This is to certify that the

dissertation entitled

THE ROLE OF MENTAL SIMULATION IN CAUSAL DECISION MAKING

presented by

Keith E. Niedermeier

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Psychology

Major professor

Date 4/26/99

THE ROLE OF MENTAL SIMULATION IN CAUSAL DECISION MAKING

By

Keith E. Niedermeier

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

1999

ABSTRACT

THE ROLE OF MENTAL SIMULATION IN CAUSAL DECISION MAKING

By

Keith E. Niedermeier

The theoretical importance of alternative hypotheses in causal decision making has been recognized by several scholars (Dougherty, Gettys, & Thomas, 1997; Einhorn & Hogarth, 1986; Pennington & Hastie, 1992, 1993). However, empirical work on the subject to date paints an incomplete picture. The purpose of the present investigation was to explore how the presence and properties of alternative hypotheses can affect belief in a primary hypothesis. Four experiments were conducted in which participants acted as jurors in a case in which a bus company was being sued for killing a dog. Across experiments, participants generally ruled against the defendant less when they were able to mentally simulate an alternative explanation than when simulation was more difficult. In other words, the ability to simulate an alternative hypothesis reduced belief in the primary hypothesis.

This dissertation is dedicated to the many great teachers who have influenced me over the years. Foremost among them are Irwin Horowitz, Norbert Kerr, James Livecchi, Allen McConnell, Lawrence Messé, Alexander von Eye, Kipling Williams, and my parents. I further dedicate this work to Michelle Niedermeier.

ACKNOWLEDGMENTS

I would first like to thank my advisor and the chair of my dissertation committee, Dr. Norbert Kerr. Dr. Kerr has provided unparalleled direction and support throughout my graduate career. Another person worthy of special thanks is Dr. Lawrence Messé, who played a crucial role in this line of research and often served as a surrogate advisor to me. I would also like to thank Dr. Bruce Burns, Dr. Thomas Carr, and Dr. Allen McConnell, all of whom served on my dissertation committee. Additionally, Dr. Irwin Horowitz at Oregon State University provided needed input and support. Dr. Gwen Wittenbaum of the Department of Communications and Dr. Robert Hymes of the University of Michigan-Dearborn assisted with elements of this paper. Also, Kelly Sandor deserves recognition for her assistance with data collection.

The following deserve special acknowledgment: Apple Computer, Inc. for making products that work easily and seemlessly, the Michigan State University Basketball and Hockey teams for advancing to the final four of their respective NCAA tournaments and providing a needed uplifting diversion, Pennsylvania State University for hiring me early and eliminating my job anxiety, my students for showing me why I want to do this in the first place, Tala for being there, and Michelle for everything.

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

Everyday we make judgments about the causes of events. Often we are presented with a possible or likely cause and we must decide if that explanation is, in fact, valid. For instance, jurors are routinely given the primary hypothesis that a defendant is responsible for a certain negative outcome. The jurors must then decide whether to accept this hypothesis as true or not. Much of the existing research on this type of decision making has focused on the amount of evidence or the likelihood that the primary hypothesis is true. For example, a variety of juror decision making models focus on the amount and the weight given to the evidence (see Hastie, 1993, for a review). However, much less attention has been paid to the effects of alternative hypotheses on the perceived validity of a primary hypothesis. The purpose of the present investigation was to explore how the presence and properties of alternative hypotheses can affect belief in a primary hypothesis.

Einhorn and Hogarth (1986), in their review of the literature on judging causality, discussed the importance of alternative explanations for the cause of a particular outcome. Einhorn and Hogarth stated that the strength of a primary explanation is discounted by the presence of specific alternative explanations. Furthermore, they proposed that the power of a specific alternative explanation lies in the potential for it to replace the primary explanation. Although

Einhorn and Hogarth emphasized the importance of alternative explanations in the process of causal reasoning and decision making, they did not offer any empirical data to support their notion.

Pennington and Hastie's (1992, 1993) story model represents another attempt to integrate alternative hypotheses into a model of causal reasoning. According to the model, jurors initially encode evidence and implications in long term memory. Pennington and Hastie state that the evidence is stored in a meaningful structure that represents a coherent, narrative summary of the testimony. This coherent structure is referred to as the story. In the second stage, jurors learn the verdict categories from the judge's instructions. Finally, the juror matches the story to the category of best fit. The degree to which the story is complete, matches a verdict category, and is coherent yields a degree of confidence in the verdict.

Research by Pennington and Hastie (1986, 1988, 1992) has generally supported the story model. For instance, the researchers used extensive interviewing techniques to establish that jurors were, in fact, creating narrative structures of the evidence. In their studies, jurors who reached similar verdicts possessed similar narrative accounts of the evidence. Furthermore, jurors incorporated a substantial amount of implied information into their stories. The researchers also demonstrated that story summaries were the key factor in reaching a verdict by varying the order of

evidence. Changing the order of evidence presumably altered the construction of the jurors' stories and ultimately the verdicts.

Pennington and Hastie (1992, 1993) state that multiple stories can be constructed, but one story will usually be accepted as the best story. The emergence of the "best" story is based on the principles of coverage, coherence, and uniqueness. Coverage is the extent to which a story accounts for all of the evidence. Coherence is the extent to which a story is consistent, plausible, and complete. If only one story is judged to be coherent, the story is said to be unique. It is the concept of uniqueness that is of most interest to the present investigation.

Pennington and Hastie (1992, 1993) state that the presence of more than one coherent story will lessen the belief in any one particular story. However, this idea is given little attention in the various tests of this perspective. It is assumed that one story will emerge as the best and the other stories merely will be dismissed. It is important to note that no explicit basis for alternative stories are given in the trials in the cases used in Pennington and Hastie's (1992) research. In one case, for example, a defendant was charged with embezzlement and there were various pieces of incriminating and exonerating evidence presented, but there was no explicit basis by which to create an alternative story that could account for the missing money. In other words, all of the evidence was either

incriminating or exonerating to the defendant. No evidence introduced specific alternatives to the primary hypothesis that the defendant embezzled the money.

In real court cases, explicit alternative hypotheses often are introduced by the defense. For instance, the defense attorney in a murder case may demonstrate that a specific person other than the defendant had the motive and ability to commit the murder in question. In a civil case, an attorney may make the case that someone other than his client, like the plaintiff or a specific third party, is responsible for the outcome that led to the law suit. In both examples, there is a specific basis for creating an alternative story or hypothesis.

Empirical Investigations of Alternative Explanations

Most of the previous research on alternative hypotheses has focused on the effect of multiple causal routes on the perceived likelihood of different outcomes. Much of this research has demonstrated that simulating many causal routes to a number of different outcomes reduces likelihood estimates of any particular outcome (Hirt & Markman, 1995; Koehler, 1994). The focus of this paper, however, is on the effect of multiple causal routes to a single outcome. In other words, what is the effect of simulating alternatives to the primary cause (hypothesis) of a given outcome? Unfortunately, there have been few empirical tests of the effects of alternative explanations on belief in a primary hypothesis.

In a few studies, strength of belief in a given hypothesis was found to be independent of belief in competing hypotheses (Robinson & Hastie, 1985; Robinson Van Wallendael, 1989). In these investigations, subjects read murder mysteries in which various incriminating and exonerating clues were introduced about a primary suspect. At various points, one or more additional suspects were introduced. Across a number of experiments, the introduction of alternative hypotheses in the form of additional viable suspects did not reliably decrease the strength of belief in a primary hypothesis.

In contrast, other research has found that the strength of belief in a given explanation is directly related to the number of available alternative explanations (Dougherty, Gettys, & Thomas, 1997; Fischhoff, Slovic, & Lichtenstein, 1978; McDonald, 1998). For instance, McDonald presented subjects with an unusual event and provided them with varying numbers of equally coherent explanations for the cause. The average probability assigned to a specified hypothesis approximated the proportion of the number of hypotheses presented. For example, if subjects were presented with three possible explanations for an event, the average probability assigned to a specified hypothesis was approximately 33%. Similarly, if subjects were presented with five possible explanations, the average probability assigned to the same specified hypothesis was approximately 20%.

In a similar investigation, Fischhoff et al. (1978) presented subjects with organized lists of possible causes for a car that failed to start. Omitting possible causes from the lists increased estimates of probability given to the remaining explanations, despite the presence of a category labeled "All other problems." The researchers concluded that increasing the number of specific alternative explanations decreased the estimated probability of any given explanation being the cause for the car's failure to start. Furthermore, they stated that the subjects were quite insensitive to possible explanations that were not specifically listed.

In a more recent investigation, Dougherty et al. (1997) had subjects read short vignettes in which the cause of an outcome was unclear. For instance, one vignette described a firefighter who was killed on duty. Subjects were then asked to judge the probability of a given cause (e.g., smoke inhalation) and list all the thoughts they had while making the probability judgment. The results showed that subjects who considered multiple causes reported lower probabilities for the given cause of an outcome.

In the aforementioned studies, the primary dependent variable was subjective probability, that is, subjects' estimate of the likelihood that a given explanation was "the" cause of the event. Although Robinson and Hastie (1985) and Robinson Van Wallendael (1989) found no differences in subjective probabilities based on the number of alternative explanations, Fischhoff et al. (1978) and McDonald (1998)

did. A possible reason for this inconsistency in findings is that the former studies specified a primary hypothesis, while the latter studies did not. In the latter studies, various numbers of explanations for particular events were all introduced simultaneously and none was presented as the primary hypothesis. In the former studies, the evidence pertained to a primary suspect and alternatives were introduced later.

The Dougherty et al. (1997) study was different in that it was unclear whether one explanation was primary. Subjects were asked to judge the probability of a given explanation only after reading the vignette. The researchers assumed that the given explanation (such as smoke inhalation) was primary, but it is possible that subjects had another hypothesis in mind, making the judged explanation actually the alternative. In any event, it is unclear if a primary hypothesis existed for all subjects in this study.

Collectively, these studies suggest that the presentation of one hypothesis as the primary hypothesis may play an important role in the decision making process. If we have no reason to consider one hypothesis as primary, considering all available alternatives as more or less equally likely is not an unreasonable strategy (Pennington & Hastie, 1992, 1993). However, when we are presented with a primary hypothesis, we are likely to consider that hypothesis and retain it as the "default" explanation unless it can be

replaced or discounted by a plausible alternative (Einhorn & Hogarth, 1986).

Primary Hypotheses

Given this distinction between primary and alternative hypotheses, what psychologically defines a hypothesis as primary becomes an important question. In many situations, one hypothesis may be overtly presented as primary. As mentioned before, jurors are routinely given the primary hypothesis that a defendant is the cause of a certain negative outcome.

Alternatively, a situation may be structured such that one explanation is presented prior to other possible explanations, as in the Robinson and Hastie (1985) and Robinson Van Wallendael (1989) studies. In these studies, one suspect in a murder mystery was introduced and only later were additional suspects revealed. In these situation, the explanation presented first may be considered primary because a story is constructed around this explanation and retained in memory (Einhorn & Hogarth, 1986). Unless this explanation is replaced or discounted later, it should ultimately be accepted.

Furthermore, the mere mention of a single possibility may be enough for it to be considered a primary hypothesis. This line of reasoning is consistent with the conversational convention literature, which suggests that things are not generally stated in conversation unless they are relevant and important. Therefore, the mention of one causal hypothesis in

a scenario may serve as cue that it should be primary (Hilton, 1995; Schwarz, 1994).

Another possibility is that one explanation is considered primary because of previous experience. Imagine, for example, that a mother walks into the kitchen and finds that the vase that normally sits on the table is broken. There are numerous possible hypotheses for the broken vase, but if the mother has a six-year old with a history of playing ball in the house, this causal schema may be the most accessible cognitively and thus emerge as a primary hypothesis.

More generally, people may naturally construct one primary hypothesis for a given outcome (Dougherty et al., 1997). One common example of this type of process is the hindsight bias in which people overestimate the predictability of an outcome (Fischhoff, 1975; Hawkins & Hastie, 1990). This bias may occur because of the tendency to focus on one causal route that occurred and the failure to consider alternative routes. The confirmation bias is another example of this tendency. This bias is the inclination to seek out evidence that supports a given hypothesis (Wason, 1960). Similarly, the confirmation bias may occur because of the tendency to focus on one primary hypothesis and the failure to consider alternatives. A relevant condition in which people do seek out disconfirming information is when alternative causal hypotheses are made salient (Doherty,

Chadwick, Garavan, Barr, & Mynatt, 1996; Jou, Shanteau, & Harris, 1996).

Continuous versus Categorical Measures

As previously mentioned, the primary dependent variable in most studies of alternative hypotheses has been subjective probability estimates. Subjective probabilities are a continuous measure of people's estimates of the likelihood that an event occurred in a given manner. However, people often are required to make categorical decisions about cause. For instance, jurors must ultimately decide if the defendant is guilty or not guilty, regardless of subjective probability. In an investigation of jurors' use of statistical evidence, Wells (1992) collected both continuous and categorical measures in a study that presented participants with a primary hypothesis.

In Wells's (1992) study, participants read about a civil trial in which Mrs. Prob was suing the Blue Bus Company for running over her dog. Participants were informed that Mrs. Prob was walking her dog along a county road when a swerving bus hit and killed her dog. Testimony revealed that only two bus companies, the Blue Bus Company and the Gray Bus Company, operated in the area. Both companies used the road to run empty buses back to their respective terminals.

Unfortunately, Mrs. Prob was color blind so she could not identify the color of the bus. In a condition termed the tire-tracks version, a county transportation official testified that tire tracks at the scene were compared to all

10 of the buses owned by the Blue Bus Company and all 10 of the buses owned by the Gray Bus Company. The tracks matched 80% of the blue buses and 20% of the gray buses. Participants reported continuous subjective probabilities and categorical verdicts indicating whether or not they would find the Blue Bus Company liable.

Like the murder mystery studies (Robinson & Hastie, 1985; Robinson Van Wallendael, 1989), a primary hypothesis was implicit in that the Blue Bus Company was being sued by Mrs. Prob. In the tire-tracks version described above, participants reported than it was near 80% likely that a blue bus was responsible for killing the dog. The participants were instructed to follow the preponderance of evidence decision rule and thus should have ruled against the Blue Bus Company. However, Wells (1992) found that fewer than 20% of the participants ruled against the Blue Bus company in this condition.

In another condition termed the tire-track-belief version, the transportation official testified that tire tracks at the scene were compared to all 10 of the buses owned by the Blue Bus Company and all 10 of the buses owned by the Gray Bus Company using a method that is correct 80% of the time. Based on the results of this matching technique, the official testified that he believed that a blue bus killed Mrs. Prob's dog. Just as in the tire-tracks version, participants reported that it was nearly 80% likely that a Blue Bus Company bus killed Mrs. Prob's dog. In this latter

condition, however, participants ruled against the Blue Bus Company nearly 70% of the time.

The Wells (1992) study illustrates the important point that subjective probabilities do not necessarily translate into functionally equivalent categorical choices. The tire-tracks and tire-track-belief versions produced similar subjective probabilities, a pattern that is consistent with the Robinson and Hastie (1985) and Robinson Van Wallendael (1989) studies, which also presented participants with a primary hypothesis. However, the two conditions produced very different categorical verdicts indicating that continuous and categorical measures do not necessarily produce parallel results, at least in some conditions.

Furthermore, Wells (1992) reasoned that this pattern disconfirmed the family of threshold models of decision making. These models (e.g., Dane, 1985; Kaplan, 1977) view continuous subjective probability as driving categorical verdicts. Generally, these models maintain that an initial prior probability is updated by entering each piece of new information into a mental equation that continually revises subjective probability of guilt. This process occurs over and over until there is no new evidence to consider. The final subjective probability is then compared to a decision criterion or threshold. If subjective probability surpasses the decision threshold, the juror will rule against the defendant.

Wells (1992) stated that the mock jurors is his study were all operating under the same decision rule (preponderance of evidence), so they should have had the same threshold. Given the same level of subjective probability but the very different rates of verdicts against the Blue Bus Company in the two versions of the case, Wells reasoned that the data disconfirmed these types of threshold models. This view of threshold models further suggests that continuous and categorical measures do not necessarily produce parallel results.

Additionally, Wells's paradigm nicely lends itself to the question at hand because there is a clear primary hypothesis (a blue bus killed the dog), and there is a clear alternative hypothesis (a gray bus killed the dog). Also, both subjective probabilities and categorical decisions can be examined while holding actual probabilities constant.

Consideration of Alternatives: The Role of Mental Simulation

Wells (1992) investigated a number of explanations for the very different verdicts against the Blue Bus Company in the two different versions of the case. He ruled out a number of explanations and proposed fact-to-evidence reasoning as the operative mechanism. The fact-to-evidence explanation states that, "in order for evidence to have a significant impact on people's verdict preferences, one's hypothetical belief about the ultimate fact must affect one's belief about the evidence" (p. 746). In other words, if the evidence presented at the trial can be true regardless of what

actually happened, then that evidence is seen as non-diagnostic. For instance, the evidence in the tire-tracks version stated that the tire tracks matched 80% of the blue buses and 20% of the gray buses. This evidence could be accepted as true regardless of which bus actually killed the dog. Thus, such evidence fails the fact-to-evidence criterion and should not have a significant impact on the verdict.

The evidence in the tire-track-belief condition, however, stated that the official believed a Blue Bus Company bus killed the dog based an test that is reliable 80% of the time. The official is correct in his belief if a blue bus killed the dog, but would be incorrect if it were a gray bus that did so. In this version of the case, which bus company's bus actually killed the dog (the fact) bears on the validity of the witness's testimony (the evidence). Because the witness's testimony is true only if a blue bus killed the dog, this evidence passes the fact-to-evidence test and affects the verdict

In contrast to the fact-to-evidence reasoning explanation, Niedermeier, Kerr, and Messé (1999) suggested that mock jurors' ability to mentally simulate a situation in which the Blue Bus Company was <u>not</u> guilty (i.e., an alternative hypothesis) influenced their willingness to rule against the defendant. For instance, in the tire-tracks version, the matching technique produced a complete match with 8 of 10 Blue Bus Company buses and 2 of 10 Gray Bus Company buses. In this version, there indisputably are two

concrete gray buses that could have plausibly killed the dog. Because it is easy to imagine that one of these two gray buses may have killed the dog, jurors may be more reluctant to hold the Blue Bus Company responsible. In the tire-tracks-belief version, however, the transportation official testified that he believed that a blue bus killed the dog based on an 80%-reliable matching technique. This scenario requires a much more complex cognitive mutation to plausibly imagine that a gray bus killed the dog. Although this testimony could create some doubt about a blue bus's guilt, it offers no explicit basis to create a story consistent with a gray bus's guilt. Thus, mock jurors were more willing to rule against the Blue Bus Company.

Niedermeier et al. (1999) tested this ease-of-simulation explanation by altering the critical testimony in the case used by Wells (1992). In a condition termed the complete-match version, the transportation official testified that tire tracks from the scene of the accident completely matched 8 of 10 blue buses and 2 of 10 gray buses, as in Wells tire-tracks version. In another condition termed the partial-match version, the official testified that the tire tracks partially matched one blue bus and one gray bus, and test results indicated an 80% probability that the blue bus killed the dog and a 20% probability that the gray bus was the one.

In both the complete-match and the partial-match conditions, mock jurors reported that it was about 80% likely that a blue bus killed the dog. However, jurors were much

less willing to hold the Blue Bus Company responsible for killing the dog in the complete-match condition (25%) than in the partial-match condition (47%). In the complete-match condition, there are two gray buses that completely matched the tire tracks and could have killed the dog. In this condition it is not difficult to imagine that one of these gray buses killed the dog. In the language of the story model (Pennington & Hastie, 1992, 1993), the possibility of a blue bus being responsible lacks uniqueness because the possibility of a gray bus being responsible has coherence. In the partial-match condition, however, there is only a vague 20% probability match to a single gray bus. In this latter condition, it is more difficult to simulate a situation in which a gray bus could have killed the dog. The possibility of a blue bus being responsible has uniqueness because the possibility of a gray bus being responsible lacks coherence.

In fact, mock jurors reported that it was significantly easier to imagine a scenario in which a gray bus killed the dog in the complete-match condition than in the partial-match condition. Thus, jurors' ability to mentally simulate an alternative hypothesis, that is, a gray bus being responsible for killing the dog, significantly affected jurors' willingness to rule on the primary hypothesis, that is, the liability of the Blue Bus Company. Using a very similar methodology to Niedermeier et al. (1999), Sykes and Johnson (in press) found comparable patterns of verdicts and reached a similar conclusion.

It has been suggested that the ease of mental simulation of alternatives may be used as a heuristic to determine the validity of a given hypothesis (McDonald, 1998; Miller, Turnbull, & McFarland, 1989). In a series of studies, Miller, et al. demonstrated that the more easily people could imagine something occurring by one route, the less suspicious they were that it occurred by some other route. Miller et al. presented subjects with a scenario in which a child takes only chocolate chip cookies from a cookie jar while leaving oatmeal cookies to go stale. To avoid stale cookies, the child's mother instructs him to close his eyes before reaching into the cookie jar and take whichever cookie he grabs. The child agrees to the method, goes to the kitchen and returns with a coveted chocolate chip cookie. Subjects were asked to rate their own suspicion that the child peeked. In one condition, subjects were informed that there was 1 chocolate chip cookie and 19 oatmeal cookies in the jar. In another condition, subjects were informed that there were 10 chocolate chip cookies and 190 oatmeal cookies.

Subjects expressed greater suspicion when there was only 1 chocolate chip cookie than when there were 10, despite the fact that drawing a chocolate chip cookie at random was equally likely in both conditions. When there was only one chocolate chip cookie, there was only one way in which the outcome could have occurred by chance. In contrast, it was easier to imagine that the outcome had occurred by chance when there were 10 chocolate chip cookies. The researchers

concluded that the easier it was simulate that an outcome occurred in one manner (chance), the less suspicious subjects were that it occurred in another manner (peeking).

The Miller et al. (1989) study demonstrates that the relative frequency of alternatives may play a role in the decision making process. When presented with the primary hypothesis that the child is peeking, participants were less suspicious of the infraction when there were relatively more alternative possibilities than when there were fewer alternatives. Although the researchers suggested that it was easier to imagine drawing a chocolate chip cookie when there were 10 such cookies than when there was one, they provided no direct evidence of ease of simulation.

General Objectives

To summarize, the theoretical importance of alternative hypotheses in decision making has been recognized by several scholars (Dougherty et al., 1997; Einhorn & Hogarth, 1986; Pennington & Hastie, 1992, 1993). However, empirical work on the subject to date paints an incomplete picture. The current investigation sought to address four issues related to the role of mental simulation of alternative hypotheses in causal decision making.

First, the notion that mental simulation plays a role in casual decision making was further tested by attempting to make alternative hypotheses more or less easily simulated. Second, the importance of categorically eliminating potential alternatives was investigated by altering the order in which

participants considered various alternative explanations. Third, the relative frequency of alternative possibilities were altered to determine if one coherent alternative is sufficient to reduce belief in the primary hypothesis as opposed to several alternatives. Lastly, the heuristic nature of mental simulation was investigated by manipulating participants' motivation to consider alternatives in order to determine if mental simulation is heuristic or a systematic but intrinsically flawed method of reasoning. In all experiments, both categorical and continuous measures were evaluated.

Experiment 1

Using Wells's (1992) paradigm, it was demonstrated that the ability to mentally simulate alternatives to a primary hypothesis reduced jurors' willingness to believe the primary hypothesis (Niedermeier et al., 1999; Sykes & Johnson, in press). As previously described, jurors in the Niedermeier et al. study were much less willing to hold the Blue Bus Company responsible for killing a dog when it was easy to simulate a gray bus doing so than when it was more to difficult to simulate, as in the partial-match version. Again, there was only a vague 20% match to a single gray bus in the partial-match condition compared to full matches to two gray buses in the complete-match condition.

In Experiment 3 of the Niedermeier et al. (1999) study, one group of mock jurors read the partial-match version of

the blue bus case while another group read the partial-match case with additional testimony in which the defense attorney noted that the accident scene was not closed off to traffic. The attorney asked if a second bus could have come by after the accident and obscured the tracks of the responsible bus, leaving its own tracks in place of the responsible bus's tracks, thus making the innocent bus look responsible. The transportation official, who was testifying about the tire tracks, acknowledged that this was possible, but stated that there was no indication of a second set of tracks on the dog.

The scenario proposed by the defense attorney provided a means for simulating a situation in which a blue bus appears responsible when another bus actually killed the dog, without adding any substantive evidence that changed actual (or subjective) probabilities. However, when jurors heard the scenario posited by the defense attorney, they were much less willing to rule against the Blue Bus Company and reported that it was easier to think of a way that a bus from the Gray Bus Company could have been responsible for killing the dog than when they heard the partial-match case alone.

The data from Niedermeier et al. (1999, Experiment 3) suggest that, given a framework to simulate a coherent alternative hypothesis, people will be less willing to rule against the defendant. It follows that anything that will increase jurors' ability to simulate a coherent alternative to the primary hypothesis should reduce rulings against the defendant. For instance, consider some of the work on

pretrial publicity by Greene and others (Greene, 1990; Greene & Loftus, 1984; Greene & Wade, 1988). In these studies, mock jurors read newspaper articles about eyewitness testimony that led to a conviction in a rape trial. The testimony was later recanted and the convicted defendant was shown to be innocent. Jurors then were asked to render verdicts in a similar trial based on eyewitness testimony. Subjects who were exposed to the pretrial publicity had significantly lower conviction rates than subjects who were not exposed.

Although Greene and her colleagues (1984, 1988, 1990) interpreted their findings differently, it is plausible that the pretrial publicity provided jurors with a means to simulate a scenario in which the defendant was not guilty when they rendered verdicts in a similar trial. Research has demonstrated that hearing stories, scenarios, and analogies can create causal schematic structures that guide the interpretation of information provided in a similar framework (Johnson & Seifert, 1994; McKoon, Ratcliff, & Seifert, 1989; Schank, 1982).

This possibility was tested using the partial-match version of the blue bus/gray bus case used in Niedermeier et al. (1999). It was predicted that given the same evidence, mock jurors would rule against a defendant less when provided a framework to mentally simulate an alternative explanation than when simulation was more difficult. In other words, the ability to simulate an alternative hypothesis should reduce belief in a primary hypothesis. As in Niedermeier et al.

(1999) and Wells (1992), it was predicted that this difference would be manifest in categorical verdicts, but not in continuous measures of subjective probability.

Method

Participants

Participants were 72 undergraduate psychology students who took part in the study as a classroom demonstration. The mean age of participants was 21.9 years and 63% were female.

Design and Procedure

Participants were informed that they were taking part in a study on juror decision making and that they would act as jurors in a civil case and make several judgments. They were further told that the case they were going to read included some technical evidence about tire track matching. They were then given a technical information sheet that described how tire tracks can be lifted from an accident scene and later matched to actual tires on a vehicle. Although the text was fictitious, no participant questioned the veracity of the information. The text read as follows:

> Unknown vehicles can be identified by tire-track identification. This technique involves matching the tracks left at an accident scene to the tires on a vehicle. Tire-track identification is effective because each vehicle's tires display unique patterns of wear in addition to the fact that there are hundreds of different types of tires and tread patterns.

> Tire-tracks are usually lifted from the pavement using photosensitive paper and an ultraviolet light emission gun. The paper is laid over the track and exposed to the ultraviolet light. The light is absorbed by the rubber left by a recent tire-track but reflected by the pavement. Thus, the exposed paper produces a negative image of the track. This image is scanned into a computer, which prints out a positive image of the track that can then be compared to any suspect vehicles' tires. Matches are graded from 0% (no match) to 100% (perfect match).

For approximately half of the participants, text was added describing situations in which tire tracks can be obscured by vehicles passing by after the accident has occurred, sometimes making uninvolved vehicles appear responsible. This text was added to facilitate simulation of an alternative explanation. The additional text read as follows:

> It is important that the track being lifted is the most recent track on the pavement. If a second vehicle passes over the track in question, the technique can produce a mistaken match. For instance, if a car involved in a hit and run is trying to be identified, it is important that no other cars drive through the scene. If an innocent

car takes the same path as the offending car before the tracks are lifted, the technique could produce a better match with the innocent vehicle than with the offending vehicle.

All participants then read the partial-match version of the blue bus/gray bus trial. The text was as follows:

Mrs. Prob is suing the Blue Bus Company for having caused the death of her dog. At the trial, the following evidence was given:

Mrs. Prob testified that she was walking her dog along county road #37 when she heard a large vehicle behind her. She turned around and saw a bus serving recklessly down the road. She jumped out of the way but the bus swerved and hit her dog, killing him instantly. The incident occurred at 11:40 A.M. The bus continued at a high speed down the road. Unfortunately, Mrs. Prob is color blind and thus does not know the color of the bus.

A county transportation official took the stand, was sworn in as a witness, and testified that there are only two bus companies that travel in the county: the Blue Bus Company and the Gray Bus Company. Each company uses the road to run empty buses back to their stations after dropping off passengers. Therefore, one of the two bus

companies had to be responsible for the death of Mrs. Prob's dog.

A second county transportation official, who was the transportation department's chief investigator, took the stand, was sworn in as a witness, and reported that he examined the dead dog and took prints of the tire tracks. These prints were then transferred onto paper and compared to all 10 of the 10 buses owned by the Blue Bus Company, and all 10 of the 10 buses owned by the Gray Bus Company. He testified that the technique used for matching produced a partial match with one of the Blue Bus Company buses and a partial match with one of the Gray Bus Company buses. The results of the test indicated that there was an 80% chance that the blue bus was the one that ran over the dog, and a 20% chance that the gray bus was the one. Based on this matching evidence, the official concluded that it was a blue bus that killed the dog.

Mrs. Prob's attorney argued that the jury must find the Blue Bus Company liable for damages because in all likelihood, it was the Blue Bus Company that killed Mrs. Prob's dog.

After reading the trial, participants indicated whether they would rule against the Blue Bus Company and make them pay damages to Mrs. Prob. Next, they reported subjective

probabilities by estimating the probability that the Blue Bus Company actually killed the dog. Participants also rated confidence in their individual verdicts on an 11-point scale ranging from not at all confident (1) to completely confident (11). In order to assess the ease-of-simulation, participants were asked to rate how easy or difficult it was to think of a situation in which a bus from the Gray Bus Company could have been responsible for killing the dog on an eleven point scale ranging from very easy (1) to very difficult (11).

Results

The verdicts were in the predicted direction although the trend was not statistically significant. Participants who read the simulation framework ruled against the Blue Bus Company less (24%) than participants who read the partial-match case alone (39%), $\chi^2(1, N = 72) = 1.71$, $p < .20$, $\phi = .15$. Effect sizes for categorical measures were calculated using phi (Rosenthal & Rosnow, 1991). A one-tailed test was also not significant, $z = 1.31$, $p < .10$.

Participants who read the simulation framework reported that it was easier to think of a way that a bus from the Gray Bus Company could be responsible for killing the dog ($M = 3.80$) than participants who read the partial-match case alone ($M = 5.13$), $F(1,70) = 4.39$, $p < .05$, $\omega^2 = .05$. All effect sizes for continuous measures were calculated using omega squared (Tabachnick & Fidell, 1996). Additionally, a significant relationship between the ease-of-simulation measure and verdicts was revealed such that the easier

participants reported it was to simulate that a Gray Bus Company bus may have killed the dog, the less willing they were to hold the Blue Bus Company responsible, $r(69) = .39$, $p < .01$.

Verdicts were combined with the confidence measure to yield a scale that ranged from 1 (completely confident in a verdict against liability of the Blue Bus Company) to 22 (completely confident in a verdict for liability of the Blue Bus Company). Again, the means were in the predicted direction in that participants who read the partial-match case alone had more pro-conviction verdicts against the Blue Bus Company ($M = 10.16$) than participants who read the simulation framework ($M = 7.85$). However, the trend was not significant, $F(1,70) = 1.86$, ns, $\omega^2 = .01$.

Participants who read the simulation framework also reported that it was less likely that a bus from the Blue Bus Company killed the dog ($M = 61.83$) than participants who read the partial-match case alone ($M = 70.13$), $F(1,70) = 4.82$, $p < .05$, $\omega^2 = .05$. A significant relationship between the subjective probabilities and verdicts was revealed such that participants who reported lower subjective probabilities were less likely to rule against the Blue Bus Company, $r(69) = .42$, $p < .001$. Additionally, a significant relationship between ease-of-simulation and subjective probability was revealed such that participants who reported that it was easier to think of a way that a bus from the Gray Bus Company could be responsible for killing the dog reported that it was

less likely that the Blue Bus Company killed the dog, $r(69) = .33$, $p < .01$.

As the significant difference in subjective probabilities was unexpected, a regression of the ease-of-simulation measure on the verdicts was performed while controlling for the effects of subjective probability. This procedure was carried out to demonstrate that it was not simply subjective probabilities driving the categorical verdicts. Rather, ease-of-simulation was operating independently. The ease-of-simulation measure still accounted for a significant amount of the variance of verdicts after controlling for subjective probability, $t(69) = 2.52$, $p < .02$, $R^2 = .08$.

Discussion

The data from Experiment 1 demonstrated that the easier it was to simulate that a Gray Bus Company bus killed the dog, the less willing participants were to hold the Blue Bus Company responsible. Although this verdict pattern was not significant between conditions, a significant correlation showed that the relationship was evident across conditions. Thus, participants' ability to simulate an alternative hypothesis (that a gray bus killed the dog) was related to belief in the primary hypothesis (that a blue bus killed the dog).

It is possible that the manipulation intended to make simulation easier was not strong enough. Given that the potential effect size is "small," power analysis revealed

that the ability to detect the effect with the current sample size would be less than .15 (Cohen, 1988). It would take a sample of 400 to detect the effect at p = .05 with power of .50. Thus, strengthening the ease-of-simulation manipulation and increasing the sample size would aid in detecting the effect.

Furthermore, it may be that the effect was limited by a functional floor effect. It seems reasonable to assume that some people will rule against the Blue Bus Company simply because of the extreme likelihood that they are responsible, regardless of other factors. Additionally, the rate of verdicts against the Blue Bus Company for the partial-match case alone was 39% in this experiment. This rate was the lowest observed rate using the partial-match case in this study and in previous research (Niedermeier et al., 1999). This unusually low rate of verdicts against the Blue Bus Company in the partial-match condition and the likely functional floor effect may have suppressed the magnitude of the effect in this experiment, thus making it difficult to detect with the current sample size.

However, subjective ease-of-simulation may be more important than the particular manipulation in that individuals who were able to easily simulate a Gray Bus Company bus being responsible for killing the dog ruled against the Blue Bus Company less than those who found simulation more difficult, regardless of condition.

Although it was unexpected, participants who read the simulation framework reported that it was less likely that the Blue Bus Company killed the dog than participants who read the partial-match case alone. Moreover, a significant relationship between the ease-of-simulation measure and subjective probability was revealed such that participants who reported that it was easier to think of a way that a bus from the Gray Bus Company could be responsible for killing the dog reported that it was less likely that the Blue Bus Company killed the dog. These findings are inconsistent with Niedermeier et al. (1999) and Wells (1992), who both found that subjective probabilities did not differ across conditions and were unaffected by ease-of-simulation.

In the current framework, both continuous and categorical measures of the Blue Bus Company's responsibility are viewed as outcome variables, and in fact, both variables were related to ease-of-simulation. As mentioned before, some threshold models of decision making (e.g., Dane, 1985; Kaplan, 1977) could view the continuous subjective probabilities as driving the categorical verdicts. Although the pattern of data from this experiment could support such a model, ease-of-simulation still independently accounted for a significant amount of the variability when the effects of subjective probability were partitioned out. This is consistent with the argument that mental simulation affects categorical judgments above and beyond any effect of subjective probability.

Experiment 2

Given that the presence of alternative hypotheses can decrease the belief in a primary hypothesis, it follows that the elimination of such alternatives should increase belief in the primary hypothesis. As previously noted, the current paradigm is useful for investigating alternative hypotheses because actual probabilities can be held relatively constant at 80% (or any other value) for the Blue Bus Company and 20% for the Gray Bus Company. In Experiment 2, mock jurors made judgments about both the Blue Bus Company and the Gray Bus Company after reading either the complete-match or partial-match versions of the case. However, it was not made explicit that the Blue Bus Company was being sued. Jurors were merely told that they were going to read about a civil case in which a woman's dog was killed. Participants were asked to make continuous and categorical judgments while considering either the Blue Bus Company first or the Gray Bus Company first.

Considering the Gray Bus Company first may suggest to jurors that a gray bus being responsible is the primary hypothesis. As it is clearly unlikely that a gray bus is responsible (20%), jurors should easily reject this hypothesis. Thus, when asked to rule on the Blue Bus Company, the Gray bus hypothesis will already have been dismissed, increasing rulings against the Blue Bus Company. This should produce a main effect for order such that ruling against Blue will increase when Gray is considered first. This effect should be particularly evident in the complete-match

conditions because rulings against the Blue Bus Company should be very low when Blue is considered first, as in Niedermeier, et al. (1999). Furthermore, it was expected that previous findings (Niedermeier et al.) would be replicated such that jurors would find it easier to simulate a gray bus being responsible in the complete-match conditions and rule against the Blue Bus Company less than jurors in the partial-match conditions. As in Experiment 1, a relationship also was expected such that the easier participants report it is to simulate that a Gray Bus Company bus killed the dog, the less willing they would be to hold the Blue Bus Company responsible.

The predicted main effect for order, however, could occur because of a simple contrast effect. It could be that considering the Blue Bus Company's likely (80%) responsibility after the unlikely (20%) Gray Bus Company could increase rulings against the Blue Bus Company merely because of the contrast of probabilities. To test this notion, additional participants read a separate case prior to reading either the complete-match or partial-match case. This separate case involved a man who was suing a vitamin company because he suffered seizures after using the company's product. Testimony in the case revealed that there was only a 15% chance that the product could have caused seizures. If considering the Gray Bus Company first raises rulings against the Blue Bus Company because of a contrast effect, the

consideration of a separate, unlikely case should also increase ruling against the Blue Bus Company.

Method

Participants

Participants were 87 undergraduate psychology and communications students who received class credit for taking part in the study. The mean age of participants was 21.2 years and 72% were female.

Design and Procedure

Participants were told that they would be taking part in a study on juror decision making and that they would act as jurors in a case and make several judgments. Half of the participants read the partial-match case, as in Experiment 1. Again, the critical testimony in the partial-match case read as follows:

> A second county transportation official, who was the transportation department's chief investigator, took the stand, was sworn in as a witness, and reported that he examined the dead dog and took prints of the tire tracks. These prints were then transferred onto paper and compared to all 10 of the 10 buses owned by the Blue Bus Company, and all 10 of the 10 buses owned by the Gray Bus Company. He testified that the technique used for matching produced a partial match with one of the Blue Bus Company buses and a partial match with one of the Gray Bus Company buses. The results of the test

> indicated that there was an 80% chance that the blue bus was the one that ran over the dog, and a 20% chance that the gray bus was the one. Based on this matching evidence, the official concluded that it was a blue bus that killed the dog

The other half of the participants read the complete-match case. The complete-match case was identical to the partial-match case except for the fourth paragraph which reads as follows:

> A second county transportation official, who was the department's chief investigator took the stand, was sworn in as a witness, and reported that he examined the dead dog and took prints of the tire tracks. These prints were then transferred onto paper and compared to all 10 of the 10 buses owned by the Blue Bus Company, and all 10 of the 10 buses owned by the Gray Bus Company. The tracks matched 8 of the 10 Blue Bus Company's buses and 2 of the 10 Gray Bus Company's buses.

Participants were not told that Mrs. Prob was suing the Blue Bus Company. They were merely told that they were going to read about a civil case in which a woman's dog was killed. After reading the trial, participants indicated whether or not they would rule against the Blue Bus Company and make

them pay damages to Mrs. Prob and estimated the probability that the Blue Bus Company actually killed the dog. As in Experiment 1, confidence and ease-of-simulation measures were also collected. Participants made the same judgments about the Gray Bus Company. The measures for the different bus companies were on separate sheets and participants considered either the Blue Bus Company first or the Gray Bus Company first.

Approximately one third of the subjects participated in the contrast control conditions. These participants first read a case in which a man was suing the Herbrite Vitamin Company. The text of this control case read as follows:

> Mr. Smith had a very severe chest cold. His lungs were congested and he was coughing. In order to alleviate his symptoms, he took a herbal medicine that he bought at a local nutrition store. The medication, manufactured by the Herbrite Company, claimed it could help alleviate chest cold symptoms. Approximately six hours after he took the medicine, he began having respiratory seizures and was rushed to the hospital.
>
> The doctor who treated Mr. Smith testified that the patient was placed on a respirator and given medication to control the seizures. He further testified that Mr. Smith would likely have permanent damage to his lungs.

An independent biochemist conducted tests on the herbal medicine and Mr. Smith and concluded that there was only a 15% chance that the herbal medicine could have caused the seizures. In other words, he stated that there was an 85% chance that the seizures were caused by something other that than the herbal medicine.

Participants who read this control case uniformly reported that the probability that the herbal medicine caused Mr. Smith's seizures was quite low (M = 29%). Furthermore, most participants in the control conditions reported that they would not rule against the vitamin company (83%). After reading this case, participants read either the complete-match or partial-match case and filled out the dependent measures for only the Blue Bus Company.

Results

Table 1 reveals the verdicts against the Blue Bus Company, subjective probabilities, verdict/confidence means, and ease-of-simulation means for all conditions.

Initially, the verdicts against the Blue Bus Company in the complete-match version were compared to those in the partial-match version only for the participants who considered the Blue Bus Company first. This test was conducted as a replication of the basic contrasts in previous research (Niedermeier et al., 1999: Wells, 1992). As expected, participants who read the complete-match condition ruled against the Blue Bus Company significantly less than

Table 1

Verdicts, Subjective Probabilities, Verdict/Confidence, and Ease of Simulation by Condition (Experiment 2)

	Blue First		Gray First		Control	
DV	Complete	Partial	Complete	Partial	Complete	Partial
n	19	11	15	13	13	16
V	21%	64%	40%	54%	31%	44%
SP	74.2	82.6	75.3	78.3	77.3	75.0
V/C	7.7	13.4	9.7	12.0	10.6	11.2
EOS	6.1	7.2	6.1	7.6	5.8	6.8

Note. The numbers refer to percentage of verdicts against the Blue Bus Company (V), mean subjective probability (SP), mean verdict/confidence (V/C), and mean ease-of-simulation (EOS). Higher numbers on the verdict/confidence measure represent more confident verdicts against the Blue Bus Company. Higher numbers on the ease-of-simulation measure represent more reported difficulty.

those who read the partial-match version, $\chi^2(1, N = 30) = 5.44$, $p < 05$, $\phi = .42$ ($z = 2.46$, $p < .01$, one-tailed). This pattern did not hold when the Gray bus company was considered first, as participants ruled against the Blue Bus company at the same rate in the complete-match condition as the partial-match condition, $\chi^2(1, N = 28) = .54$, ns, $\phi = .14$.

Next, the pattern of verdicts when the participants considered the Blue Bus Company first was compared to the pattern of verdicts when the participants read the control case prior to ruling on the Blue Bus Company. If contrast with a very low probability case is enough to raise verdicts against the Blue Bus Company, a main effect should emerge such that participants who first read the control case should rule against the Blue Company more often than those who did not. However, participant ruled against the Blue Bus Company at the same rate when they read the control case (38%) as when they did not (37%), $\chi^2(1, N = 59) = .01$, ns, $\phi = .00$. The data from the control conditions were not considered in any further analyses.

The verdicts when participants considered the Blue Bus Company first were then compared to the conditions in which the Gray Bus Company was considered first. A main effect was expected such that participants who considered the Gray Bus Company first should rule against the Blue Bus company at a higher rate. Although the pattern was in the predicted direction, participants ruled against the Blue Bus Company at statistically the same rate when they considered the Gray Bus

Company first (46%) as they did when they considered the Blue Bus Company first (37%), $\chi^2(1, N = 58) = .56$, ns, $\phi = .10$, ($z = .70$, ns). Next, the verdicts for those who considered the Blue Bus Company first were compared to those who considered the Gray Bus Company first only for participants who read the complete-match case. Although the trend was in the predicted direction, the effect was not significant, $\chi^2(1, N = 58) = 1.45$, ns, $\phi = .21$, ($z = 1.20$, ns).

There was a main effect for case such that participants who read the complete-match case ruled against the Blue Bus Company significantly less (29%) than those who read the partial-match case (58%), $\chi^2(1, N = 58) = 4.85$, $p < 05$, $\phi = .29$ ($z = 2.20$, $p < .02$, one-tailed). Case version and order did not interact, $\chi^2(1, N = 58) = 1.38$, ns, $\phi = .15$.

Considering the possibility of a ceiling effect for verdicts against the Blue Bus Company in the partial-match conditions, it is plausible that the only cell that would have a low conviction rate is the complete-match case when Blue is considered first. Convictions rates should be high in both partial-match conditions, so there may not have been room for the partial-match conviction rate to get any higher when Gray was considered first. Also, it was predicted that conviction rates should be high in the complete-match when Gray is considered first. Therefore, it is reasonable to compare the complete-match case when Blue was considered first to the other three cells. Participants who considered Blue first and read the complete-match case in fact ruled

against the Blue bus Company significantly less than participants in the other three conditions, $\chi^2(1, N = 58) = 4.81$, $p < .05$, $\phi = .29$.

The combined verdict/confidence measures revealed a similar pattern to verdicts in that a main effect was revealed such that participants who read the partial-match case reported more pro-conviction verdicts against the Blue Bus Company ($M = 12.63$) than those who read the complete-match case ($M = 8.89$), $F(1,54) = 4.27$, $p < .05$, $\omega^2 = .06$. No other significant effects were revealed on this measure.

As Table 1 shows, subjective probabilities did not differ and were appropriately high across conditions. However, a significant relationship between the subjective probability and verdicts was revealed such that participants who reported lower subjective probabilities were less likely to rule against the Blue Bus Company, $r(56) = .28$, $p < .05$.

Next the ease-of-simulation measure was evaluated. Those who read the partial-match case reported that is was more difficult to think of a situation in which the Gray Bus Company could be responsible for killing the dog ($M = 7.33$) than those who read the complete-match case ($M = 6.10$), $F(1,54) = 4.64$, $p < .05$, $\omega^2 = .06$. Ease-of-simulation was unaffected by order. Furthermore, a significant relationship between the ease-of-simulation measure and verdicts was revealed such that the easier participants reported it was to simulate that a Gray Bus Company bus may have killed the dog, the less willing they were to hold the Blue Bus Company

responsible, $r(56) = .33$, $p < .02$. Additionally, ease-of-simulation still accounted for a significant amount of variance in the verdicts after controlling for the effects of subjective probability, $t(56) = 2.83$, $p < .01$, $R^2 = .09$. There was no relationship between ease-of-simulation and subjective probabilities, $r(56) = .15$, ns.

Lastly, the intervening role of ease-of-simulation was tested by entering case version, ease-of-simulation, and verdicts into a general linear model. When the effects of ease-of-simulation were controlled, the significant effect of case version ($p = .03$) was reduced below the level of significance ($p = .10$). Furthermore, the effect size of case version ($\omega^2 = .06$) was reduced by 50% ($\omega^2 = .03$).

Discussion

First, the basic pattern of findings in previous research (Niedermeier et al., 1999: Wells, 1992) was replicated. Participants who read the complete-match condition ruled against the Blue Bus Company significantly less than those who read the partial-match version. Moreover, participants' ability to simulate an alternative hypothesis (that a gray bus killed the dog) was related to belief in the primary hypothesis (that a blue bus killed the dog) across conditions.

Furthermore, it was demonstrated that ease-of-simulation plays an intervening role between case version and verdicts. When the effects of ease-of-simulation were partialed out of the model, the effect of case on verdicts was no longer

significant. This finding is suggestive that ease-of-simulation mediated verdicts.

Unlike Experiment 1, subjective probabilities did not differ across conditions and were appropriately high across conditions. This pattern is consistent with previous research (Niedermeier et al., 1999: Wells, 1992). There was an unexpected relationship between subjective probability and verdicts, but ease-of-simulation was still a significant predictor of verdicts after controlling for the effects of subjective probability. Furthermore, there was no relationship between subjective probability and ease-of-simulation, indication that both mechanisms independently affected verdicts. Overall, these findings further suggest that continuous and categorical measures do not necessarily produce parallel results and are not necessarily measuring the exact same thing.

The main hypothesis in this experiment was not supported in that verdicts against the Blue Bus Company were not higher when participants considered the Gray Bus Company first. However, the pattern of verdicts were in the predicted direction and may have become evident with additional power. Given that the potential effect size is small, power analysis revealed that the ability to detect the effect with the current sample size would be less than .10 (Cohen, 1988). It would take a sample of 400 to detect the effect at $p = .05$ with power of .50.

However, two findings did indicate that considering the alternative hypothesis first may have had an effect. First, more participants ruled against the Blue Bus Company in the complete-match condition than the partial-match condition when Blue was considered first, but there was no difference when Gray was considered first. Thus, considering Gray first did reduce the effect of case, despite the fact that a significant interaction failed to emerge. Second, the contrast of the complete-match condition when Blue was considered first to the other three conditions suggests that the consideration of Gray did affect verdicts by raising the conviction rate in the complete-match condition when Gray was considered first.

Additionally, a pilot study was conducted which was exactly the same as Experiment 2, but did not include ease-of-simulation measures. This pilot revealed the predicted main effect for order such that participants who considered Gray first ruled against the Blue Bus Company significantly less (26%) than those who considered Blue first (45%), $\chi^2(1, N = 126) = 4.63$, $p < 05$, $\phi = .19$ ($z = 2.25$, $p < .02$, one-tailed). The unusual fact that the pilot study had more participants than the actual experiment was due to an unforeseen shortage of subjects at the time the experiment was conducted. The combined order effect for the pilot and Experiment 2 was $z = 2.09$, $p < .02$.

Experiment 2 clearly replicated the basic contrast between the easily simulated alternative in the complete-

match case and the less easily simulated alternative in the partial-match case. Second, the relationship between ease-of-simulation and verdict was again demonstrated and ease-of-simulation was shown to play a mediating role. Lastly, Experiment 2, when combined with the data from the pilot study, provided some suggestive evidence that eliminating an hypothesis by considering it first and categorically rejecting it may increase acceptance of the primary hypothesis.

Experiment 3

The research by Miller et al. (1989) suggests that increasing the number of alternatives will undermine belief in a given hypothesis. In their study, the proportion of the alternatives is constant (10%), only the absolute number is changed. For instance, the existence of 10 chocolate chip cookies reduced subjects' suspicion that the child peeked in the cookie jar compared to when there was one chocolate chip cookie. Similarly, other research (Fischhoff et al., 1978; McDonald, 1998) also suggests that the likelihood that people will accept a given explanation for an event is directly related to the number of available alternatives. Again, none of these studies collected categorical measures.

Alternatively, the Niedermeier et al. (1999) study suggests that a single coherent alternative is enough to undermine the primary hypothesis. In their study, the lone possibility of a second bus company was enough to reduce

rulings against the Blue Bus Company in the complete-match condition. Additional alternatives would not necessarily enhance ease of simulation, especially if the total probability of all alternatives were held constant.

In Experiment 3, participants read an adaptation of the complete-match trial in which the Blue Cab Company was being sued for killing Mrs. Prob's dog. There were either one or five other cab companies that produced matches to the tire tracks at the scene. Cabs were used instead of buses in this experiment because it is more plausible that multiple cab companies would operate in one area than multiple bus companies. Additionally, some subjects read a version in which they were aware of five alternative matches, but were unaware to which company the matches belonged. In this study, participants were told that there were 100 cabs that matched the tire tracks at the scene. Ninety-five Blue Cab Company cabs matched the tracks, as did five cabs from other companies. The overall probability of a blue cab being responsible was 95% in this experiment (compared to 80% in the previous studies) in order to avoid a possible floor effect and better differentiate between the conditions. This was deemed necessary because pretests in which 80% of the implicated cabs were blue produced very few findings against the Blue Cab Company in any condition.

If increasing the absolute number of alternatives reduces belief in a primary hypothesis, participants should have ruled against the Blue Cab Company less when there were

five alternative cab companies that produced matches than when there was only one. Research would suggest that increasing alternative routes to a particular outcome should reduce belief in given hypothesis (Fischhoff et al., 1978; McDonald, 1998; Miller et al., 1989). However, if a single easily simulated alternative is all that is necessary, rulings against the Blue Cab Company should be equally low in all three conditions. As in Experiments 1 & 2, a relationship was expected such that the easier participants report it is to simulate that an alternative cab company may have killed the dog, the less willing they would be to hold the Blue Cab Company responsible. Furthermore, no differences were expected on the continuous measure of subjective probability.

Method

Participants

Participants were 141 undergraduate psychology students who took part in the study as a classroom demonstration. The mean age of participants was 20.0 years and 74% were female.

Design and Procedure

Participants were told that they would be taking part in a study on juror decision making and that they would act as a juror in a civil case and make several judgments. Approximately one third of the subjects read a case entitled the one-alternative version. The case was the same as the case in Experiment 1, except that it referred to cabs instead of buses. The critical fourth paragraph read as follows:

A second county transportation official, Mr. Jones, took the stand, was sworn in as a witness, and reported that he examined the dead dog and took prints of the tire tracks. The tire tracks were checked against all cabs that operate out of the central dispatching station on County Road #37. Mr. Jones produced a table that showed the results of his investigation. As the table shows, the prints matched 95 of the cabs owned by the Blue Cab Company and 5 of the cabs owned by the Gray Cab Company. None of the cabs from the other companies produced a match [see Appendix A].

Approximately one third of the participants read the five-separate-alternatives version. The only difference in this case was the critical fourth paragraph which read as follows:

A second county transportation official, Mr. Jones, took the stand, was sworn in as a witness, and reported that he examined the dead dog and took prints of the tire tracks. The tire tracks were checked against all of the cabs that operate out of the central dispatching station on County Road #37. Mr. Jones produced a table that showed the results of his investigation. As the table shows, the prints matched 95 of the cabs owned by the Blue Cab Company and one cab from five other cab companies (that is, one match each from the Gray,

Yellow, Red, Orange, and Purple Cab Companies) [see Appendix A].

The remaining third of the participants read the five-grouped-alternatives version. The only difference in this case was the critical fourth paragraph which will read as follows:

> A second county transportation official, Mr. Jones, took the stand, was sworn in as a witness, and reported that he examined the dead dog and took prints of the tire tracks. The tire tracks were checked against all of the cabs that operate out of the central dispatching station on County Road #37. Mr. Jones produced a table that showed the results of his investigation. As the table shows, the prints matched 95 of the cabs owned by the Blue Cab Company and 5 of the cabs owned by five other cab companies [see Appendix A].

After reading the trial, participants indicated whether or not they would rule against the Blue Cab Company and make them pay damages to Mrs. Prob and estimated the probability that the Blue Cab Company actually killed the dog. As, in Experiment 1 and 2, confidence measures were also gathered. To measure ease-of-simulation, participants were asked to rate how difficult it was to think of a way that a cab other than one from the Blue Cab Company could have been

responsible for killing the dog on an 11-point scale from very easy (1) to very difficult (11).

Results

Table 2 reveals the results for Experiment 3. Rulings against the Blue Cab Company did not differ among conditions, $\chi^2(2, N = 141) = 1.63$, ns, $\phi = .10$. Likewise, there were no differences among conditions on the verdict/confidence measure, $F(2,138) = .73$, ns, $\omega^2 = .00$. Also, subjective probabilities did not differ and were appropriately high across conditions, $F(2,138) = 1.13$, ns, $\omega^2 = .00$. Furthermore, there was no relationship between subjective probability and verdicts, $r(139) = .15$, ns.

Participants who read the one-alternative version reported that it was more difficult to think of a way that a cab from a company other than the Blue Cab Company could have been responsible for killing the dog than those who read the five-separate-alternatives version or the five-grouped-alternatives version, $F(2,138) = 12.41$, $p < .001$, $\omega^2 = .14$.

Finally, a significant relationship between the ease-of-simulation measure and verdicts was revealed such that the easier participants reported it was to simulate that a cab from a company other than the Blue Cab Company may have killed the dog, the less willing they were to hold the Blue Cab Company responsible, $r(139) = .34$, $p < .001$. There was no relationship between ease-of-simulation and subjective probability, $r(139) = .11$, ns.

Table 2

Verdicts, Subjective Probabilities, Verdict/Confidence, and Ease of Simulation (Experiment 3)

	Alternatives		
Measure	One	Five	Five grouped
n	48	47	46
V	29%	36%	24%
SP	90.9	90.0	85.9
V/C	9.00	9.94	8.09
EOS	6.46_a	3.77_b	4.15_b

Note. The numbers refer to percentage of verdicts against the Blue Bus Company (V), mean subjective probability (SP), mean verdict/confidence (V/C), and mean ease-of-simulation (EOS). Higher numbers on the verdict/confidence measure represent more confident verdicts against the Blue Bus Company. Higher numbers on the ease-of-simulation measure represent more reported difficulty. Means with different subscripts are different at $p < .01$.

Discussion

Rulings against the Blue Cab Company were equally low in all conditions, indicating that any coherent alternative is enough to suppress verdicts of liability. Across all three conditions, participants ruled against the Blue Cab Company at a rate of 30%. This seems quite low when compared to the partial-match cases in Experiment 2 and Niedermeier et al. (1999) in which participants ruled against the defendant at a rate of approximately 55%. Also, keep in mind that the actual probability of Blue's responsibility was 80% in the previous studies compared to 95% in the current experiment.

An informative condition would have been a partial-match version of the cab case in which there were incomplete matches with the Blue Cab Company and alternative cab companies that resulted in a 95% chance the Blue was responsible. This would have allowed a clear contrast with the three existing conditions. Provided that the partial-match condition was high, as in previous studies, this contrast would bolster the conclusion that any number of easily-simulated coherent alternatives would reduce willingness to rule against the Blue Cab Company.

The differences between conditions on the ease-of-simulation measure was unexpected. However, the measure in this experiment concerned simulation of any company other than blue, whereas the previous experiments specified simulation of gray. Therefore, the object of simulation in this experiment was not constant among the conditions. For

instance, participants must pick out the Gray Cab Company to answer the question in the one-alternative condition, but they may consider any of five companies in the five-separate-alternatives condition. Therefore participants in the one-alternative condition must take the extra mental step of separating the Gray Cab Company from the others. Alternatively, participants in the one-alternative condition may have mentally averaged across all of the available alternatives, which included five companies with no matches. In either case, this could account for the elevated ease-of-simulation scores in the one-alternative condition. However, the basic relationship between ease-of-simulation and verdicts held in this experiment, despite the potential ambiguity in the ease-of-simulation measure.

Experiment 4

Much of the research on lay hypothesis testing assumes that reliance on the availability of alternatives to make judgments is a heuristic (Kahneman & Tversky, 1982; McDonald, 1998; Niedermeier et al., 1999). Kahneman and Tversky dubbed the phenomenon currently under study "the simulation heuristic." Additionally, McDonald claims that beliefs in many extraordinary phenomena such as telekinesis and alien crop circles are bolstered by the inability to easily simulate alternative explanations. McDonald termed this effect the "can-you-think-of-anything-else heuristic."

This raises the question of whether the reduction of belief in a primary hypothesis when alternative hypotheses are easily simulated is indeed a heuristic rather than inherently flawed, mindful processing. Popular dual-processing models, such as Petty and Cacioppo's (1984, 1986) elaboration likelihood model or Chaiken's (1987) systematic/heuristic model of attitude change may lend some insight. One might conceive of evidence as a persuasive communication and presume that the recipients of this communication will only process it systematically (i.e., carefully and thoroughly analyzing all available relevant information) if they have both the capacity and the motivation to do so. Otherwise, they might be expected to process the communication heuristically, that is, to utilize cues to simplify judgment and decision making.

In the Blue Bus Company case, it is plausible that mock jurors process the available evidence heuristically. Specifically, mock jurors may be using an easily simulated alternative as a heuristic cue to find the Blue Bus Company not liable. If this is true, the stronger their motivation to consider the available evidence systematically, the less they would be expected to rely upon such a heuristic cue. So given motivation to process the case systematically, jurors would be less willing to rely on the simulation heuristic and become more likely to rule against the Blue Bus Company. If participants process in a mindful fashion, verdicts should be

more in line with actual and reported subjective probabilities.

In this experiment, participants read either the standard complete-match case, in which there is an easily simulated alternative hypothesis, or the partial-match case, in which there is not an easily simulated alternative. Half of the participants were told at the outset of the study that, after they made their individual choices, they would be asked to report and defend their private verdict choices in front of the entire group in order to promote systematic processing. Previous research (e.g., Petty, Cacioppo, & Goldman, 1981) has demonstrated the efficacy of such a "public accountability" treatment for increasing systematic processing. The other half was given no such instructions and should presumably rely more on mental simulation as in the previous experiments. If mental simulation is indeed used as a heuristic cue, an interaction is expected such that rulings against the Blue Bus Company should be more common when jurors are cognitively engaged than when they are not or when mental simulation is difficult. When engagement was low, the original pattern of few rulings against blue in the complete-match should have been observed. Alternatively, if mental simulation is part of an intrinsically flawed reasoning process, then verdicts against the Blue Bus Company should be lower in the complete-match cases than the partial-match cases, regardless on engagement.

Method

Participants

Sixty introductory psychology students received extra class credit for taking part in the study. Participants were run in groups of three to thirteen. Group size showed no effects on any individual measures. The mean age of participants was 20.1 years and 83% were female.

Design and Procedure

The cases used in this experiment were the standard complete-match and partial-match cases used in Experiment 2. Crossed with this treatment was a cognitive engagement (motivation to process systematically) treatment. All participants were told that they would be taking part in a study on juror decision making and that they would act as jurors in a case and make several judgments. Half of the participants were told at the outset of the study that after they made their individual choices, they would have to report and defend their private verdict choices in front of the entire group. The other half was given no such instruction.

Thus, the presence of an easily simulated alternative (complete-match) or less easily simulated alternative (partial-match) was crossed with the cognitive engagement manipulation (high engagement vs. low engagement) to yield a 2 X 2 factorial design.

After reading the trial, participants indicated whether or not they would rule against the Blue Bus Company and make them pay damages to Mrs. Prob and reported subjective

probability by estimating the likelihood that the Blue Bus Company actually killed the dog. As in the previous experiments, confidence and ease-of-simulation measures were also gathered. In addition, participants responded to an open-ended question that asked them to explain how they reached their verdict.

Results

Table 3 reveals the verdicts against the Blue Bus Company, subjective probabilities, verdict/confidence means, and ease-of-simulation means for all conditions.

Initially, the verdicts against the Blue Bus Company in the complete-match version were compared to those in the partial-match version only for the participants who did not receive the engagement instructions. This test was conducted as a replication of the basic contrasts in previous research (Experiment 2, Niedermeier et al., 1999: Wells, 1992). An effect was revealed such that participants who read the complete-match condition ruled against the Blue Bus Company less than those who read the partial-match version, $\chi^2(1, N = 31) = 3.48$, $p < 07$, $\phi = .33$ ($z = 1.96$, $p < .05$). As Table 3 reveals, the pattern was the same in the high engagement condition, $\chi^2(1, N = 31) = 6.43$, $p < 02$, $\phi = .46$ ($z = 2.70$, $p < .01$). When participants in both engagement conditions were considered, those who read the complete-match condition ruled against the Blue Bus Company significantly less (14%) than those who read the partial-match version (52%), $\chi^2(1, N = 60) = 9.64$, $p < .001$, $\phi = .40$ ($z = 3.33$, $p < .001$).

Table 3

Verdicts, Subjective Probabilities, Verdict/Confidence, and Ease of Simulation by Condition (Experiment 4)

	Complete-match		Partial-match	
Measure	High	Low	High	Low
n	14	15	15	16
V	14%	13%	60%	45%
SP	77.9	73.3	74.3	69.1
V/C	6.1	7.1	13.1	11.0
EOS	3.7	5.0	5.5	5.1

Note. The numbers refer to percentage of verdicts against the Blue Bus Company (V), mean subjective probability (SP), mean verdict/confidence (V/C), and mean ease-of-simulation (EOS). Higher numbers on the verdict/confidence measure represent more confident verdicts against the Blue Bus Company. Higher numbers on the ease-of-simulation measure represent more reported difficulty.

As Table 3 also reveals, participants ruled against the Blue Bus Company at statistically the same rate when they received engagement instructions (38%) as when they did not (29%), $\chi^2(1, N = 60) = .53$, ns, $\phi = .09$. Additionally, there was no interaction between case type and engagement instructions, $\chi^2(1, N = 60) = .32$, ns, $\phi = .07$.

There were no significant effects for subjective probability, as estimates were appropriately high across conditions. There was no relationship between subjective probability and verdicts, $r(58) = .02$, ns.

The verdict/confidence measure revealed a main effect such that participants in the partial-match conditions had more pro-conviction verdicts against the Blue Bus Company ($M = 12.03$) than participants in the complete-match conditions ($M = 6.62$), $F(1,56) = 8.62$, $p < .01$, $\omega^2 = .12$. There were no other significant effects on this measure.

There were no main effects or interaction involving the ease-of-simulation measure. However, there was a significant relationship between the ease-of-simulation measure and verdicts such that the easier participants reported it was to simulate that a bus from the Gray Bus Company may have killed the dog, the less willing they were to hold the Blue Bus Company responsible, $r(58) = .29$, $p < .05$. There was no relationship between ease-of-simulation and subjective probability, $r(58) = .05$, ns.

The open-ended question was coded for consideration of the Gray Bus Company. Participants were scored as either

considering The Gray Bus Company in their decision or not. Any mention of the Gray Bus Company (or a reference to the "other" company) was scored as considering Gray. Table 4 reveals the percentage of participants who considered Gray in each condition. There was a non-significant trend that suggested participants were more likely to consider Gray when they received engagement instructions (69%) than when they did not (52%), $\chi^2(1, N = 60) = 1.88$, ns, $\phi = .18$. Participants who read the complete-match case considered the Gray equally under both engagement conditions, $\chi^2(1, N = 60) = .01$, ns, $\phi = .00$. However, participants who read the partial-match case were more likely to consider Gray when they received engagement instruction than when they received no such instructions, $\chi^2(1, N = 60) = 3.89$, $p < 05$, $\phi = .25$ ($z = 2.02$, $p < .05$). Although these findings suggest an interaction, the trend did not reach significance, $\chi^2(1, N = 60) = 2.42$, $p = .20$, $\phi = .33$ ($z = 1.54$, $p < .13$).

Discussion

First, the basic pattern of findings in previous research (Niedermeier et al., 1999: Wells, 1992) and in Experiment 2 was replicated in that participants were more willing to rule against the Blue Bus Company when they read the complete-match case than when they read the partial-match case. Furthermore, increased ability to simulate an alternative hypothesis (that a gray bus killed the dog) was again reliably related to belief in the primary hypothesis (that a blue bus killed the dog).

Table 4

Percentage of Participants Who Considered Gray

Engagement	Complete-match	Partial-match
High	79	60
Low	80	25

Additionally, subjective probability was consistently high in all conditions and showed no relationship with verdicts, indicating that continuous and categorical measures do not necessarily produce parallel results.

The analysis of the open ended question did suggest that participants were more likely to consider an alternative hypothesis when they received engagement instructions than when they did not. This pattern was evident for participants who read the partial-match case. Consideration of Gray was high in both engagement conditions for those who read the complete-match case, which stands to reason as simulating Gray in this case has been shown to be easy (Experiment 2, Niedermeier et al., 1999). These results suggest that the engagement manipulation did motivate participants to process more systematically and consider more of the evidence.

A potentially informative condition might be to limit engagement by making participants cognitively busy. This manipulation may limit participants ability to simulate the alternative hypothesis. Without the ability to simulate alternatives, participants may rely more heavily on the probabilistic information and rule against the Blue Bus Company in the complete-match condition.

Overall, participants in Experiment 4 ruled against the Blue Bus Company no differently when they were highly engaged than when they were less motivated to consider the evidence. This pattern suggests that the basic difference in verdicts

is not solely due to reliance on a heuristic. Rather, it suggests that the verdict differences are the result of inherently flawed reasoning, as the pattern held when engagement was high and systematic processing was likely. However, it should be noted that the sample size was small limiting the power the detect any potential effect.

General Discussion

This study began with four objectives. First, the notion that mental simulation plays a role in casual decision making was tested by making alternative hypotheses more or less easily simulated. The trend in Experiment 1 and significant main effects for case version in Experiments 2 and 4 support the notion that an easily simulated alternative reduces willingness to accept a primary hypothesis. Furthermore, all four experiments revealed significant relationships between ease-of-simulation and verdicts. Additionally, the effects of ease-of-simulation on verdicts was consistently shown to be independent of subjective probability estimates.

Second, the importance of categorically eliminating potential alternatives was investigated by altering the order in which participants considered various alternative explanations. The results of Experiment 2 (along with the pilot data) suggest that considering one alternative first and categorically rejecting it increased willingness to accept an hypothesis presented subsequently.

Third, the relative frequency of alternative possibilities were altered to determine if one coherent alternative is sufficient to reduce belief in the primary hypothesis as opposed to several alternatives. The results of Experiment 3 suggest that a single coherent alternative hypothesis is sufficient to reduce belief in a primary hypothesis. Increasing the number of alternatives beyond a single one did not have a significant effect on the willingness to accept the primary hypothesis.

Lastly, the heuristic nature of mental simulation was investigated by manipulating participants' motivation to consider alternatives in order to determine if mental simulation is simply a cue or a systematic but intrinsically flawed method of reasoning. Experiment 4 demonstrated that verdicts were not altered by increasing motivation to carefully process evidence, indicating that the use of mental simulation is not so much a reliance on a heuristic as it is intrinsically flawed reasoning.

Although manipulating moderators of simulation proved more difficult than expected across experiments, it was clear in all four experiments that ease-of-simulation was related to the acceptance of a primary hypothesis. More specifically, the easier it was to simulate that a Gray Bus Company bus killed the dog (the alternative hypothesis), the less willing participants were to hold the Blue Bus Company responsible (the primary hypothesis). The combined effect of ease-of-simulation for all four experiments was $\underline{r} = .35$, indicating a

significant and reliable relationship. Whatever made simulation easy or difficult for individuals across experiments clearly related to verdicts and was independent of any effect of subjective probability.

Furthermore, the results of Experiments 2 and 4 directly replicated previous work in this area (Niedermeier et al., 1999, Wells, 1992). Participants who read the complete-match case, in which simulation of the alternative was easy, were more likely to rule against the Blue Bus Company than those who read the partial-match case, in which simulation was more difficult.

Although ease-of-simulation is theoretically viewed as driving verdicts in this context, it could be argued that participants were simply justifying their verdicts through the ease-of-simulation measures. However, this argument seems implausible for a number of reasons. First, when ease-of-simulation was controlled for in Experiment 2 and Niedermeier et al. (1999), the effect of case version on verdicts was reduced, indicating that ease-of-simulation indeed plays an mediating role. Second, it seems unlikely that participants in experiments such as these would feel so compelled to justify their verdicts that they would do so on a measure that is not obviously related to the verdict. If justification were a motivating factor, one would think this would become evident on the subjective probability measure, which is more clearly linked to the verdict.

Another contribution of this study is the demonstration that continuous and categorical measures do not necessarily produce parallel results. In much of the previous work (Dougherty et al., 1997; Fischhoff et al., 1978; McDonald, 1998; Robinson & Hastie, 1985; Robinson Van Wallendael, 1989), only continuous measures of belief in the primary hypothesis were collected. Some of these studies showed effects of alternative hypotheses and other did not. The current experiments demonstrate that continuous subjective probabilities do not necessarily translate into like categorical responses, as subjective probabilities had no relationship with categorical verdicts in Experiments 3 and 4. There was such a relationship in Experiments 1 and 2, but it was shown to be independent of the effects of mental simulation on categorical verdicts. These findings, along with those of Niedermeier et al. (1999) and Wells (1992), converge to show that a seemingly precise continuous measure may fail to capture a psychological effect evident in a dichotomous categorical measure, at least in cases in which there is a clear primary hypothesis.

Methodologically, it is important to recognize the non-parallel nature of the measures. Clearly, studies that look at continuous or categorical judgments alone may be missing an important piece of the decision making puzzle. Furthermore, a dichotomous measure, in some instances, may actually be more precise and meaningful than a continuous subjective probability. For instance, decisions to vote

guilty or not guilty, to marry or not marry, or to buy or not buy, are the types of categorical choices that have meaning and impact in our lives, regardless of our subjective probability estimates.

As alluded to earlier, many of the models of juror decision making fail to adequately consider the power of alternative explanations and the non-parallel nature of continuous and categorical measures. For instance, the story model (Pennington & Hastie, 1992, 1993) assumes that one story will emerge as the best and the other stories merely will be dismissed. However, the current research suggests that the power of a given explanation relies in large part on the presence and ease-of-simulation of alternative explanations. It stands to reason that the story model could benefit by considering ease-of-simulation of alternative stories.

Additionally, threshold models of decision making (e.g., Dane, 1985; Kaplan, 1977) generally view continuous subjective probability as driving categorical verdicts. These models maintain that if subjective probability surpasses some psychological threshold of guilt, jurors will rule against the defendant. This approach fails to capture the pattern of results found in this study and previous research in which categorical verdicts changed while continuous subjective probabilities remained static (Niedermeier et al., 1999, Wells, 1992). A way in which these findings could be rectified with the family of threshold models is if mental

simulation affects the level of the threshold criterion. It may be that the ability to easily simulate an alternative hypothesis raises the threshold that subjective probability must surpass in order to rule against a defendant. Thus, subjective probability may more accurately predict verdicts if the effect of mental simulation of alternatives on thresholds could be assessed. This is a proposition that merits further investigation.

Although the paradigm used in this study has many merits and has been an effective tool for studying causal decision making, it may inject some unwanted elements into the process. For instance, the fact that there are legal standards of liability that are consider by many participants may create a default verdict of not liable. Although there are likely subjective thresholds and defaults involved in any decision, the legal standards inherent in this paradigm may limit generalizability. Another problem with the current study is power. Some of the trends that are suggested by the data may have been clearer and emerged as significant given larger sample sizes. Future research should attempt to both replicate these findings with greater power and generalize the ease-of-simulation phenomenon using non-legal scenarios.

Conclusion

The purpose of the present investigation was to explore how the presence and properties of alternative hypotheses affect belief in a primary hypothesis. Across experiments,

participants generally ruled against the defendant less when they were able to mentally simulate an alternative explanation than when simulation was more difficult. In other words, the ability to simulate an alternative hypothesis was associated with reduced belief in the primary hypothesis. However, the factors that influence individuals' ability to simulate alternative hypotheses remain unclear. Further work is needed to identify specific moderators of mental simulation and refine current models of decision making.

APPENDIX

APPENDIX A

STIMULI USED IN EXPERIMENT 3

One-Alternative Version

Cab Company	Total # of matches
Blue	95
Gray	5
Yellow	0
Red	0
Orange	0
Purple	0
Green	0

Five-Separate-Alternatives Version

Cab Company	Total # of matches
Blue	95
Gray	1
Yellow	1
Red	1
Orange	1
Purple	1
Green	0

Five-Grouped-Alternatives Version

Cab Company	Total # of matches
Blue	95
Gray	5
Yellow	
Red	
Orange	
Purple	
Green	0

REFERENCES

REFERENCES

Chaiken, S. (1987). The heuristic model of persuasion. In M.P. Zanna, J.M. Olsen, & C.P. Herman (Eds.), Social influence: The Ontario symposium (Vol. 5). Hillsdale, NJ: Erlbaum.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Erlbaum.

Dane, F. (1985). In search of reasonable doubt: A systematic examination of selected quantification approaches. Law and Human Behavior, 9, 141-158.

Doherty, M.E., Chadwick, R., Garavan, H., Barr, D., & Mynatt, C.R. (1996). On people's understanding of the diagnostic implications of probabilistic data. Memory and Cognition, 24, 644-654.

Dougherty, M.R.P., Gettys, C.F., & Thomas, R.P. (1997) The role of mental simulation in judgments of likelihood. Organizational Behavior and Human Decision Processes, 70, 135-148.

Einhorn, H.J., & Hogarth, R.M. (1986). Judging probable cause. Psychological Bulletin, 99, 3-19.

Fischhoff, B. (1975). Hindsight foresight: The effect of outcome knowledge on judgment under uncertainty. Journal of Experimental Psychology: Human Perception and Performance, 1, 288-299.

Fischhoff, B., Slovic, P. & Lichtenstein, S. (1978). Fault trees: Sensitivity of estimated failure probabilities to problem representation. Journal of Experimental Psychology: Human Perception and Performance, 4, 330-344.

Greene, E., & Loftus, E.G. (1984). What's new in the news? The influence of well-publicized news events on psychological research and courtroom trials. Basic and Applied Social Psychology, 5, 211-221.

Greene, E. (1990). Media effects on jurors. Law and Human Behavior, 14, 439-450.

Greene, E., & Wade, R. (1988). Of private talk and public print: General pre-trial publicity and juror decision-making. Applied Cognitive Psychology, 2, 123-135.

Hastie, R. (1993). Inside the juror: The psychology of juror decision making. New York: Cambridge University Press.

Hawkins, S.A., & Hastie, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known. Psychological Bulletin, 107, 311-327.

Hilton, D.J. (1995). The social context of reasoning: Conversational inference and rational judgment. Psychological Bulletin, 118, 248-271.

Hirt, E.R., & Markman, K.D. (1995). Multiple explanations: A consider-an-alternative strategy for debiasing judgments. Journal of Personality and Social Psychology, 69, 1069-1086.

Johnson, H.M., & Seifert, C.M. (1994) Sources of the continued influence effect: When misinformation in memory affects later inferences. Journal of Experimental Psychology: Learning, Memory, and Cognition, 20, 1420-1436.

Jou, J., Shanteau, J., & Harris, R.J. (1996). An information processing view of framing effects: The role of causal schemas in decision making. Memory & Cognition, 24, 1-15.

Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), Judgment under uncertainty: Heuristics and biases. New York: Cambridge University Press.

Kaplan, M.F. (1977) Judgment by juries. In M.F. Kaplan & S. Schwartz (Eds.), Human judgment and decision processes in applied settings. San Diego, CA: Academic Press.

Koehler, D.J. (1994). Hypothesis generation and confidence in judgment, Journal of Experimental Psychology: Learning, Memory, and Cognition, 20, 461-469.

McDonald, J. (1998). 200% probability and beyond: The compelling nature of extraordinary claims in the absence of alternative explanations. Skeptical Inquirer, 22, 45-49, 64.

McKoon, G., Ratcliff, R., & Seifert, C.M. (1989). Making the connection: Generalized knowledge structures in story understanding. Journal of Memory and Language, 28, 711-734.

Miller, D.T., Turnbull, W., & McFarland, C. (1989). When a coincidence is suspicious: The role of mental simulation. Journal of Personality and Social Psychology, 57, 581-589.

Niedermeier, K.E., Kerr, N.L., & Messé, L.A. (1999) Juror's use of naked statistical evidence: Exploring the bases and implications of the Wells effect. Journal of Personality and Social Psychology, 76, 533-542.

Pennington, N., & Hastie R. (1986). Evidence evaluation in complex decision making. Journal of Personality and Social Psychology, 51, 242-258.

Pennington, N., & Hastie R. (1988). Explanation-based decision effects: Effects of memory structure on judgment, Journal of Experimental Psychology: Learning, Memory, and Cognition, 14, 521-533.

Pennington, N., & Hastie R. (1992). Explaining the evidence: Tests of the story model for juror decision making. Journal of Personality and Social Psychology, 62, 189-206.

Pennington, N., & Hastie R. (1993). The story model for juror decision making. In R. Hastie (Ed.), Inside the juror: The psychology of juror decision making, (pp.192-221), New York: Cambridge University Press.

Petty, R.E., & Cacioppo, J.T. (1984). The effect of involvement on responses to argument quantity and quality: Central and peripheral routs to persuasion. Journal of Personality and Social Psychology, 46, 69-81.

Petty, R.E., Cacioppo, J.T., & Goldman, R. (1981). Personal involvement as a determinant of argument-based persuasion. Journal of Personality and Social Psychology, 41, 847-855.

Robinson, L.B., & Hastie, R. (1985). Revision of beliefs when a hypothesis is eliminated from consideration. Journal of Experimental Psychology: Human Perception and Performance, 11, 443-456.

Robinson Van Wallendael, L.R. (1989). The quest for limits on noncomplimentarity in opinion revision. Organizational Behavior and Human Decision Processes, 43, 385-405.

Rosenthal, R. and Rosnow, R.L. (1991). Essentials of behavioral research: Methods and data analysis (2nd ed.). New York: McGraw-Hill.

Schank, R.C. (1982). Dynamic memory: A theory of reminding and learning in computers and people. New York: Cambridge University Press.

Schwarz, N. (1994). Judgment in social context: Biases, shortcomings, and the logic of conversation. Advances in Experimental Social Psychology, 26, 123-162.

Sykes, D.L., & Johnson, J.T. (in press). Probabilistic evidence vs. the representation of an event: The curious case of Mrs. Prob's Dog. Basic and Applied Social Psychology.

Tabachnick, B.G., & Fidell, L. (1996). Using multivariate statistics (3rd ed.). New York: Harper & Row.

Wason, P.C. (1960). On the failure to eliminate hypotheses in a conceptual task. The Quarterly Journal of Experimental Psychology, 12, 129-140.

Wells, G.L. (1992). Naked statistical evidence of liability: Is subjective probability enough? Journal of Personality and Social Psychology, 62, 739-752.