of dreaming have shown that 80% of adults of both sexes report that they dream, if only occasionally, whereas 20% claim they do not remember ever dreaming. Accordingly, people are classified by dream investigators as 'Dreamers' or 'Nondreamers'. In close to 70% of all married couples, husband and wife share the same classification, i.e., both are Dreamers or both are Nondreamers, whereas slightly more than 30% of couples are made up of one Dreamer and one Nondreamer. Mrs. X is a Nondreamer. What do you think are the chances that her husband is also a Nondreamer? ? (Bar-Hillel, 1980)

During the Vietnam War, a fighter plane made a non-fatal strafing attack on a US aerial reconnaissance mission at twilight. Both Cambodian and Vietnamese jets operate in the area. The US pilot identified the fighter as Cambodian. The pilot's aircraft recognition capabilities were tested under appropriate visibility and flight conditions. When presented with a sample of fighters (half with Vietnamese markings and half with Cambodian) the pilot made correct identifications 80 percent of the time and erred 20 percent of the time. 85 percent of the jet fighters in that area are Vietnamese; 15 percent are Cambodian. What is the probability that the fighter was Cambodian rather than Vietnamese? (Heuer, 1999)

In a city of 1,000 inhabitants there are 100 spies and 900 non-spies. In an attempt to catch the spies, the city installs an alarm system with a surveillance camera and automatic facial recognition software. The software isn't perfect though. If the camera scans a terrorist, a bell will ring 99% of the time, and it will fail to ring 1% of the time. If the camera scans a non-terrorist, a bell will not ring 99% of the time, but it will ring 1% of the time. What is the probability that the person triggering the alarm is a terrorist?

A particular security dog at an airport searches for drugs in the baggage. On average, 5% of the baggage at this airport contains drugs. This dog is pretty good at his job, and only makes a mistake 1% of the time. Of course, there are two types of mistakes this dog can make. He can miss the drugs when there are some, or he can falsely signal that drugs are present when they are not. If one of the security officers at the airport randomly sampled one of the bags the dog has indicated contained drugs, what is the probability it would actually contain drugs?

A cashier at a grocery store has to mark $50 and $100 dollar bills with a marker than shows if they are fake or not. Only 1% of the $50 and $100 dollar bills are fake. This marker is accurate 99% of the time, meaning it marks real bills as fake 1% of the time, and fake bills as real 1% of the time. What is the probability that of the bills marked as fake is actually fake?

Annie's doctor tells her that she has tested positive for a certain disease. 10% of the population has the disease. The test she took has a 95% accuracy rate, meaning it will tell 5% of people with the disease that they don't have it, and it will tell 5% of people without the disease that they do have it. What is the probability that Annie has the disease?