

6-19-2015


# It Must Have Been Him: Coherence Effects within the Legal System

Jonathan N. Carbone

Florida International University, jcarb010@fiu.edu

**DOI:** 10.25148/etd.FIDC000095

Follow this and additional works at: <https://digitalcommons.fiu.edu/etd>

 Part of the [Cognitive Psychology Commons](#), [Criminal Law Commons](#), [Criminal Procedure Commons](#), [Law and Philosophy Commons](#), [Law and Psychology Commons](#), [Other Law Commons](#), [Other Psychology Commons](#), and the [Social Psychology Commons](#)

---

## Recommended Citation

Carbone, Jonathan N., "It Must Have Been Him: Coherence Effects within the Legal System" (2015). *FIU Electronic Theses and Dissertations*. 2204.

<https://digitalcommons.fiu.edu/etd/2204>

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact [dcc@fiu.edu](mailto:dcc@fiu.edu).

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

IT MUST HAVE BEEN HIM: COHERENCE EFFECTS WITHIN THE LEGAL  
SYSTEM

A dissertation submitted in partial fulfillment of

the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PSYCHOLOGY

by

Jonathan Nicholas Carbone

2015

To: Dean Michael Heithaus  
College of Arts of Sciences

This dissertation, written by Jonathan Nicholas Carbone, and entitled It Must Have Been Him: Coherence Effects within the Legal System, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

---

Ryan J. Winter

---

Ronald P. Fisher

---

Jamie L. Flexon

---

Steve D. Charman, Major Professor

Date of Defense: June 19, 2015

This dissertation of Jonathan Nicholas Carbone is approved.

---

Dean Michael Heithaus  
College of Arts and Sciences

---

Dean Lakshmi N. Reddi  
University Graduate School

Florida International University, 2015

## ABSTRACT OF THE DISSERTATION

# IT MUST HAVE BEEN HIM: COHERENCE EFFECTS WITHIN THE LEGAL SYSTEM

by

Jonathan Nicholas Carbone

Florida International University, 2015

Miami, Florida

Professor Steve D. Charman, Major Professor

The present series of studies examine how jurors and public defenders evaluate different pieces of evidence and integrate them into a coherent conclusion within the context of a criminal case. Previous research has shown that in situations where both sides of the case are compelling, decision-makers nevertheless come to highly confident and polarized decisions, called coherence shifts (Simon, 2004). The present research sought to expand on coherence effects, improve upon the methodology of previous studies, and explore potential moderators of coherence. In Study 1, mock jurors ( $n = 306$ ) read about a criminal case and evaluated multiple pieces of evidence at various points throughout the case. Results indicated that participants exhibited pronounced coherence shifts (i.e., their evaluations of the various pieces of evidence (a) became more consistent as the case progressed, and (b) were evaluated in line with their initial leanings) using an improved methodology that randomized evidence order and evidence valence. Furthermore, participants' interim leanings of guilt or innocence biased their subsequent evaluations of ambiguous evidence. The direction and magnitude of participants' coherence shifts were predicted by their pretrial dispositions towards prosecution and

defense. Participants lacked awareness of how their perceptions of the evidence have shifted. Coherence shifts were not, however, moderated by asking mock jurors to justify their decisions, or by asking mock jurors to play devil's advocate while considering each piece of evidence, underscoring the pervasiveness of this cognitive bias. Study 2 examined whether actual public defenders experience coherence shifts and how those shifts relate to the plea bargaining process; however, no coherence shifts were observed. Study 3 examined whether the timing of the defense's presentation of their case could reduce coherence effects; results indicated that reading about the defense's case immediately after the prosecution's case (c.f. following a delay) marginally ( $p = .09$ ) reduced coherence effects among jurors who acquitted the defendant, suggesting one potential strategy to mitigate this bias.

## TABLE OF CONTENTS

CHAPTER	PAGE
INTRODUCTION .....	1
Broad Theories of Legal Decision-Making .....	1
Cognitive Coherence .....	5
Coherence Effects in relation to other findings within the Psychological Literature .....	9
Goals of the Present Study .....	14
Hypotheses .....	24
METHOD .....	30
RESULTS .....	38
Coherence shifts in evidence evaluation .....	38
Coherence shifts in evidence integration .....	42
Moderators of Coherence .....	45
Coherence effects among Switchers .....	51
DISCUSSION .....	53
STUDY 2 .....	56
Goals.....	56
Hypotheses .....	56
METHOD .....	58
RESULTS .....	60
DISCUSSION .....	62
STUDY 3 .....	64
Goals.....	64
Hypotheses .....	65
METHOD .....	66
RESULTS .....	69
DISCUSSION .....	72
GENERAL DISCUSSION .....	74
Coherence effects in jurors.....	74
Overview of decision-making in the legal system .....	91
Shift in focus .....	93

REFERENCES .....	95
APPENDICES .....	104
VITA.....	219

## INTRODUCTION

Psychologists are in the business of understanding and predicting human behavior and yet it remains very challenging to successfully predict complex decisions. Even the techniques of decision-making *experts*, such as trial consultants, are often ineffective at predicting verdicts when their techniques have been subjected to empirical scrutiny (Costanzo & Krauss, 2012). On the one hand, it is perhaps no surprise that human decision-making is so difficult to understand; even the most powerful computers do not begin to approach the processing speed or seemingly infinite storage capacities of the human brain (Schwartz, 2011). On the other hand, when so much of our society hinges on decision-making, including the legal system, economics, medicine, and so forth, it is ironic that decision-making is not better understood. In the legal system, decision-making occurs throughout almost every step of the process, ranging from criminal investigations, to plea bargaining among lawyers, to jury deliberations, and on through the appeals process.

### **Broad Theories of Legal Decision-Making**

There have been two predominant schools of thought in terms of legal decision-making: rationalists and critics (Simon, 2004). Rationalists argue that logic and deduction are the guiding principles of a decision and that decision-making proceeds from a given rationale to a given conclusion. Rationalists point to Bayesian reasoning as an example of how decision-making works. The idea behind Bayesian reasoning is that one should be able to numerically determine the value of a particular piece of information or evidence. Furthermore, when all of the evidence is summed up, if the total value of that evidence surpasses a certain threshold, then a decision will be reached accordingly.



Thus, a rationalist would suggest that one could examine a juror and establish a ‘likelihood of guilt’ threshold, whereby if the evidence surpasses that threshold, that juror will find the defendant guilty. For example, a given juror may have a ‘likelihood of guilt’ threshold of 90%, whereby if the evidence increases the odds, or more accurately, the perception of the odds, that the defendant is guilty beyond 90%, then the juror will find the defendant guilty. The baseline odds of this juror finding the defendant guilty begin at 1%. If the juror then hears evidence regarding an incriminating eyewitness testimony, which increases the likelihood of guilt by 30%, a dubious alibi, which increases the likelihood of guilty another 30%, and a positive fingerprint match, which increases the likelihood of guilt an additional 35%, then the total of 96% surpasses the threshold of 90% and the juror comes to believe the suspect is guilty.

Another way to conceptualize the rationalists’ standpoint is that:  $a + b = c$ . Here ‘a’ and ‘b’ represent evidence and ‘c’ represents a decision. Each piece of evidence, then, has a specific value and does not influence the decision *beyond* its value. The implication is that the value of ‘a’ is *independent* and does not affect the value of ‘b.’ The prominence of this viewpoint can be seen throughout the wording in the legal system, for example, in judicial instructions. Jurors are instructed to consider the value of evidence and to weigh the various pieces of evidence in order to come to a conclusion. Moreover, during the appeals process, the doctrine of harmless error states that even if there was a problem with a certain piece of evidence or if a procedural protocol was violated, that error would most likely not be important enough to have influenced the outcome of the case or the other aspects of the case (Scheb & Scheb, 2011). As will be plain by the end of this manuscript, that assumption is not true, evidence can have a prejudicial ‘snowball

effect.’ Even the merest mention that the defendant was a ‘bully’, during, say, pre-trial publicity, can entirely sway the outcome of trial (Simon, 2012).

Unlike rationalists, legal critics would suggest an alternate position: people are not rational decision-makers (Simon, 2004). The argument that critics put forth is that people come to a conclusion that considers factors that they are unaware of, such as biases, and then develop a post-hoc explanation for their conclusions. Critics could cite racial profiling in the legal system as evidence of their viewpoint. If an officer makes an arrest on the basis of a racial stereotype, the officer would never state, nor indeed even realize, that the arrest was made because of the suspect’s race. The officer would likely cite other reasons for the arrest, such as the suspect was behaving suspiciously, but in reality, the conclusion would be made using a factor (race) that the officer was unaware of on a conscious level.

There are aspects of the both the rationalist and critical viewpoints that are compelling, but, in isolation, there are also aspects of each standpoint that are problematic. A rough analogue of the distinction between rationalist and critical viewpoints concerns studies on explicit versus implicit decision-making (Dijksterhuis, Bos, & Nordgren, 2006; Dijksterhuis & Nordgren, 2006). Dijksterhuis et al. have studied explicit versus implicit decision-making via asking participants to choose between two cars, one of which is the better choice, considering factors like gas mileage, safety rating, insurance, and so forth. Some participants were asked to make their decision rationally and explicitly. These participants were told to break the decision-making task down into its component parts and list each aspect of the cars that they find appealing. However, other participants were encouraged to make their decision implicitly. Those participants

were given a distracting task and told that they would come to a conclusion later on. Participants who were distracted outperformed those who used an explicit, rational approach to making their decision. This finding supports the critics' view of decision-making in that non-conscious, implicit factors are integral to decisions.

However, Dijksterhuis and Nordgren (2006) found that implicit decision-making outperformed explicit decision-making only when the task was *complex*. In the complex task conditions, each car had over ten features that were being compared. In the *simple* task conditions, though, there were only four features being compared between cars. Here, participants who engaged in explicit decision-making outperformed participants who engaged in implicit decision-making. Thus, the simple-decision conditions of Dijksterhuis' study supported the rationalist view in that decision-making was improved by explicitly considering each independent factor of the decision.

Dijksterhuis' studies (Dijksterhuis, Bos, & Nordgren, 2006; Dijksterhuis & Nordgren, 2006) represent a concrete example of how, in the real world, both the rationalist and critical views of decision-making have some merit. Indeed, debating which of the two positions is right would likely lead to stalemate and a stagnant study of decision-making (Simon, 2004). Therefore, Simon has proposed that decision-makers engage in what he calls *coherence-based reasoning*, a process that incorporates both rationalist and critical views of decision making. For instance, he proposes that people are somewhat rational in their decisions, in that they evaluate and weigh various pieces of evidence. Simultaneously, though, coherence-based reasoning is also a largely non-conscious process, in that people's opinions and evaluations of that evidence tend to shift over the course of the decision, a shift that they are later unable to accurately recall

having made. Furthermore, the extent to which the decision-making process is non-conscious leaves room for biases and outside influences to affect the decision as well.

### **Cognitive Coherence**

**Theoretical background of Coherence.** Simon and colleagues (Holyoak & Simon, 1999; Simon, 2004; Simon, 2012; Simon, Pham, Le, & Holyoak, 2001) have proposed that when people are presented with a complex decision, a coherence process takes place whereby, because of shifting evaluations of evidence, the decision-maker begins to increasingly favor one side of the decision over the other. Importantly, coherence-based reasoning is a bi-directional process, in that a given rationale can lead to a conclusion, but emerging conclusions also creep backwards and influence the perception of one's rationale (see Table 1 for a comparison between rationalist, critical, and coherence-based views on decision-making). For example, if a juror hears exculpatory evidence from an eyewitness, that evidence could cause the juror to tentatively lean towards believing the defendant is not guilty. Subsequently, the belief that the defendant might not be guilty could skew interpretations of new evidence, such as an alibi.

Coherence is a constraint-satisfaction mechanism, whereby each aspect of the decision constrains the other aspects of the decision (Simon, 2004). For example, if a juror is told that fingerprint evidence points towards the defendant's guilt, then that knowledge of the fingerprint evidence 'constrains' the jurors' view of subsequent evidence that may arise, such as an alibi. If fingerprint evidence is perceived as incriminating, then, logically, the suspect's alibi cannot be true, and the juror will tend to become much more skeptical of the alibi. Alternatively, if a juror learns about an alibi

witness who saw the defendant far from the scene of the crime, then that evidence should also constrain the juror's understanding of the incriminating fingerprints. In this fashion, the perceptions of the various pieces of evidence come to cohere with one another.

Eventually, one side of the decision comes to dominate the other side, which results in a confident and comfortable decision. Moreover, although the juror may not have had an initially strong opinion about whether the defendant is guilty or not guilty, by the end of the coherence process, the juror's perceptions of the defendant and of the evidence end up polarized.

**Methodology of Coherence studies.** In order to demonstrate coherence effects, Simon conducted a series of studies (Holyoak & Simon, 1999; Simon, Pham, Le, & Holyoak, 2001; Simon, 2004) designed to track the shifting, or cohering, of opinions over the course of the decision-making process. Participants read a series of unrelated, abstract vignettes about evidence and then provided ratings of the probative value of that evidence. For example, the participants read a vignette about an eyewitness identification and were then asked to rate the likelihood that the identification was correct.

After participants had rated the abstract vignettes, they read about a trial. The evidence presented at the trial was designed to be complex and somewhat ambiguous as to whether the suspect was guilty or not guilty. Both sides of the case were reasonably compelling. Participants were told that they were playing the role of a young judge and would need to decide the case, but were instructed to withhold making a final decision until they had the opinion of a veteran judge. Participants read through the arguments from each attorney, which described the evidence in this case. Critically, this evidence was the same evidence that the participants rated earlier during the vignette portion of the

study, only now it was fleshed out with the specific names of people and situations involved in the case. Once the participants read through the evidence, they were asked to provide an initial leaning about the suspect's guilt or innocence and then they were asked to rate each piece of evidence. After participants provided their initial leanings, they expected to receive feedback from a veteran judge but were told that the judge would not be providing any feedback after all. Subsequently, the participants were then asked to render a verdict, rate their confidence in that verdict, and then provide post-decision ratings of all of the evidence one last time.

**Coherence results.** Simon (2004) found that participants' ratings of the evidence during the abstract vignettes phase of the study were largely uncorrelated with each other. For example, a participant's rating of 'how likely an eyewitness' identification was correct' did not correlate with the participant's ratings of 'how likely that it was that the suspect's fingerprints matched the fingerprints on the weapon.' However, over the course the study, which was a single session that lasted under an hour, participants' views of the evidence began to correlate more and more, painting an increasingly coherent picture of guilt or innocence. To examine this coherence process, Simon divided participants into those who acquitted the defendant and those who convicted the defendant. The predicted patterns of coherence were observed (see Figure 1). When looking at the ratings of the evidence by people who ultimately acquitted the defendant, for example, it was observed that as they progressed through the trial, evidence that suggested innocence became more believable, whereas evidence that suggested guilt became less believable. In short, the ratings of the various pieces of evidence began to cohere with one another. From a logical perspective, one's rating of how probative a

piece of evidence was in the abstract should not change dramatically when that same piece of evidence is presented again in the context of the trial; after all, the evidence during the abstract vignettes phase of the study is functionally identical to the evidence in the later phases of the study. However, the ratings of the evidence did change dramatically over the course of the trial, in such a manner as to cohere with the participants' initial leaning.

Another important finding from Holyoak and Simon's (1999) study is that, despite the complexity of the trial, nearly all participants indicated maximal or next-to-maximal confidence in their decisions. Their confidence is disturbing insofar as the study was designed to be ambiguous and, in actuality, about half of the participants found the defendant guilty whereas the other half found the defendant not guilty. In effect, coherence-based reasoning made the decision seem much more obvious than it actually should have been.

In sum, Simon and colleagues (Holyoak & Simon, 1999; Simon, Pham, Le, & Holyoak, 2001; Simon, 2004; Simon, 2012) found that: (a) individual pieces of evidence underwent a shift in value over the course of the trial, such that evidence that supported one's leaning became more believable and evidence that contradicted one's leanings became less believable, (b) groupings of evidence that supported one side or the other became increasingly correlated in accordance with a participant's eventual verdict, and (c) all participants were highly confident in their decision despite the challenging and ambiguous nature of their task.

## **Coherence Effects in relation to other findings within the Psychological Literature**

**Confirmation bias and attitude polarization.** Given the similarities between coherence-based reasoning and other psychological phenomena, it is important to understand the relationship between them. For instance, confirmation bias (Wason, 1968) is a common psychological bias concerning information evaluation and decision-making. In fact, some researchers have argued that it is a ubiquitous phenomenon, affecting people in essentially any line of work (Nickerson, 1998). In the legal realm, it is often used to explain tunnel vision on the part of investigators and erroneous decisions on the part of judges and jurors.

The confirmation bias occurs when people evaluate incoming evidence or information in accordance with their prior beliefs, resulting in biased conclusions. Specifically, any information that supports an already-established belief or hypothesis tends to be viewed as especially compelling. However, information that contradicts established beliefs tends to be subject to intense scrutiny and ends up mostly disregarded.

Confirmation bias can lead to attitude polarization. Lord, Ross, and Lepper (1979) evaluated people who harbored pro- or anti-death penalty attitudes. They presented participants with a series of fictional studies, some that supported the effectiveness of the death penalty, and some that failed to support its effectiveness. Participants were asked to evaluate each of the studies. The authors found that the evaluations of the studies were highly dependent on participants' initial attitudes. They found that participants who were pro-death penalty, for example, believed that the studies that supported the death penalty were valid, well-conducted, and were derived from representative areas of the country. In contrast, they believed the studies that described



the drawbacks of the death penalty were invalid, of shoddy methodology, and unrepresentative. Participants who were anti-death penalty evaluated the exact same set of studies but came to the opposite conclusions. As a consequence, being exposed to the exact same contradictory information caused participants' initially moderate death-penalty beliefs to become even more polarized or entrenched.

Confirmation bias and attitude polarization occur throughout the legal system. For example, if a judge has a repeat offender show up at court, that judge may interpret innocuous or ambiguous actions on the part of the defendant as implying guilt because of his criminal history. Alternatively, if juror A believes the defendant is not guilty, that juror will tend to discount reasonable counter-arguments from juror B during deliberations.

Confirmation bias (Wason, 1968; Nickerson, 1998) and attitude polarization (Lord, Ross, & Lepper, 1979) are compelling phenomena. However, they both presume that decision-makers have a belief or a hypothesis prior to encountering some new piece of information. In many real world situations, people may have no such strong prior opinions. Jurors, for instance, should tend to approach a trial with few if any prior beliefs concerning the defendant's guilt. Nonetheless, coherence effects (Simon, 2004) demonstrate that people can start off with neutral attitudes and wind up extremely polarized. Specifically, participants in Simon's coherence studies did not demonstrate strong attitudes during the first phase of the study, which indicates a lack of prior beliefs. However, they did wind up strongly polarized by the end of the study. This occurs because participants begin to form emerging conclusions throughout the process of evaluating the various pieces of evidence (even despite instructions to avoid doing so),

which then further constrain the interpretation of other evidence, resulting in coherence. Simon argues that a great deal of the polarization that decision-makers experience is not genuinely based on evidence, but rather is an artifact of the cognitive *process* of coherence, whereby the mind is motivated to arrive at a simple, confident conclusion.

Coherence effects (Simon, 2004) can be thought of as a broad theoretical umbrella under which more specific effects, like confirmation bias and attitude polarization, would be subsumed. Coherence describes the process of polarization through which a decision-maker begins a decision task from a relatively neutral standpoint, but winds up polarized in his or her conclusions, whereas confirmation bias and attitude polarization describe specific instances in which an already-polarized belief taints the perception of new information and remains unchanged by contradictory information.

**Story model theory.** Story Model Theory (Pennington & Hastie, 1992) states that jurors come to decisions via a three-stage process: (a) they construct competing stories of the incident, which are based largely on their own personal biases and beliefs, (b) they become aware of the verdict options, and (c) they eventually choose the story that has the most explanatory power. Critically, decision-makers bring to bear their biases, opinions and experiences, perhaps non-consciously, on the construction of stories. The stories, then, are integral to their understanding of the case and their eventual decision. Pennington and Hastie found that jurors who heard testimony from witnesses in a story-driven order were 70% likely to convict, whereas witnesses who heard the same testimony in a different order were only 30% likely to convict. Thus, when participants were able to easily construct a compelling story of guilt (i.e., because it was presented in

a chronological order), the ‘guilty story’ was chosen over the ‘not guilty story’ and the participants then selected their verdicts accordingly. Another example of competing stories is that Pennington and Hastie looked at the qualitative reports of mock jurors who reviewed the evidence in the OJ Simpson trial. They found that jurors who developed a narrative around ‘racial profiling and police misconduct’ were likely to find OJ not guilty, whereas jurors who focused on the ‘evidence and the victim’ were likely to find OJ guilty.

In some ways, Story Model Theory and coherence-based reasoning make similar predictions. Specifically, as the decision develops (either because of increasing coherence or because of the increasingly compelling story that is created), it comes to constrain the interpretation of other evidence. However, Simon and colleagues’ studies (Holyoak & Simon, 1999; Simon, Pham, Le, & Holyoak, 2001; Simon, 2004; Simon, 2012) directly measured ratings of the evidence, whereas Pennington and Hastie (1998) evaluated qualitative accounts of verdict choice. As such, coherence effects are better able to explain the bi-directional *process* of decision-making whereby evidence evaluations lead to initial leanings, and initial leanings feed backward to affect evidence evaluation, whereas Story Model Theory essentially describes the *result* of the decision-making process, that one side of the story has come to dominate the other side. Story Model Theory, like confirmation bias and attitude polarization, fits nicely under the theoretical umbrella of coherence effects.

**Cognitive Dissonance.** Cognitive dissonance (Festinger & Carlsmith, 1959) is a powerful negative state that is created when one’s behaviors are inconsistent with one’s attitudes. People are highly motivated to reduce this negative state; consequently, when

dissonance arises, people often alter their attitudes or their behavior in such a way as to bring the dissonant elements of the situation into harmony. For example, when participants were asked to perform an incredibly boring task and then asked to tell the next participant that the task was fun, the inconsistency between the participants' attitudes (the task was boring) and behaviors (telling the next participant it was fun) produced dissonance. To resolve the dissonance they experienced, participants actually changed their attitudes and came to convince themselves that the task was more fun than it really was.

Interestingly, many of the findings in Simon's studies would be predicted by dissonance. Holyoak and Simon (1999) found that post-decision ratings of evidence closely resembled the decision that had been reached. For example, participants who found the defendant guilty rated the evidence as highly inculpatory. Dissonance predicts that *after* a decision-maker arrives at a decision, he or she wants to avoid experiencing doubt and thus is motivated to bring any conflicting components of the decision in line with the final decision. Thus, dissonance would have predicted the same result. For example, if a juror finds a defendant guilty, he or she will be motivated to view exculpatory evidence as insubstantial, in order to prevent dissonance.

Critically, dissonance theory states that any changes in attitudes about the components of a decision will occur *after* the decision has been reached, in an attempt to justify one's decision; without a decision having been made, there is nothing to justify.

In contrast, coherence effects can *precede* the decision. To demonstrate this, participants in Simon (2004) provided their 'initial leanings' towards a verdict, prior to actually rendering a verdict. He showed that coherence effects still occurred at that stage,

despite the fact that participants had not actually made a final decision. In a follow-up study, participants were not asked to make a decision at all. Rather, the study was phrased as a memory task. Nonetheless, coherence effects were still observed, even though participants were simply processing the information while trying to memorize it. Dissonance would not have predicted that opinions concerning the evidence would correlate in the absence of a decision, because participants had no motivation to avoid dissonance.

Thus, although some of the findings from cognitive dissonance and cognitive coherence do coincide, the mechanisms and purpose behind the two phenomena appear to be different. Simon (2004) found that coherence arises simply as a way of understanding and processing information; whereas dissonance necessarily has a motivational component to it, such as restoring self-esteem (Gilbert & Malone, 1995).

### **Goals of the Present Study**

There are three primary goals of the present study: (1) to expand the general research on coherence effects by examining how coherence-based reasoning can affect the perception of ambiguous evidence, (2) to resolve important methodological flaws in past research in this area, and (3) to investigate the potential moderators of coherence effects.

**Expanding the research on Coherence Effects.** One of the primary goals of the present study is to expand the knowledge base regarding coherence effects and decision-making in general. For example, decision-makers in coherence studies have been shown to polarize towards either guilt or innocence over the course of a trial, but it still remains to be seen whether that polarization can taint their perception of novel, ambiguous

evidence. In other words, as participants lean towards either conviction or acquittal, does their leaning act as a *self-generated* bias that can taint the perception of subsequent, ambiguous evidence? If so, that would be preliminary evidence that *self-generated* bias that arises because of coherence shifts is an important consideration alongside *other-generated* bias, which has been firmly established in the legal psychological literature.

The majority of legal psychology studies thus far have looked at specific evidence evaluation effects whereby one piece of evidence has a corrupting influence on other pieces of evidence (Ask, Rebelius, & Granhag, 2008; Hasel & Kassin, 2009; Charman, Gregory, & Carlucci, 2009; Kassin, Bogart, & Kerner, 2011). Charman et al. found that mock investigators' beliefs in the guilt of a suspect changed their perceptions of the similarity between the suspect and a composite of the perpetrator. Hasel and Kassin found that a false confession could dramatically bias an eyewitness identification, such that witnesses were more likely to change their previous identification if they learned that a non-identified lineup member confessed to the crime. Additionally, Kassin et al. replicated similar findings in the real world by conducting an archival study on DNA exonerations obtained by the Innocence Project. They found that when cases contained a strong piece of evidence, such as a false confession, there were also likely to be other instances of eyewitness misidentifications and forensic malpractice. Kassin et al. argued that their findings demonstrate a contamination process whereby one strong piece of evidence taints how other evidence is collected, administered, and perceived.

Recently, Charman (2013) differentiated between two similar, but distinct, processes that may be affected by coherence effects: evidence evaluation and evidence integration. Evidence evaluation refers to how a decision-maker perceives a given piece

of evidence and to what extent the decision-maker finds that piece of evidence probative. Evidence integration, though, refers to how the various pieces of evidence are related to each other and to how all of the evidence combines to form a decision. Charman described how there has been an over-emphasis on specific evidence evaluation research at the expense of researching the mechanisms behind evidence integration. Indeed, an emerging body of literature is now pointing to differences between the two mechanisms. Charman, Carbone, Kekessie, and Villalba (under review) found that the order in which evidence was presented had a differential impact on the evaluation of the evidence versus the integration of the evidence. Specifically, this study showed that there was a primacy effect for evidence evaluation (i.e., the first piece of evidence presented influenced the evaluation of subsequent pieces of evidence), but there was a recency effect for evidence integration (i.e., the last piece of evidence was most influential in participants' final verdicts, which required the integration of all evidence).

Distinguishing between evaluation and integration contributes to a more nuanced dialogue regarding how people make decisions and whether coherence shifts occur similarly with respect to evidence evaluation and evidence integration. As such, the present study will examine both evidence evaluation, such as the ratings of how inculpatory or exculpatory a given piece of evidence is, and evidence integration, such as judgments of the defendant's overall likelihood of guilt.

**Methodological issues with prior Coherence Research.** Some important methodological issues with the previous coherence studies (Holyoak & Simon, 1999; Simon, Pham, Le, & Holyoak, 2001; Simon, 2004) need to be addressed before recommendations to the legal system can be made. To begin, the past coherence studies

have asked participants to rate arguments from attorneys about the case, rather than rating actual evidence. The current studies expands the research on coherence effects to include perceptions and ratings of different pieces of evidence, instead of attorneys' arguments.

Additionally, the order of the attorneys' arguments were not randomized in the coherence studies. As Charman et al. (under review) found, evidence order can systematically impact perceptions of evidence and judgments of the defendant's likelihood of guilt. Additionally, if a study uses the same evidence in the same order across all participants, then the results of the study may be limited to situations in which the type and order of evidence is similar to that of the study. The present studies randomize the presentation order of evidence to ameliorate those concerns.

Furthermore, valence (inculping or exculpating) of the attorneys' arguments is not randomized coherence studies. This is problematic because it could be that participants who found the defendant guilty were those who were naturally compelled by attorney argument x, y, and z, instead of being determined by coherence effects or other variables. For example, a juror naturally finds motive to be a compelling factor, then if the prosecution makes the argument that the defendant had a clear motive, that will be especially important to that juror. Thus, that juror may shift towards guilt because of the value he/she places on motive, instead of coherence shifts. As such, it is important to randomize the valence of evidence so that jurors who find a specific type of argument or evidence to be compelling will not just be drawn to the side of the case that presents that type of evidence. It will be a much stronger statement of the power of coherence effects if the present study is able to show that participants still arrive at polarized conclusions



despite whatever combination of inculpatory and exculpatory evidence they receive, and despite the order in which they receive that evidence.

**Exploring potential moderators of Coherence Effects.** Coherence effects (Holyoak & Simon, 1999; Simon, et al., 2001; Simon, 2004; Simon, 2012) go a long way toward explaining decision-making; however, only a few moderators of coherence have been discovered thus far. For example, Simon (2004) found that death penalty attitudes moderate whether jurors are likely to shift towards conviction or acquittal. The present series of studies investigates more potential moderators.

***General pro-prosecution and pro-defense attitudes.*** When jurors exhibit coherence shifts, they naturally polarize towards either guilt or innocence. But it still remains unclear as to what factors cause some jurors to begin leaning toward guilt or innocence in the first place. The present study predicts that the direction toward which jurors exhibit coherence shifts will be strongly related to a juror's general pro-prosecution or pro-defense attitudes. The Juror Bias Scale (Kassin and Wrightsman, 1983) is a validated measure of general predisposition towards conviction or acquittal that will be administered at the end of the study. If jurors do shift in line with their general pro-prosecution and pro-defense attitudes, this would suggest why coherence effects are difficult to reduce and why jurors polarize so strongly. This finding could also have implications for voir dire, in that jurors' pre-trial prosecution/defense bias could be able to predict the direction of their coherence shifts that will occur over the course of the trial.

***Anticipatory justification as a moderator of coherence effects.*** One potentially problematic discrepancy between Simon's coherence studies and the real world is that

real decision-makers often have to justify their decisions to other parties. For example, an investigator must justify his or her arrests by charging the suspect with a crime and, later, providing the district attorney with the evidence that supports the arrest. Public defenders must justify their assertions about their client's innocence to the judge and the jury. Additionally, jurors will often have to justify their opinions to other jurors, in order to try and sway them. A critical question then is whether people's belief that they will have to justify their decision will affect their tendency to engage in coherence-based reasoning.

Despite all the situations in which justifying decisions might matter, there has been relatively little research on the topic; what has been conducted suggests that the belief that one will have to justify a decision changes the decision-making process. Pennington and Schlenker (1999) found that the justification behind decisions can change based on who the expected audience of the decision will be. Participants in their study were asked to adjudicate a student honor court dispute involving a student who had cheated. When participants were told they were going to meet with the student who had cheated, they recommended less severe sentences than when told they would meet with the student's professor. They content coded the justifications behind the sentencing and found qualitative differences based on who the anticipated audience would be. Thus, this study suggests that having to justify one's decision to a particular party can influence the way that a decision is reached in anticipation of an interaction with said party.

Similarly, the pressure to justify one's decision has been shown to increase the amount of information utilized in explaining the decision. Huber and Seiser (2001) constructed a task where participants had to choose one of six candidates to become the

head of a corrective home for juveniles. They found that participants who were pressured to justify their decisions mentioned more of the information relevant to the decision than did those who were simply asked to explain their decision.

The present study tests whether instructing legal decision-makers to justify their decisions will moderate coherence effects. Specifically, decision-makers will be informed at the *beginning* of the study that they will need to justify their decision. Thus, the present study is examining not just the role of justifying one's decision, but also of *anticipating* having to justify one's decision. Lerner and Tetlock (1999) refer to anticipatory justification as 'predecisional accountability' and they found that the effects of accountability were more dramatic when the decision-makers are told in advance of the formation of any opinions that they will be accountable for the eventual decision. Moreover, Lerner and Tetlock reviewed a large body of literature and found that the effects of predecisional accountability on decisions varies greatly depending on the type of decision. For example, having to justify or be accountable for one's decision can reduce reliance on heuristic information and improve systematic processing, or it can amplify bias by increasing sensitivity to irrelevant cues (e.g. age and stereotypes in hiring decision), or it can have no effect on the decision task at all, as is the case with trying to get participants to use base rate information. As Lerner and Tetlock's review demonstrates, having to justify one's opinion can have very mixed results depending on the properties of the task. Their review, for example, found that increasing decision-makers' feelings of accountability also increased their aversion to ambiguous decision choices. That finding has implications for the present study in that decision-makers who anticipate having to justify their decision should be motivated to avoid ambiguity, which

they could do by shifting their perceptions to view the case as ‘black and white’ instead of ambiguous. In other words, in order to avoid the uncomfortable feeling of ambiguity, decision-makers in the ‘justify’ conditions should experience exacerbated coherence shifts as they come to perceive the case as lopsided.

The impression management, or self-presentation, literature provides additional support for the prediction that having to justify one’s decision will exacerbate coherence effects. In order to present one’s self as a competent and consistent decision-maker, people who anticipate having to justify their decision will be motivated to arrive at firm and confident conclusions so as to appear competent (Swann, 1985; Baumeister & Bushman, 2008). As such, a juror who finds the defendant guilty and must justify his or her decision should downplay the aspects of that decision that might contradict that conclusion in order to appear socially favorable. In the present study, when jurors are told to anticipate having to justify their decision, they are also informed that they will be justifying it to a group of lawyers who have asked for their help preparing for a trial (which relates to the cover story for the present experiment). Thus, in order to present themselves as competent decision-makers, jurors who anticipate having to justify their decision should experience stronger coherence shifts in order to arrive at consistent and firm conclusions.

***Devil’s Advocate instructions as a moderator of coherence effects.*** The polarization between how decision-makers view the evidence that supports their side of the case and the evidence that does not support their side of the case represents an *artifact* of the coherence process, rather than a difference in the quality of the evidence. In that sense, the magnitude of coherence effects can be conceptualized as an error or bias that

needs to be reduced. For example, if coherence effects lead jurors to perceive evidence as more incriminating than it actually is, that could push jurors to find the defendant guilty beyond a reasonable doubt, even though the reasonable doubt standard is not really being met on the basis of the merits of the evidence. Indeed, it is alarming that in the coherence studies mock jurors did find the defendant guilty about 50% of the time, despite the study deliberately being designed to be compelling for the defense as well as for the prosecution (Simon, 2004).

Simon (2004) was able to somewhat mitigate coherence effects by asking participants to ‘consider that the other side of the case is actually right.’ He asked them to consider the other side of the case towards the end of the study, but before the final decision. He found that asking them to consider the other side of the case reduced coherence effects by about half. However, he also noted that while the coherence effects he observed were significant, they were not always extreme, and thus it was hard to observe a decrease in their effect.

The present study hopes to improve upon reducing coherence effects in two ways. First, the ‘Devil’s Advocate’ instruction will be given at the beginning of the study (as opposed to near the end of the study) and then again as decision-makers review each piece of evidence. This should increase the salience of the instruction and strengthen its effect.

Second, participants will be told to consider how each piece of evidence they are presented with may be flawed or problematic (as opposed to considering the opposite of one’s final conclusion about the case). For example, if jurors are presented with

fingerprint evidence, then jurors will be urged to consider the issues and flaws inherent in that piece of evidence, such as ‘there was only a partial print.’

The theoretical rationale behind the prediction about the reduction of coherence effects because of the ‘Devil’s Advocate’ instructions is that self-generated counterarguments are usually much more compelling than other-generated counterarguments (Carlsmith & Aronson, 1963). Studies on confirmation bias, for example, have examined the concept of ‘disconfirmation,’ whereby participants are asked to consider how they could be wrong. Disconfirmation instructions are usually highly effective at reducing confirmation bias and, in some cases, can even eliminate it (Edwards & Smith, 1996).

The literature on belief perseverance (Anderson & Sechler, 1986) also supports the predictions about the ‘Devil’s Advocate’ manipulation. Belief perseverance states that once a belief has been established, it will persist even when people are told that the belief was inaccurate all along. This occurs because people self-generate reasons to explain that belief; thus, even when the belief itself has been discredited, their reasons for that belief still exist. However, when people are asked to consider why their belief might be wrong, the tendency for their beliefs to persist after having later been discredited is reduced. For example, if people are told that risky firefighters make the best firefighters, but are later told that statement is untrue, the belief will still tend to persist, on account of the self-generated reasons participants have come up with to support the original belief. However, when participants are explicitly instructed to consider reasons why the initial belief is wrong (e.g., that cautious firefighters are better because they stand a greater chance of surviving and rescuing potential victims), then a later discrediting of the

erroneous belief succeeds in eliminating that belief. Similarly, it is expected that when participants in the present study self-generate specific ideas about how each piece of evidence is flawed and problematic, those self-generated ideas will reduce coherence effects.

***Evidence Order and Coherence Effects.*** The present study will randomize whether participants first receive three pieces of evidence from the prosecution or three pieces of evidence from the defense. Previous research by Charman et al. (under review) has demonstrated that the most recent evidence a juror receives has the strongest impact on likelihood of guilt. It will be important to evaluate whether coherence effects are moderated by evidence order because that is an area of the legal system where practical recommendations could potentially be made. In many ways, Charman et al.'s study is a precursor to the present series of studies. As such, the full manuscript is included here (see Appendix A), as it expands the discussion of evidence evaluation and evidence integration and the extent to which the presentation order of evidence moderates that relationship and coherence effects.

## **Hypotheses**

### **Hypothesis 1 – Coherence Effects and evidence evaluation**

It is expected that, in keeping with previous research (Holyoak & Simon, 1999; Simon, Pham, Le, & Holyoak, 2001; Simon, 2004; Simon, 2012), jurors will experience escalating coherence shifts throughout the course of coming to a decision. The present study seeks to expand upon this prior work by examining similar coherence effects when jurors evaluate evidence (as opposed to arguments), and when the presentation order and valence of the individual pieces of evidence are randomized.

**Hypothesis 1a. – Valenced evidence shifts.** Each juror will be presented with six pieces of evidence, which will be randomized in terms of whether they have an inculpatory valence (supporting the prosecution’s case) or an exculpatory valence (supporting the defense’s case). It is predicted that although jurors will initially have moderate perceptions of the inculpatory and exculpatory values of the evidences, their perceptions of the evidence will shift and become progressively more polarized in accordance with whether they are leaning towards a verdict of either guilty or not guilty. Specifically, the evidence that supports their initial leaning toward either guilt or innocence will be perceived as reliable and believable, whereas the evidence that contradicts their initial leaning will be increasingly perceived as less reliable and less believable. Evidence ratings will be measured at three different times throughout the trial.

**Hypothesis 1b. – Ambiguous evidence shifts.** In addition to the six pieces of valenced evidence, jurors will also be presented with two pieces of neutral evidence that were stipulated by both the prosecution and the defense. These two pieces of evidence will be presented after the first phase of the trial is finished and jurors have rated the other evidence and indicated their initial leanings towards guilty or not guilty. It is predicted that the perception and ratings of the ambiguous evidence will be ‘tainted’ by jurors’ pre-formed notions of guilt or innocence. For example, a juror who is leaning towards guilt at the interim of the trial will see more inculpatory value in the ambiguous evidence than a juror who is leaning towards innocence.

**Hypothesis 1c. – Lack of insight into coherence shifts.** At the end of the study, participants will be asked how much they thought their ratings of the evidence shifted



over the course of the study. Additionally, they will also be asked to try and *recall* the first ratings they provided at the beginning of the trial for each piece of evidence. On the basis of prior research showing that people are generally unaware of their own biases (e.g., Wilson & Brekke, 1994), it is predicted that participant's recollections of their original rating will not be accurate and that their recollections of the original ratings will be more aligned with their post-trial evidence ratings. This would demonstrate that not only are participants susceptible to coherence biases, but that they are unaware of these biases.

### **Hypothesis 2 – Coherence effects and evidence integration.**

A substantial amount of research in legal psychology has focused on how the evaluation of evidence can be tainted by bias. However, very little research has focused on how bias affects evidence integration, which refers to how various pieces of the evidence are combined to form a final conclusion. The following hypotheses refer to the relationship between the various pieces of evidence and to how perceptions of the evidence are related to jurors' final conclusions about the case (such as overall likelihood of guilt, verdict, and verdict confidence).

**Hypothesis 2a. – Increase in correlations between evidence.** The present study expects to see progressively stronger correlations among the various pieces of evidence as the trial proceeds. For example, a juror who eventually renders a guilty verdict may initially exhibit only weak correlations among the different pieces of evidence evaluated when presented in the abstract. But as the trial progresses, the inter-item correlation between the pieces of evidence should grow stronger as they cohere, such that the various pieces of evidence come to represent a singular 'construct' of guilt. If this effect is

found, it would speak to the bi-directionality of decision-making, whereby the evaluation of evidence influences the decision outcome that one is leaning towards, and, critically, the potential outcome ‘creeps backwards’ to affect the perception of evidence, resulting in increased coherence among the evidence (Simon, 2012).

**Hypothesis 2b. – Coherence shifts in likelihood of guilt.** It is expected that jurors’ perceptions of the defendant’s overall ‘likelihood of guilt’ will shift from the interim of the trial to the end of the trial, in correspondence with their verdict. Additionally, the likelihood of guilt will shift in the same direction as jurors’ ratings of evidence.

**Hypothesis 2c. – Coherence shifts and verdict.** In keeping with the general findings of coherence effects (Simon, 2004), it is predicted that there will be a significant difference between the average ratings of evidence dependent on whether a participant renders a guilty verdict or a not guilty verdict. For example, if a juror eventually renders a not guilty verdict, than that juror should have perceived the evidence as much more exculpatory than a juror who voted guilty.

**Hypothesis 2d. – Coherence shifts and confidence.** As jurors progress throughout the trial, it is predicted that their confidence will shift and increase, regardless of which verdict a juror eventually renders. As a meta-judgment, it is also expected that confidence will be highly correlated with other final conclusions about the case, such as likelihood of guilt.

### **Hypothesis 3 – Moderators of coherence effects.**

If psychologists are to eventually make recommendations to the legal system regarding coherence effects, it is important to discover moderators of coherence shifts.

**Hypothesis 3a. – Coherence shifts and pro-prosecution/pro-defense attitudes.**

Coherence shifts cause jurors to polarize towards one end of the decision or the other of their own accord. But it still remains unclear as to why some jurors begin to lean toward guilt whereas others lean toward innocence. The present study predicts that the directions jurors shift will be strongly related to a juror's general pro-prosecution or pro-defense attitudes. The Juror Bias Scale (Kassin and Wrightsman, 1983), a validated measure of general predisposition towards conviction or acquittal, will be administered at the end of the study. It is expected that there will be strong correlations between jurors' pro-prosecution/pro-defense attitudes and their: ratings of the evidence across the three points in time, the direction in which their coherence shifts occur (i.e., stronger pro-prosecution attitudes should be related to the tendency to shift towards guilt), their perceptions of the defendant's likelihood of guilt, the magnitude of those shifts, and the ratings of the ambiguous evidence.

**Hypothesis 3b. – Anticipatory Justification.** It is predicted that participants who anticipate having to 'justify' their decisions will exhibit *stronger* coherence effects when evaluating and integrating the evidence than participants in the control condition or the 'Devil's Advocate' condition. For example, the average ratings of the evidence among participants in the 'justification' condition will be more extreme (in the direction of their leanings) than among participants in the other conditions. In terms of evidence integration, it is expected that the correlation between the evaluations of the various pieces of evidence will be stronger in the 'justify' condition than it is in other conditions. Furthermore, the correlation between the defendant's 'likelihood of guilt' and the average ratings of the evidence will be strongest in the 'justify' condition. The impression

management literature (Swann, 1985; Baumeister & Bushman, 2008) provides a solid theoretical foundation for this prediction: in order to appear consistent and competent, decision-makers who anticipate having to ‘justify’ their decision will experience more coherence effects.

**Hypothesis 3c. - ‘Devil’s Advocate.’** It is predicted that participants who receive Devil’s Advocate instructions (who are asked to consider and write down the flaws, issues, and problems with each piece of evidence they are presented) will show *weaker* coherence effects. The reduction in coherence effects is expected to occur because Devil’s Advocate participants should be particularly vigilant in noticing the ways in which the evidence could be construed as inculpatory *or* exculpatory. For example, a match between two fingerprints could be construed as inculpatory information, but the realization that the match was based on only a partial, smudged fingerprint could be construed as much less inculpatory. It is therefore expected that participants in the Devil’s Advocate condition will show less extreme ratings of evidence compared to participants in the other conditions. Furthermore, compared to the other conditions, it is expected that there will be a lower correlation between the different pieces of evidence and a lower correlation between the average ratings of evidence and the defendant’s ‘likelihood of guilt.’ These predictions are based on the literature on disconfirmation (Edwards & Smith, 1996) and belief perseverance (Anderson & Sechler, 1986), which state that self-generating ideas about how a piece of evidence or an argument could be inaccurate will bring one’s position closer to neutrality.

**Hypothesis 3d. – Coherence and order effects.** The order in which jurors read about the prosecution’s evidence and the defense’s evidence is randomized. Some jurors

will read about three pieces of the prosecution's evidence first, followed by three pieces of the defense's evidence, whereas other jurors will read about the three pieces of the defense's evidence, followed by three pieces of the prosecution's evidence. It is predicted that a juror will be more likely to render a verdict for whichever side of the case presents their evidence most recently, or second. This prediction is derived from Charman et al.'s (under review) results, which found that a strongly valenced piece of evidence had a higher impact on decision outcome when it was presented second rather than first, in relation to an ambiguous piece of evidence.

#### **Hypothesis 4 – Coherence among 'switchers.'**

Although the vast majority of jurors who begin to lean towards either guilt or innocence should render a final verdict consistent with that initial leaning, Simon (2004) found that some people switch from an interim verdict of guilty/not guilty to the opposite conclusion at the end of the trial. It is predicted that these 'switchers' will not undergo the typical coherence shifts across the span of the trial. It is also predicted that switchers will have less confidence in their decisions than non-switchers.

### **METHOD**

#### **Design Overview**

The present study measured coherence shifts in evidence ratings and other judgments over the course of a trial as well as investigated the effect of moderators (pro-prosecution/pro-defense attitudes, having to justify one's decision, playing Devil's Advocate, and trial order) on coherence shifts.

The study used a 2 (Justify instructions: anticipatory justification vs. no anticipatory justification) X 2 (Devil's advocate instructions: Devil's advocate vs. no

devil's advocate) x 2 (Decision Outcome: Guilty vs. not guilty) between-subjects design. The study progressed chronologically as follows: Participants read and rated abstract evidence vignettes, they were presented with a case summary, they read about 3 pieces of inculpatory evidence and 3 pieces of exculpatory evidence (in a randomized order), they provided interim leanings of guilty or not guilty and evaluated the 6 pieces of evidence, they were presented with and evaluated 2 ambiguous pieces of evidence, and finally, participants rendered a final verdict and provided post-decision evaluations of all of the evidence. Participants were randomly assigned to one of the four instruction conditions and their decision outcome was a self-selected factor. Participants in the 'justify' conditions were told that they should anticipate having to eventually defend and explain their decision to the lawyers/researcher; participants in the 'Devil's Advocate' condition were told that they will have to write down the flaws, issues, and problems with each piece of evidence as it was presented to them. Participants in the 'justify and Devil's Advocate' condition received both sets of instructions. Participants in the control condition received neither instruction. Participants provided ratings of the evidence on four occasions: before the trial begins, during the interim of the trial, again after they have arrived at a verdict, and finally participants were asked to try and recall their original pre-trial ratings of the evidence. The extent to which participants' opinions of the trial evidence shifted and became polarized was evaluated over the course of the study via ratings of evidence and correlations among those ratings. Participants were asked to withhold making a final decision until all of the evidence has been presented. See Figure 2 for a summary of the study's procedure.

## **Participants**

Three hundred and six participants (90 males, 207 females, and 9 who declined to respond to gender; ages 14 – 52, mean age = 22 years; 60% Hispanic, 15% African American, 14% Caucasian, 5% other, 3% Asian) were recruited from a research pool of undergraduate students at a large southeastern university in the United States. Participants received extra credit in their courses for participating in the study.

## **Materials**

The materials for the case loosely followed an Innocence Project (2014) case that eventually resulted in a DNA exoneration. The case of *Tennessee v. House* involved the conviction of Paul House for the murder of a neighbor, Carolyn Muncey. This case was selected as a template for the current study's case because it involved a myriad of compelling factors for both sides of the case, including: forensic evidence, eyewitness testimony, character evidence, alibis, et cetera. The materials for the present study were deliberately developed to be a complex and ambiguous case in which either decision, guilty or not guilty, is not overwhelmingly compelling. Each participant received three pieces of inculpatory evidence, three pieces of exculpatory evidence, and two ambiguous pieces of evidence. The inculpatory and exculpatory evidence was counter-balanced; meaning, some participants may have received inculpatory eyewitness testimony ("I saw him do it"), whereas others may have received near-identical exculpatory eyewitness testimony ("I saw someone else do it"). Furthermore, the order of evidence was randomized, except that the two ambiguous pieces of evidence were always presented in the second phase of the trial, after initial leanings. The order of the two ambiguous pieces of evidence was also randomized. The six pieces of valenced evidence include

shoe-prints, eyewitness testimony, fingerprints, polygraph results, evidence of motive, and traffic camera footage. The two pieces of ambiguous evidence were phone records and the defendant's questionable alibi (see Appendix H). There was no DNA evidence or confessions, as the impact of such pieces of evidence is potentially overwhelming.

It is important to note that each participant essentially experiences a different trial based on their randomization to different combinations of inculcating and exculpating evidence. The benefit of this is that any coherence effects and the effects of the independent variables will transcend the type and valence of the evidence.

### **Recruitment**

Participants signed up for the study under the guise that local lawyers, who are graduates of the present study's university, have asked the legal psychology department to help them prepare for an upcoming murder trial (see Appendix B for recruitment document). Participants were told that their feedback would be provided to the lawyers in order to help them understand their case. The ruse was expected to increase the odds of participants taking the task seriously and it provided an excuse for why, in some conditions, the participants were asked to justify and explain their decisions. As such, participants were recruited through the use of a mild deception, however, that deception should have been innocuous and it was a necessary guise for the study to proceed. Participants began the study by filling out the informed consent form and, in so doing, they were reminded why they have been asked by the lawyers to participate in this study.

### **Procedure**

**Abstract evidence ratings.** The mock jurors began the study by evaluating some initial questionnaires (see Appendix C). These questionnaires were separate from the rest



of the study and were tied to the participants' other surveys via their participant number. The only items of interest on this survey were their ratings of certain pieces of evidence that would show up again later, during the trial. Everything else on the survey, such as other evidence or policy opinions, served as a distraction. Participants were given brief scenarios describing the details of a piece of evidence, for example, a polygraph examination. Then they were asked to rate the degree to which that evidence implies guilt or implies innocence for the suspect. These evidence ratings serve as a baseline for how probative participants perceive certain pieces of evidence to be. Critically, none of the evidence presented at this point is part of any cohesive narrative; they are simply stand-alone ratings of evidence in the abstract. No proper names were included; the evidence vignettes were simply scenarios.

**Instructions.** Participants began the next phase of the study by going over some important instructions that outlined their role as a juror in this trial (Appendix D). At the end of the instructions, in the experimental conditions, there was an extra paragraph that contains the manipulation.

**'Justify' manipulation.** Participants read the following extra paragraph at the end of the instructions section:

At the conclusion of the trial you will be asked to justify your decision to the lawyers (in writing) and to the researcher (in person) who is currently running your study. Be prepared to discuss your eventual decision and to cite the aspects of the case that support your view. When thinking about justifying your decision, it may help you to imagine that you will be trying to convince someone who sees the case differently than you do. This is important because the goal of this activity is to help the lawyers understand how jurors will perceive the case and come to a decision.

Additionally, participants in the ‘justify’ condition were reminded on two more occasions that they would have to justify their eventual decision, once prior to reading the case evidence and once prior to rendering their interim verdict leaning. After providing their final verdict and final evidence ratings, every participant, regardless of condition, was actually asked to ‘justify’ and explain their verdict. However, only participants in the ‘justify’ and ‘justify + devil’s advocate’ conditions were told from the beginning that they should anticipate having to justify their views.

**‘Devil’s Advocate’ Manipulation.** Participants read the following extra paragraph at the end of instructions section:

As you read about the various pieces of evidence presented throughout the trial, you will be asked to take some time to consider any issues with those pieces of evidence. Meaning, even if evidence is presented by one side of the case, either the prosecution or the defense, consider how that piece of evidence would be perceived from the other side of the case. If you were the opposing attorney, what problems, flaws, or issues would YOU find with each piece of evidence?

Additionally, after participants in the ‘Devil’s Advocate’ and ‘Justify + Devil’s Advocate’ condition had read a piece of the case evidence, they were asked:

For the piece of evidence you just read, imagine that you’re the opposing attorney: Are there any flaws, issues, or problems with this evidence? Are there any alternative explanations to the evidence? Are there any aspects of the evidence that might actually support your side of the case, as the opposing lawyer, instead of supporting their side of the case? Please provide your responses below:

**The trial: The people v. David Mitchell.** After reading the instructions, all participants read through an identical synopsis of the case that both sides have agreed upon (Appendix E). Following that, participants read about both the inculpatory and exculpatory evidence in this case (Appendix F). The evidence order was randomized and the valence of the evidence was counterbalanced as well. For example, one participant

may have read about a polygraph test that points towards the defendant being guilty, whereas another participant may have read about near-identical polygraph evidence, only it points towards the defendant being innocent. One additional note is that the 3 pieces of evidence that are inculpatory were presented in a row, under the heading ‘evidence from the prosecution’s case, and the 3 pieces of evidence that were exculpatory were presented in a row as well under the heading ‘evidence from the defense’s case.’ The order of which section of evidence came first was also randomized. For example, one participant’s evidence order may be: inculpatory eyewitness, polygraph, and motive, followed by exculpatory shoe print, fingerprint and traffic camera. Whereas another participant’s evidence order may be: exculpatory shoe print, traffic camera, polygraph, followed by inculpatory motive, eyewitness, and fingerprint.

Once participants read through the evidence, they were asked to provide their initial verdict leanings as well to rate the degree to which each piece of evidence implied the guilt or innocence of the suspect (see Appendix G). The rating scale chosen for this study ranged from -7 (strongly implies not guilty) to 0, neutral, to 7 (strongly implies guilty). This scale was chosen because it has a wide range of responses which should reduce floor and ceiling effects and it allows for opinions to be mobile and shift dramatically over time. If participants were in the ‘justify’ or ‘Justify + Devil’s Advocate’ conditions, they were also reminded at this point about how they will have to eventually discuss their decisions. However, participants were asked to withhold arriving at a final decision, as more evidence would be presented shortly.

After providing their initial leanings and opinions about the evidence, participants moved on to the next phase of the trial (see Appendix H). During this phase, two new

ambiguous (neutral) pieces of evidence were presented (phone records and alibi). The new evidence did not come from the prosecution's side or the defense's side, rather, the lawyers stipulated that the jury should have all of the facts. The order of the two pieces of ambiguous evidence was randomized. Participants will be asked to rate the new evidence as well.

**Post-Trial Questionnaires.** Once the trial portion of the study had been completed, participants proceeded to fill out post-trial questionnaires (see Appendix I). To begin, participants provided a final verdict and indicated their confidence level associated with their verdict. Participants were also asked to rate the *extent* to which the defendant is guilty or not guilty. Additionally, they rated all eight pieces of evidence again (the 6 valenced pieces and the 2 ambiguous pieces), at post-trial, with regard to how much each piece implies the guilt or innocence of the suspect. Following those ratings, all participants were in fact asked to justify and defend why they selected the verdict they rendered, regardless of condition. Participants were then asked to recall their original ratings of the 8 pieces of evidence:

Remember how at the beginning of this activity, in the first packet, you rated the various pieces of evidence and matters of legal policy? Try to remember how you rated each piece of evidence and indicate those ratings here. If you cannot remember the exact number, try to guess the number as closely as possible. Do not look back at your first packet to see your original ratings.

Once participants completed the 'original evidence ratings,' they also filled out the Juror Bias Scale (JBS; Kassin & Wrightsman, 1983). The JBS measures pro-prosecution and pro-defense legal attitudes and serves as a helpful tool for establishing a

baseline of beliefs which may be able to predict the direction of coherence shifts that a participant experiences, towards either guilt or innocence.

A few other questions were asked during this time as well, including: a believability check (“Do you think your response will help the lawyers?”), a manipulation check (“Which of the following instructions did you receive?”), how much pressure they felt to ‘justify’ their final verdict, and to what extent their opinions of the evidence shifted over the course of the trial, and finally demographic information including native language, age, whether or not they have served as a real juror, and so forth.

## **RESULTS**

### **General descriptives**

Overall, of the 306 participants, 111 provided guilty verdicts and 189 provided not guilty verdicts (six participants did not respond to verdict, but their responses were not deleted from the data). Of the 306 participants, there was 46 jurors who were ‘switchers’, meaning their interim verdict leaning and their final post-trial verdict were different. As Simon (2004) points out, jurors who switch their verdict may be undergoing a different cognitive process than jurors who experienced a clear shift in their decision one way or another. Thus, for all of the following analyses, these ‘switchers’ were removed and the coherence effects they experience will be analyzed separately, at the end of the results section.

### **Coherence shifts in evidence evaluation**

**Valenced evidence shifts.** As Hypothesis 1a predicted, a 2 (verdict: guilty or not guilty) X 3 (evidence ratings at: abstract phase, interim, and post trial) repeated measures

ANOVA revealed that participants experienced a significant and dramatic shift in their perceptions of evidence over the course of the trial as a function of their verdict,  $F(2, 472) = 48.03, p < .001, \eta_p^2 = .91$ . Specifically, follow-up tests revealed that among convicts, their average ratings of the abstract evidence ( $M = .47, SD = 1.42$ ) were already significantly greater than 0 at the beginning of the trial, which predicts their final verdict direction,  $t(86) = 3.11, p = .003, d = .32$ . Additionally, there was a significant increase in the perceived inculpatory value of evidence ratings from the abstract evidence ratings, to the interim evidence ratings ( $M = 1.50, SD = 1.25$ ),  $t(86) = 6.07, p < .001, d = .30$ . There was then another significant increase in the inculpatory value of the evidence from interim evidence ratings to the post-trial evidence ratings ( $M = 1.85, SD = 1.46$ ),  $t(84) = 4.55, p < .001, d = .44$ . The latter finding is different from previous studies' coherence results, in that significant shifts are not typically observed from interim to post-trial (Simon, 2004).

The same pattern was found among acquitters, such that their average ratings of the abstract evidence ( $M = -.76, SD = 1.80$ ) were significantly less than 0 at the beginning of the trial, which predicts their final verdict direction,  $t(160) = 5.38, p < .001, d = .39$ . Additionally, there was an increase in the perceived exculpatory value of the evidence from abstract to the interim phase ( $M = -1.51, SD = 1.98$ ),  $t(157) = 4.83, p < .001, d = .36$ . However, among acquitters, there was not a significant increase in the perceived exculpatory value of the evidence from interim ( $M = -1.51, SD = 1.99$ ) to post-trial ( $M = -1.46, SD = 2.11$ ),  $t(156) = -.66, p = .51$ .

The observed coherence effects demonstrate that as jurors gravitate towards either a verdict of guilty or not guilty, their perceptions of the evidence polarize in accordance

with that over the course of coming to a decision. These results are shown graphically in Figure 3.

**Ambiguous evidence shifts and biased perception.** All participants evaluated the same two pieces of ambiguous evidence – phone records and the suspect’s alibi.

Hypothesis 1b predicted that a juror’s interim verdict would ‘taint’ the way he/she would evaluate the ambiguous pieces of evidence. Consistent with this prediction, participants’ interim leanings of guilty/not guilty were significantly related to how jurors evaluated the ambiguous evidence: One-sample t-tests indicated that jurors leaning towards guilt perceived the interim ambiguous evidence to be significantly inculpatory ( $M = 2.20, SD = 2.08, t(87) = 9.96, p < .001, d = .73$  whereas jurors leaning towards innocence perceived the exact same interim ambiguous evidence to be significantly exculpatory ( $M = -.84, SD = 2.64, t(165) = 4.12, p < .001, d = .31$ ).

The degree to which jurors’ ratings of the six pieces of valenced evidence at the interim ‘tainted’ their perception of the interim ambiguous evidence was evaluated. The more that convicts perceived the valenced evidence to imply guilt ( $M = 1.50, SD = 1.24$ ), the more they perceived the ambiguous evidence to imply guilt ( $M = 2.20, SD = 2.08$ ),  $r = .25, p = .02$ . Similarly, the more acquitters perceived the valenced evidence to imply innocence ( $M = -1.49, SD = 1.97$ ), the more they perceived the ambiguous evidence to imply innocence ( $M = -.84, SD = 2.64$ ),  $r = .45, p < .001$ .

A 3 (ambiguous evidence ratings: abstract, ratings, and post-trial) X 2 (verdict: guilty or not guilty) repeated measures ANOVA revealed a significant ambiguous evidence ratings by verdict interaction,  $F(1, 250) = 27.02, p < .001, \eta_p^2 = .86$ . Follow-up analyses revealed that, initially, convicts perceived the abstract ambiguous evidence as

slightly inculpatory ( $M = 1.00$ ,  $SD = 2.33$ ), but they came to perceive significantly more guilt in the same evidence after the interim of the trial ( $M = 2.20$ ,  $SD = 2.09$ ),  $t(86) = -4.71$ ,  $p < .001$ ,  $d = .45$ . Furthermore, there was a significant increase in the inculpatory value of the ambiguous evidence between the interim ( $M = 2.20$ ,  $SD = 2.09$ ) and post-trial ( $M = 2.51$ ,  $SD = 2.03$ ) ratings,  $t(86) = -2.18$ ,  $p = .03$ ,  $d = .23$ .

Although acquitters trended in the same direction, they did not experience such strong shifts with respect to the ambiguous evidence. There was not a significant difference between the abstract ratings of the ambiguous evidence ( $M = -.55$ ,  $SD = 2.55$ ) and the interim ratings of the ambiguous evidence ( $M = -.84$ ,  $SD = 2.64$ ),  $t(165) = 1.39$ ,  $p = .18$ . Nor was there a significant shift between the interim ambiguous evidence ratings and the post-trial ambiguous evidence ratings ( $M = -1.00$ ,  $SD = 2.62$ ),  $t(165) = 1.20$ ,  $p = .23$ . Acquitters did, however, evaluate the ambiguous evidence as significantly more exonerating at post-trial compared to in the abstract,  $t(165) = 1.99$ ,  $p = .05$ ,  $d = .15$ .

**Lack of insight into coherence shifts.** Hypothesis 1c predicted that jurors would lack awareness into the extent that their perceptions of the evidence shifted from the beginning of the trial to the end of the trial. Consistent with this prediction, convicts' *recollections* of the abstract evidence ( $M = 1.66$ ,  $SD = 1.71$ ) turned out to be significantly more inculpatory than their original perceptions of the evidence really were ( $M = .47$ ,  $SD = 1.71$ ),  $t(81) = -6.00$ ,  $p < .001$ ,  $d = .55$ . In other words, convicts believed they had felt the evidence was strongly inculpatory all along. In actuality, their recollections of their abstract ratings turned out to be much more similar to their post-trial ratings ( $M = 1.85$ ,  $SD = 1.49$ ), as there was no significant difference between the two,  $t(80) = -1.56$ ,  $p = .12$ .



Convicters lacked insight into the extent to which their opinions about evidence had shifted.

The same finding was observed among acquitters: Their recollections of the abstract evidence ( $M = -1.31$ ,  $SD = 2.21$ ) were more exculpatory than their original perceptions of the evidence were ( $M = -.73$ ,  $SD = 1.78$ ),  $t(155) = -3.80$ ,  $p < .001$ ,  $d = .29$ . Again, in actuality, their recollections of their abstract ratings turned out to be much more similar to their post-trial ratings ( $M = -1.50$ ,  $SD = 2.06$ ), as there was only a marginally significant difference between the two,  $t(155) = 1.76$ ,  $p = .08$ .

### **Coherence shifts in evidence integration**

**Inter-item Correlations of evidence.** One way to conceptualize evidence integration is to evaluate the extent that the various pieces of evidence are correlated or hang together. Hypothesis 2a predicted that there should be an increase in the extent to which evidence is correlated as a juror progresses through the decision-making process.

This prediction was supported by the data. At the beginning of the trial, there was a modest correlation between jurors' ratings of all of the evidence (which includes the six pieces of valenced evidence and the two ambiguous pieces of evidence), Cronbach's  $\alpha = .47$ . As the trial progressed to the interim stage, the correlation between evidence increased significantly to Cronbach's  $\alpha = .75$ ,  $z = -5.19$ ,  $p < .001$ . Finally, at post-trial, the correlation between the various pieces of evidence was marginally higher than at the interim, Cronbach's  $\alpha = .81$ ,  $z = 1.78$ ,  $p = .08$ . The increase in inter-item correlations across the various points of the decision indicates that evidence ratings are *not* merely independent evaluations, rather, the various pieces of evidence cling together more and more and come close to representing a single factor, of either guilt or innocence.

**Correlation between overall ‘likelihood of guilt’ and evidence ratings.** As one would expect, convickers perceived a greater likelihood of guilt ( $M = 4.34$ ,  $SD = .62$ ) than acquitters ( $M = 2.39$ ,  $SD = 1.02$ ),  $t(252) = 16.44$ ,  $p < .001$ ,  $d = .72$ .

There was partial support for Hypothesis 2b, as convickers perceptions of the defendant’s likelihood of guilt increase/shifted from interim ( $M = 4.07$ ,  $SD = .50$ ) to post-trial ( $M = 4.34$ ,  $SD = .62$ ),  $t(87) = 3.64$ ,  $p = .001$ ,  $d = .36$ . Unlike convickers, acquitters perceptions of the defendant’s likelihood of guilt did not significantly shift from interim ( $M = 2.49$ ,  $SD = .91$ ) to post-trial ( $M = 2.39$ ,  $SD = 1.02$ ),  $t(165) = 1.37$ ,  $p = .17$ .

As Hypothesis 2b predicted, there was a significant correlation between jurors’ average interim evidence ratings ( $M = -4.30$ ,  $SD = 2.26$ ) and their interim judgments of the defendant’s likelihood of guilt ( $M = 3.04$ ,  $SD = 1.10$ ),  $r = .58$ ,  $p < .001$ . There was also significant correlation between jurors’ average post-trial evidence ratings ( $M = -.30$ ,  $SD = 2.47$ ) and their post-trial judgments of the defendant’s likelihood of guilt ( $M = 3.06$ ,  $SD = 1.30$ ),  $r = .65$ ,  $p < .001$ . There was not a significant difference in the correlations between average evidence ratings and likelihood of guilt from interim ( $r = .58$ ) to post-trial ( $r = .65$ ),  $z = 1.24$ ,  $p = .21$ . The relation between evidence ratings and likelihood of guilt speaks to the fact that evidence becomes integrated to form an overall conclusion about the defendant’s guilt or lack thereof.

**Relationship between perceptions of evidence and verdict.** To expand on the relationship between evidence evaluation and evidence integration, one can compare the average post-trial ratings of evidence to the final verdict of the jurors. As Hypothesis 2c predicted, there was indeed a significant relationship; jurors who voted guilty perceived much more inculcating value in the evidence ( $M = 1.85$ ,  $SD = 1.46$ ) than jurors who

voted not guilty ( $M = -1.45$ ,  $SD = 2.10$ ),  $t(242) = 12.90$ ,  $p < .001$ ,  $d = .64$ . Note that the directionality of the relationship between decision outcome (verdict) and evidence evaluation is difficult to discern. In other words, it is not clear whether evidence evaluations lead to judgments of guilt or whether perceived judgments of guilt are biasing evidence evaluations.

**Coherence and confidence.** In general, most participants indicated moderate confidence in their decisions ( $M = 3.13$ ) on a scale that ranged from -7 (not very confident) to 7 (very confident). Among jurors who voted guilty, post-trial confidence ( $M = 3.63$ ,  $SD = 2.61$ ) was significantly correlated with the overall perceived likelihood of the defendant being guilty ( $M = 4.24$ ,  $SD = .72$ ),  $r = .44$ ,  $p = .04$ . Similarly, the less acquitters perceived the defendant to be guilty ( $M = 2.51$ ,  $SD = 1.06$ ), the more confident they became in their verdicts ( $M = 2.85$ ,  $SD = 2.97$ ,  $r = -.39$ ,  $p < .001$ ).

A 2 (confidence interim and post-trial) X 2 (verdict) repeated measure ANOVA revealed that convicters and acquitters were not significantly different in terms of their shifts in confidence,  $F(1, 251) = .77$ ,  $p = .38$ . A test for main effects, however, demonstrated that jurors experienced a significant shift in confidence from the interim confidence ratings ( $M = 2.49$ ,  $SD = 2.76$ ) to their post-trial confidence ratings ( $M = 3.32$ ,  $SD = 2.78$ ),  $F(1, 251) = 24.13$ ,  $p < .001$ ,  $\eta_p^2 = .84$ .

To further analyze confidence and coherence, one can examine how the magnitude of jurors' coherence shifts over the course of the trial correlate with their confidence. For each participant, a 'change in evidence score' was calculated by taking the average rating of the 6 pieces of valenced evidence, as taken at post-trial, and subtracting the average rating of the 6 pieces of valenced evidence during the initial

abstract rating of evidence scenarios. The more convicts shifted in their ratings of the evidence from abstract to post-trial toward guilt ( $M = 1.34$ ,  $SD = 1.80$ ), the more confident they became ( $M = 3.89$ ,  $SD = 2.33$ ),  $r = .24$ ,  $p = .03$ . Similarly, more acquitters shifted in their ratings of the evidence from abstract to post-trial toward innocence ( $M = -.74$ ,  $SD = 2.25$ ), the more confident they became ( $M = 3.05$ ,  $SD = 2.96$ ),  $r = -.19$ ,  $p = .02$ .

### **Moderators of Coherence**

**Pro-prosecution and pro-defense attitudes.** The Juror Bias Scale (JBS; Kassin & Wrightsman, 1983) is designed to measure a juror's proclivity towards favoring the prosecution or favoring the defense, and thus it serves as a measure of predicting verdict. For the purposes of the present study, higher scores on the JBS indicate greater pro-prosecution attitudes (the present study reverse coded the JBS, which originally coded higher JBS scores as indicating pro-defense attitudes). As Hypothesis 3a predicted, JBS scores were correlated with ratings of evidence and overall conclusions about the trial.

There was a marginally significant correlation between jurors' JBS score and their average ratings of the valenced abstract evidence, such that the more pro-prosecution jurors were, the more the more inculpatory they perceived the valenced abstract evidence to be,  $r = .12$ ,  $p = .06$ . For the ambiguous abstract evidence ratings, there was a significant effect such that, the more pro-prosecution jurors were, the more the more inculpatory they perceived the ambiguous abstract evidence to be,  $r = .15$ ,  $p = .02$ .

Among participants who underwent a shift in evidence evaluation from abstract to post-trial, their JBS score significantly predicted the direction in which they shifted; the more pro-prosecution jurors were, the more likely they were to shift their evidence evaluations towards guilt,  $r = .18$ ,  $p = .007$ . Additionally, the more pro-prosecution

jurors were, the greater the magnitude of the shift in their ratings from abstract to post-trial,  $r = .15, p = .02$ .

Furthermore, as they progressed throughout the trial, the more pro-prosecution jurors were, the more inculpatory they perceived the valenced interim evidence to be,  $r = .25, p < .001$  (and vice versa, the more pro-defense jurors were, the more exculpatory they perceived the interim evidence to be). Similarly, the more pro-prosecution jurors were, the more inculpatory they perceived the valenced post-trial evidence to be,  $r = .24, p < .001$ . In terms of the ambiguous evidence, the more pro-prosecution jurors were, the more inculpatory they perceived the ambiguous interim evidence to be,  $r = .27, p < .001$ . Also, the more pro-prosecution jurors were, the more inculpatory they perceived the ambiguous post-trial evidence to be,  $r = .94, p < .001$ .

In terms of jurors' overall conclusions about the trial, jurors who voted guilty had significantly higher pro-prosecution attitudes ( $M = 36.31, SD = 7.07$ ) than jurors who voted not guilty ( $M = 32.80, SD = 6.22$ ),  $t(245) = 4.01, p < .001, d = .25$ . In keeping with that finding, the more pro-prosecution jurors were, the more they perceived a greater the likelihood of the defendant's guilt,  $r = .23, p < .001$ .

**Coherence and instruction type.** Hypotheses 3b and 3c predicted that compared to jurors in the control condition, jurors in the justify condition would show larger coherence effects whereas those in the Devil's Advocate condition would show weaker coherence effects on a variety of measures, including valenced evidence ratings, ambiguous evidence ratings, inter-item correlations, likelihood of guilt, and confidence.

**Valenced evidence ratings by condition.** A 2 (justify instructions: present or absent) X 2 (Devil's Advocate instructions: present or absent) X 2 (verdict: guilty or not

guilty) ANOVA revealed no significant three way interaction between justify instructions, Devil's Advocate instructions, and verdict on the evaluations of the post-trial valenced evidence,  $F(1, 236) = .00, p = .98$ . There was also no interaction significant interactions between justify instructions and verdict,  $F(1, 236) = .00, p = .95$  or justify instructions and Devil's Advocate Instructions,  $F(1, 236) = .61, p = .44$ . There was a marginally significant interaction between Devil's Advocate Instructions and verdict,  $F(1, 236) = 2.83, p = .09$ . A test of simple main effects revealed that among convicters, there was no difference in ratings of post-trial evidence based on whether jurors received Devil's Advocate instructions ( $M = 2.05, SD = 1.46$ ) or not ( $M = 1.72, SD = 1.45$ ),  $t(83) = 1.02, p = .31$ . Similarly, among acquitters, there was no difference in ratings of post-trial evidence based on whether jurors received Devil's Advocate instructions ( $M = -1.71, SD = 2.11$ ) or not ( $M = -1.17, SD = 2.07$ ),  $t(157) = 1.60, p = .11$ .

Because of the use of legal language and terminology, a secondary analysis was conducted to see if there was a difference in the average post-trial ratings of evidence for people whose native language was English versus not English. However, there was no difference in post-trial evidence ratings based on condition and primary language,  $F(1, 243) = 1.15, p = .33$ .

Participants were then split by whether they passed the manipulation check or not (i.e., whether they correctly answered a multiple-choice question about which instructions they read and followed throughout the trial). However, no significant differences were found,  $F(1, 225) = .65, p = .71$ . In other words, whether participants passed the manipulation check and accurately remembered which instructions they received had no effect on the relationship between condition and evidence ratings.

Beyond their memory for which instructions they received, there are other ways to examine the extent to which participants engaged with the manipulations. Participants' responses to the Devil's Advocate manipulation (where they had to list the flaws inherent in each piece of evidence) were coded by research assistants in terms of the number of flaws, issues, or problems a participant raised with a given piece of evidence. The total number of such issues raised was that participant's *Devil's Advocate Effort* score. For example, if in response to the eyewitness evidence a participant stated "the witness saw him at dusk, when it is getting dark and could be hard to see," that participant would receive a score of '1' for his/her *Devil's Advocate Effort* score. The average number of points raised by participants in the Devil's Advocate conditions was 7.44 ( $SD = 2.83$ ). A new dichotomous variable was created that represented the amount of effort participants put in, with low effort being less than 7 points raised and high effort being 7 or more points raised. However, among convicts, there was no difference in average post-trial evidence ratings between high effort participants ( $M = 2.06$ ,  $SD = 1.50$ ) and low effort participants ( $M = 2.08$ ,  $SD = .37$ ),  $F(1, 30) = .002$ ,  $p = .97$ . Among acquitters, there was also no significant difference in average post-trial evidence ratings between high effort ( $M = -1.72$ ,  $SD = 2.33$ ) compared to low effort participants ( $M = -1.56$ ,  $SD = 1.88$ ),  $F(1, 73) = .09$ ,  $p = .76$ .

One final check to make sure there was no difference in post-trial evidence ratings based on condition was performed. All participants provided a rating of how much pressure they felt to justify their decision throughout the trial ( $M = 1.05$ ;  $SD = 3.76$ ), measured on a scale ranging from -7 (lack of pressure) to 7 (extreme pressure). Responses were dichotomized such that participants with a positive score were

considered ‘high pressure’ and participants with a negative (or zero) score were considered ‘low pressure.’ However, there was no significant difference in post-trial evidence ratings between participants who felt pressure to justify ( $M = -.14$ ,  $SD = .21$ ) and those who felt no pressure ( $M = -.46$ ,  $SD = 2.48$ ),  $F(1, 190) = .81$ ,  $p = .37$ .

***Ambiguous evidence rating by condition.*** A 2 (justify instructions: present or absent) X 2 (Devil’s Advocate instructions: present or absent) X 2 (verdict: guilty or not guilty) ANOVA once again revealed no significant three way interaction between justify instructions, Devil’s Advocate instructions, and verdict on the evaluations of the ambiguous post-trial evidence,  $F(1, 245) = .35$ ,  $p = .55$ . Similarly, there was no interaction between: justify instructions and verdict,  $F(1, 245) = .98$ ,  $p = .32$ , nor Devil’s Advocate and verdict  $F(1, 245) = .65$ ,  $p = .42$ , nor justify and Devil’s Advocate instructions,  $F(1, 245) = .66$ ,  $p = .42$ .

***Inter-item correlations by condition.*** Additionally, the extent to which the post-trial evidence ratings were correlated was similar across conditions: control (Cronbach’s  $\alpha = .76$ ), justify (Cronbach’s  $\alpha = .78$ ), Devil’s Advocate (Cronbach’s  $\alpha = .84$ ), and justify and Devil’s Advocate (Cronbach’s  $\alpha = .85$ ). Even the most dramatic difference, between control (Cronbach’s  $\alpha = .76$ ) and justify and Devil’s Advocate (Cronbach’s  $\alpha = .85$ ) was non-significant,  $z = 1.40$ ,  $p = .32$ .

***Likelihood of guilt by condition.*** A 2 (justify instructions: present or absent) X 2 (Devil’s Advocate instructions: present or absent) X 2 (verdict: guilty or not guilty) ANOVA revealed no significant interactions between instruction type and verdict on the perceptions of the defendant’s likelihood of guilt,  $F(1, 246) = .11$ ,  $p = .74$ . There was no difference in likelihood of guilt based on condition and primary language,  $F(1, 245) =$



1.36,  $p = .23$ . Similarly, participants' average ratings of evidence was compared across condition and whether or not they passed the manipulation check, however, no significant differences were found,  $F(1, 235) = .51, p = .82$ . Lastly, as with the earlier analyses, there was no difference in likelihood of guilt based on the effort one put into the Devil's Advocate manipulation,  $F(1, 110) = 1.71, p = .19$ .

**Verdict by condition.** There was no significant difference in the proportion of guilty verdicts based on whether participants were in the control condition (40% guilty, 60% not guilty), the 'justify' condition (39% guilty, 61% not guilty), the Devil's Advocate condition (38% guilty, 62% not guilty), or the 'justify' and Devil's Advocate condition (30% guilty, 70% not guilty),  $\chi^2(3) = 2.00, p = .57$ .

**Confidence by condition.** A 2 (justify instructions: present or absent) X 2 (Devil's Advocate instructions: present or absent) X 2 (verdict: guilty or not guilty) ANOVA revealed no significant three way interaction between justify instructions, Devil's Advocate instructions, and verdict on confidence,  $F(1, 246) = 1.54, p = .22$ . Similarly, there was no interaction between: justify instructions and verdict,  $F(1, 245) = .07, p = .79$ , nor Devil's Advocate and verdict  $F(1, 245) = .17, p = .68$ , nor justify and Devil's Advocate instructions,  $F(1, 245) = .43, p = .51$ .

**Coherence and order effects.** There was a marginally-significant effect of order (i.e., whether participants read about the prosecution's evidence most recently or the defense's evidence most recently) on verdict. When the defense's evidence was presented *second* there were more not guilty verdicts (67% not guilty, 33% guilty) than when the defense's evidence was presented *first* (59% not guilty, 41% guilty),  $\chi^2(1) = 2.33, p$  (one-sided) = .08. Upon further analysis, the 'justify,' Devil's Advocate, and

‘justify’ and Devil’s Advocate conditions showed no difference based on trial order (all  $\chi^2$ s < .80, all  $p$  values > .40). However, the control condition demonstrated a significant effect for trial order such that when the defense’s evidence was presented *second* there were significantly more not guilty verdicts (73% not guilty, 27% guilty) than when the defense’s evidence was presented *first* (48% not guilty, 52% guilty),  $\chi^2(1) = 5.19$ ,  $p$  (two-tailed) = .02. In other words, participants’ final verdicts tended to align with the *second* set of evidence that was presented. This result is consistent with the recency effect found in Charman et al.’s (under review) study, whereby DNA evidence presented *second* had a stronger impact on perceptions of the defendant’s guilt than DNA evidence presented *first*.

### **Coherence effects among Switchers**

**Lack of coherence shifts among switchers.** There was an approximately equal proportion of jurors who switched from an initial leaning of guilty (40%) to a final verdict of not guilty (44%), as there were jurors who switched from a verdict of not guilty (46%) to a verdict of guilty (44%).

As Hypothesis 4 predicted, switchers (i.e., those participants whose final verdicts did not match their interim leanings) were less likely to experience coherence shifts. Convicters did not experience a significant shift in their ratings of evidence from abstract ( $M = -.01$ ,  $SD = 1.90$ ) to interim ( $M = -.45$ ,  $SD = 1.57$ ),  $t(22) = 1.49$ ,  $p = .15$ . Acquitters also did not experience a significant shift in their ratings of evidence from abstract ( $M = -.10$ ,  $SD = 1.64$ ) to interim ( $M = -.04$ ,  $SD = 1.29$ ),  $t(22) = -.18$ ,  $p = .86$ . However, convicters did experience a significant shift from interim ( $M = -.45$ ,  $SD = 1.57$ ) to post-trial ( $M = .33$ ,  $SD = 1.57$ ),  $t(21) = -4.07$ ,  $p = .001$ ,  $d = .66$ . Acquitters, however, did not

undergo a significant shift from interim ( $M = -.04$ ,  $SD = 1.29$ ) to post-trial ( $M = .08$ ,  $SD = 1.34$ ),  $t(22) = -1.05$ ,  $p = .31$ .

One explanation for why the switchers changed verdicts from interim to post-trial is that those jurors who changed their final verdict to guilty perceived more inculpatory value in the interim ambiguous evidence ( $M = 1.93$ ,  $SD = 2.55$ ) than those jurors who changed their verdict to not guilty ( $M = .33$ ,  $SD = 2.21$ ),  $t(44) = 2.29$ ,  $p = .03$ ,  $d = .33$ . It bears repeating that the presentation and rating of the interim ambiguous evidence occurred in between the interim verdict selection and the final verdict selection and the ambiguous evidence is the only new information between interim and post-trial. Thus, jurors' switch in verdict choice is cohering with their perceptions of the ambiguous evidence.

**Comparing switchers to non-switchers.** There was a significantly higher proportion of convicts among switchers ( $n = 46$ , proportion guilty = 50%) than among non-switchers ( $n = 254$ , proportion guilty = 34.65%),  $\chi^2(1) = 3.94$ ,  $p = .04$ . In keeping with that finding, switchers also thought that Mitchell was more likely to be guilty ( $M = 3.63$ ,  $SD = .96$ ) than did non-switchers ( $M = 3.06$ ,  $SD = 1.30$ ),  $F(1, 300) = 8.17$ ,  $p = .005$ ,  $d = .43$ . Examining the difference between switchers and non-switchers is necessary if coherence effects are to be fully understood, as one cannot ignore a subset of decision-makers who experience coherence effects differently. The finding that switchers have more inculpatory views of the defendant than non-switchers is one way in which the two populations seem to differ.

Finally, switchers were significantly less confident ( $M = 2.02$ ,  $SD = 3.04$ ) in their eventual verdict than non-switchers ( $M = 3.34$ ,  $SD = 2.78$ ),  $t(300) = 2.96$ ,  $p = .003$ . This,

of course, makes intuitive sense, as someone who is feeling less confident is more likely to be swayed by the subsequent ambiguous evidence or simply by changing his/her mind.

## **DISCUSSION**

The present study's results demonstrate that coherence shifts have a potent impact on juror decision-making; mock jurors tended to arrive at polarized and confident verdicts despite the fact that the evidence presented in the case was ambiguous. Additionally, the coherence shifts observed in this study remained significant even though the evidence order and evidence valence were randomized.

The coherence shifts begin with the evaluation of abstract pieces of evidence. Using the Juror Bias Scale (Kassin & Wrightsman, 1983), it was shown that the way jurors perceived the abstract evidence was predicted by their general predisposition towards favoring the prosecution or favoring the defense. For example, jurors who had general pro-prosecution attitudes judged abstract vignettes about various pieces of evidence to be more inculpatory than jurors with general pro-defense attitudes.

As the trial proceeded, jurors began to develop initial leanings towards either guilt or innocence. Again, the JBS revealed that the direction jurors shifted – towards either prosecution or defense – was related to their pro-prosecution or pro-defense attitudes. As jurors developed leanings, their opinions of the evidence (which were once relatively unpolarized in the abstract phase of the study) began to cohere in a manner consistent with those leanings, and consequently further polarized. For example, a pro-prosecution juror who initially rated fingerprint evidence as having an inculpatory value of 3 (out of 7) during the abstract phase, may come to re-rate the evidence as a 4 during the interim of the trial, and then re-rate that same evidence as 5 after he has rendered a guilty verdict.

The shift in perceptions of the evidence is concurrent with shifts in overall case judgments, such as likelihood of guilt. According to Simon (2004), these coherence shifts are a result of a bi-directional process, in which perceptions of the evidence affect outcome leanings, which then further influence the perceptions of evidence.

Importantly, the present study found that the direction jurors lean towards also caused them to have a biased interpretation of ambiguous evidence that was presented during a later phase of the trial. In this way, jurors experienced a self-generated bias: their own inclinations towards prosecution or defense resulted in a subsequent biased interpretation of ambiguous evidence. For example, when pro-defense jurors who began to lean towards innocence reached the phase of the trial where new ambiguous evidence was presented, such as phone records, they tended to perceive the phone records as exculpatory (compared to pro-prosecution jurors, who tended to perceive the exact same evidence as inculpatory).

On account of these coherence effects, when jurors reached a final verdict, they tended to do so confidently. Importantly, jurors' confidence in their verdict also shifted throughout the course of the trial along with the evaluations of evidence and along with the overall the likelihood of the defendant's guilt. Jurors' high confidence in their verdicts is somewhat troubling in light of the fact that their purest ratings of the probative value of the evidence (the abstract ratings) are markedly different from how they perceive the evidence at the end of the trial. Moreover, jurors lack awareness of the extent and the direction that their evaluations have shifted; in fact most jurors believed their evaluation did not shift much at all. By comparing jurors' post-trial evidence ratings to their memory of their abstract evidence ratings, the present study showed that jurors come to

believe how they feel about the evidence at the end of the trial was how they felt all along.

Additionally, the current study found that these robust coherence shifts were not easily mitigated or reduced. For example, participants who were in the Devil's Advocate conditions could have reasonably been expected to have more neutral ratings of the evidence, as they had to list the problems inherent in each piece of evidence (which were numerous because the evidence was designed to be flawed). Nonetheless, neither the Devil's Advocate nor the justify instructions had any effect on coherence effects, which speaks to how ingrained these effects are in the cognitive processes underlying decision-making. In fact, other researchers have also noted the difficulty with reducing coherence effects using a variety of manipulations (Simon, 2012).

Perhaps one reason coherence shifts are so hard to reduce is that all of the perceptions of the evidence become highly integrated automatically, without conscious control. The present study demonstrated that the inter-item correlations between the pieces of evidence grew stronger as the trial progressed, and that jurors were unaware of this change, suggesting that this coherence process occurs at a non-conscious level. At the conclusion of the trial, jurors perceived the various pieces of evidence more like a single construct representing either guilt or innocence.

Study 1 showed the existence of coherence shifts using mock jurors; however, it remains to be seen whether coherence shifts affect other members of the legal system, such as public defenders, in the same way that they affect jurors. This is the purpose of Study 2.

## **STUDY 2**

### **Goals**

The primary goal of Study 2 is to examine the decision-making of public defenders. There have been almost no studies on the decision-making processes of public defenders. This is unfortunate for several reasons. First, the vast majority of criminal cases are resolved via plea bargains and never make it to trial (Simon, 2012). As such, it is important to understand the decision-making of lawyers, who negotiate those plea bargains. Public defenders, specifically, may have unique biases that affect their decisions to take or reject plea offers; unfortunately, little is known about the interplay between a public defender's beliefs about the suspect's actual guilt and his or her negotiations of the plea bargain. Second, the majority of legal decision-making research has focused on outcome variables such as verdict and evidence evaluation; no legal psychology research has looked at plea bargaining as an outcome variable. The translation of the perceived evidence strength against the defendant into a decision regarding an appropriate plea bargain to accept (i.e., concerning the number of potential years in prison) would certainly constitute a complex and multifaceted decision. As such, the potential to find coherence effects is quite strong. The present study strives to explore these new areas and, in general, to have a degree of external validity that is not often found in decision-making research.

### **Hypotheses**

#### **Hypothesis 1 – Coherence Effects and evidence evaluation**

**Hypothesis 1a. – coherence shifts in valenced evidence.** Similar to Hypotheses 1a from Study 1, it is predicted that public defenders' average ratings of the valenced

evidence will shift over the course of the study and will shift in the direction of their perceptions of the defendant's guilt.

**Hypothesis 1b. – coherence shifts in ambiguous evidence.** It also predicted that the ratings the ambiguous evidence will shift over the course of the study in accordance with their perceptions of the defendant's guilt.

### **Hypothesis 2 – Coherence shifts in inter-item correlations among evidence**

Similar to Hypothesis 2a. from Study 1, it is predicted that the inter-item correlation among public defenders' ratings of evidence will become stronger over the course of the trial.

### **Hypothesis 3 – Plea bargaining and coherence effects**

**Hypothesis 3a. – plea bargaining and ratings of evidence.** It is predicted that the maximum number of years a public defender is willing to recommend that their client accept will be correlated with their average ratings of the evidence.

**Hypothesis 3b. – plea bargaining and likelihood of guilt.** It is predicted that the maximum number of years a public defender is willing to recommend that their client accept will be correlated with their overall beliefs about the defendant's likelihood of guilt.

### **Hypothesis 4 – Pro-prosecution versus pro-defense attitudes in public defenders**

In keeping their role as public defenders, it is predicted that they will have strongly pro-defense attitudes. Furthermore, it is also predicted that their pro-defense or pro-prosecution attitudes will be related to the direction of the coherence shifts they experience, toward perceiving the defendant as guilty or innocent.



## **METHOD**

### **Participants**

The population of the second study consisted of actual public defenders. For logistical reasons, only 9 participants were able to be recruited (4 males, 5 females; ages 28 – 55, mean age = 34 years), but because of the nature of a within-subjects design in which ratings of eight pieces of evidence were each provided at three different occasions, there are still 216 data points that were collected.

### **Procedure & Materials**

The stimulus materials and procedure for Study 2 were almost identical to those of Study 1, with a few key differences. The general procedure was the same: the lawyers rated various evidence in the abstract, read about a case summary, read through 6 pieces of randomized, valenced evidence, provided their interim ratings, read through 2 more pieces of ambiguous evidence, rated the 2 ambiguous pieces, and then provided their final opinions and ratings on the case.

The public defenders volunteered to attend a 2-phase afternoon symposium in which they first participated in the present study and then heard a presentation by a professor in the Legal Psychology program at the present study's university, regarding decision-making biases (see Appendix J for recruitment blurb). The first difference in methodologies between Study 2 and Study 1 pertains to the cover story as to why the public defenders were being asked to participate. The public defenders were told that the National Bar Association is conducting a study in conjunction with the present research team that will investigate whether lawyers in different parts of the country use the same decision-making criteria related to their cases (see Appendix K for the 'cover story' and

general instructions given to the public defenders). Thus, as in Study 1, there was mild deception involved in that the Bar was not really conducting a study. Again, this deception should have been innocuous and it served as a hopefully believable reason for asking public defenders to explain their decisions.

A second difference was that, because of the low sample size, the present researchers opted to only run the control condition. As such, there were no additional instructions pertaining to justifying their decisions or playing Devil's Advocate. Fortunately, the within-subjects design involving numerous data points (the ratings of the evidence) across the decision allows for some data analysis in terms of determining if coherence effects occur in defense attorneys.

There were a few other minor changes as well, for example, the distracter questions in the abstract evidence phase of the study and the recollection of the 'original evidence ratings' were both removed in order to reduce the duration of the study. The location of the fictitious event was changed to Rockford, Illinois, instead of Florida, to reduce the likelihood that the attorneys in this area would be suspicious about never having heard of this controversial case (see Appendix L for updated Case Summary). Additionally, there were small alterations in the wording of instructions, such as asking the lawyers to remain as objective as possible before reviewing the case evidence (see Appendix M for pre-evidential instructions).

One of the most important changes between Study 1 and Study 2 concerns the dependent variables in this iteration of the study. Although the lawyers were asked to speculate about the defendant's true guilt or innocence, verdict was not the main dependent variable of interest. Instead, the public defenders assessed the case

information and decided whether or not they would recommend David Mitchell take a plea bargain or go to trial. Furthermore, they were asked to indicate the maximum number of years in prison that they would recommend Mitchell accept in a plea bargain. Finally, the lawyers were asked to respond to an open-ended question regarding their beliefs about Mitchell's guilt/innocence and their reasoning behind their recommended plea bargains (see Appendix N for the list of dependent variables and other questions given to the lawyers at the end of the study).

## **RESULTS**

### **Verdict, confidence, and other descriptives**

Eight attorneys provided a not guilty verdict and one provided a guilty verdict. Their confidence in their verdicts (measured on a 0 – 100 scale) was relatively low ( $M = 48.89$ ,  $SD = 34.80$ ), as was their beliefs in the suspect's likelihood of guilt ( $M = 2.33$ ,  $SD = 1.00$ , measured on a 1 to 7 scale, with higher numbers representing greater beliefs in the likelihood of the suspect's guilt). In keeping with their low perceptions of the defendant's guilt, eight of the nine public defenders recommended the defendant proceed to trial, instead of taking a plea bargain. All lawyers still provided a response to the maximum number of years they would have recommended Mitchell accept had he decided to go to trial, the average for which was 6.67 ( $SD = 7.00$ , range = 1 to 20).

### **Evidence Evaluation and coherence shifts**

Because the primary goal was to examine the shift in evidence over the course of the decision, and attorneys leaning towards innocence should shift in a direction opposite that of attorneys leaning towards guilt, the one public defender who voted guilty had items reverse scored.

Contrary to what Hypothesis 1a predicted, there was no significant shift from public defenders' average valenced abstract ratings of evidence ( $M = -1.83$ ,  $SD = 1.27$ ) to their valenced interim evidence ratings ( $M = -1.63$ ,  $SD = .48$ ),  $t(8) = -.56$ ,  $p = .59$ . Nor was there significant shift from their average valenced interim ratings of evidence to their average valenced post-trial ratings of the evidence ( $M = -2.10$ ,  $SD = 1.31$ ),  $t(8) = 1.04$ ,  $p = .33$ . However, public defenders' valenced abstract evidence ratings ( $M = -1.83$ ,  $SD = 1.27$ ) were significantly different from 0,  $t(8) = 4.32$ ,  $p = .003$ ,  $d = .84$ .

Similarly, contrary to Hypothesis 1b, there was no significant shift from public defenders' average ambiguous abstract ratings of evidence ( $M = -1.78$ ,  $SD = 1.42$ ) to their ambiguous interim evidence ratings ( $M = -2.28$ ,  $SD = 1.56$ ),  $t(8) = 1.16$ ,  $p = .28$ . Nor was there significant shift from their average ambiguous interim ratings of evidence to their average ambiguous post-trial ratings of the evidence ( $M = -2.28$ ,  $SD = 1.00$ ),  $t(8) = 1.60$ ,  $p = .15$ . However, public defenders' abstract ratings of the ambiguous evidence ( $M = -1.78$ ,  $SD = 1.42$ ) were significantly different from 0,  $t(8) = 3.77$ ,  $p = .006$ ,  $d = .80$ .

### **Inter-item correlations and coherence shifts**

Contrary to Hypothesis 2, there was also no significant difference in the extent to which evidence was correlated across the different points in time; the average correlation of abstract evidence (Cronbach's  $\alpha = .38$ ) was not significantly different from the average correlation of post-trial evidence (Cronbach's  $\alpha = .47$ ),  $z = .19$ ,  $p = .84$ .

### **Plea bargains and coherence**

Contrary to Hypothesis 3a, there was not a significant correlation between the maximum acceptable years in prison and the average ratings of the post-trial evidence ( $M = -2.09$ ,  $SD = 1.31$ ),  $r = .47$ ,  $p = .20$ . Contrary to Hypothesis 3b, there was no correlation

between the maximum number of years in prison the public defender would recommend accepting in a plea deal ( $M = 6.67$ ,  $SD = 7.00$ ) and their perceptions of the likelihood of the defendant's guilt ( $M = 2.44$ ,  $SD = .88$ ),  $r = .33$ ,  $p = .39$ .

### **Public defenders' predisposition towards pro-prosecution or pro-defense attitudes**

In support of Hypothesis 4, public defenders held extremely high pro-defense attitudes, as measured by the Juror Bias Scale (Kassin & Wrightsman, 1983). Again, the JBS was reversed-scored for the present study, such that lower scores indicate pro-defense attitudes and higher scores indicate pro-prosecution attitudes. The average JBS rating for the public defenders was 12.89 ( $SD = 3.59$ ), with the maximum possible score being 85. The JBS was uncorrelated with any other items of interest, such as average post-trial ratings ( $r = -.20$ ,  $p = .96$ ), likelihood of guilt ( $r = -.33$ ,  $p = .38$ ), and verdict confidence ( $r = .27$ ,  $p = .48$ ).

## **DISCUSSION**

Even though the vast majority of criminal cases do not make it to trial, there is comparatively much more empirical research on jury decision-making than there is on the decision-making of lawyers. Indeed, the factors behind how lawyer engage in plea bargaining are not yet understood. Most lawyers report that they base their plea bargains off of what they predict the jury will do (Simon, 2012). For example if they believe the jury would convict, they would recommend a plea bargain. It is worth noting that how lawyers claim they make plea bargains and how they actually make plea bargains may be two different things. If one assumes that lawyers do have insight into how they make decisions and they are basing their decisions off of the predicted verdict of the jury, then the question remains as to how they predict a jury's verdict.

The present study expected to show that public defenders experience coherence shifts in a similar fashion as jurors, however, that result was not found. This null effect could have occurred for at least two reasons. First, it could be that even using a within-subjects design, the present study lacked sufficient sample size to find effects. Second, one would expect (and their scores on the JBS indicated) that public defenders have extremely strong pro-defense attitudes. It makes intuitive sense that because public defenders spend their careers focused on how evidence is or could be construed as exculpatory, that they would tend to be unmoving in their perceptions of evidence. Effectively, they may be so used to playing the role of Devil's Advocate, constantly questioning the validity of each piece of seemingly inculpatory evidence, that they show less of a coherence effect. Future studies could manipulate the inculpatory value of evidence (low, medium, and high), to see if perhaps public defenders will experience coherence shifts when the case against the defendant is overwhelming.

There are other important aspects of lawyering that future studies should address as well. Perhaps lawyers experience coherence effects somewhat differently; for instance, if they come to believe the trial outcome looks bleak, might they put less effort into defending their client? Future research could examine coherence shifts and beliefs about the defendant's guilt or the trial outcome, and then evaluate how those factors predict the length and quality of the closing statement prepared by the defense. It could be the case that even well-intentioned, passionate public defenders succumb to coherence shifts and biased reasoning. That finding would not be surprising considering that biased reasoning, such as confirmation bias, has been shown to be a ubiquitous phenomenon (Nickerson, 1998).

Regardless of whether attorneys exhibit coherence effects, results do strongly suggest that jurors experience coherence effects. Thus, it may be beneficial for attorneys to develop strategies to attempt to minimize the likelihood that jurors will experience coherence effects. This is the purpose of Study 3.

### **STUDY 3**

#### **Goals**

The purpose of Study 3 is to examine a possible real-world manipulation that may reduce coherence effects. Study 1 allowed participants to generate their own narratives of the suspect's guilt by presenting them with both incriminating and exonerating information. In contrast, Study 3 involved creating beliefs that the suspect is guilty by exposing participants to an inculcating prosecution summary of the case. Participants will then be randomly assigned to hear about the defense's case either immediately afterwards (after which they will hear about and rate various pieces of evidence), or will only hear about the defense's case after they learn about the various pieces of evidence and rate them (after which they will then rate the various pieces of evidence again). It is expected that the prosecution's story will produce coherence effects by tainting the subsequent interpretation of evidence when participants do not hear the defense's case before evaluating the evidence. However, if participants hear the defense's story immediately after the prosecution's story, it is expected that coherence effects will disappear.

Indeed, the hypothesis that hearing about an alternative explanation will reduce coherence effects is supported by Hirt and Markman's (1995) research on considering alternative choices as a means of debiasing judgments. Hirt and Markman gave

participants a variety of information about two competing sports teams (win rates, player injuries, expert opinions, and so forth) and eventually asked them to predict who would win an upcoming game. Prior to making that prediction, participants in the ‘one explanation’ condition were asked to describe a scenario in which either: Team 1 won by a landslide, Team 1 narrowly won, Team 2 won by a landslide, or Team 2 narrowly won. Those who were asked to describe Team 2 winning, for example, were then more likely to predict Team 2 would win the upcoming game. Participants in the ‘multiple explanations’ condition were first asked to describe one of the aforementioned scenarios, but were then asked to describe one of the other outcomes. Hirk and Markman found that just by virtue of generating different plausible stories, the likelihood of being biased by the initial description of one team winning went down to baseline. It is important to note that this research is different from Disconfirmation strategies (Edwards & Smith, 1996) in which participants are told to ‘consider-the-opposite.’ In Hirk and Markman’s study, even when participants were asked to describe a landslide victory for Team 1 followed by a narrow victory for Team 1, those participants were less likely to predict a Team 1 victory to the *same extent* that participants who were asked to describe a Team 1 landslide and then a Team 2 landslide.

## **Hypotheses**

### **Hypothesis 1 – Coherence effects and the *timing* of the defense’s story**

It is predicted that coherence effects, as measured by the average ratings of evidence post-trial, will be reduced when participants read about an alternative account of the case (specifically, the defense’s explanation) immediately after the prosecution’s account and before evaluating subsequent evidence. This prediction is derived from Hirk



and Markman's (1995) research that suggests alternative explanations of a scenario reduce initial biases.

### **Hypothesis 2 – Anchoring Effect**

It is predicted that the beneficial effect of hearing about an alternative account of the case (in terms of reducing coherence effects) will be eliminated if participants first evaluate evidence before hearing about the alternative account. In other words, it is predicted that mock jurors will anchor on their initial ratings, rendering subsequent alternative accounts less potent in eliminating coherence effects.

## **METHOD**

### **Overview**

Study 1 demonstrated that participants form coherent narratives of what happened during the incident and that their beliefs about the evidence shift in correspondence with their perceptions of the defendant's guilt. Study 3, however, investigated what happens if instead of allowing decision-makers to form their own narratives, the prosecuting attorney and the defense attorney provide narratives for them. Moreover, the aim of the study was to see if presenting a compelling defense narrative immediately after the prosecution's narrative would disrupt the prosecution's narrative from forming and consolidating (as opposed to if the defense's narrative is presented later in the activity, after the prosecution's story has had time to 'sink in' and be anchored by ratings of evidence).

To examine this, Study 3 used a 2 (Defense Timing: Defense-Immediate or Defense-Delayed) x 2 (Verdict: Guilty or Not Guilty) design, with verdict as a self-selected factor. Figure 4 displays a flowchart of the study's design. All participants first

learn about the prosecution's interpretation of various evidence against a suspect in a criminal case. In the defense-delayed condition, participants then read and rated the evidence. Following that, they were presented with the defense's story and asked to rate the evidence once more, at post-trial. However, participants in the defense-immediate condition are presented with the defense's story immediately after the prosecution's story, after which they proceeded to read about the evidence and then rate the evidence at post-trial (thus defense-immediate participants only have one set of evidence ratings, at post-trial).

### **Participants**

This study was conducted online via Qualtrics and the duration of time-spent on the study was recorded. As such, 16 participants were removed for spending less than 5 minutes on a task that took most participants around 25 minutes (mean duration = 23.55).

After those participants had been removed, the remaining sample of Study 3 consisted of two hundred and seventy six undergraduate 'mock jurors' (101 males, 174 females, and 1 transgendered individual; ages 17 – 46, mean age = 22 years) from a large university located in the southeastern United States.

### **Procedure & Materials**

Participants signed up to participate in the study online. The guise was once again that local graduates of the law school at the present study's university were preparing for an upcoming trial and needed their help in evaluating the arguments and evidence in the case (see Appendix O for recruitment document). Participants who participated in Study 1 were not allowed to participate in Study 3.

The procedure for Study 3 is somewhat different than Studies 1 and 2. Participants first read about the cover story, how they are being asked to help the graduates prepare for a controversial trial. Following that, they read about the case summary, which both sides have stipulated (the case summary is the same as in Study 1). Participants were then randomly assigned to either the Defense-Immediate or the Defense-Delayed condition. Participants in the Defense-Delayed condition were presented with the Prosecution's account of the case, which painted David Mitchell as a hot-headed, dangerous coworker of the victim who had been slighted on an important project (see Appendix P for the Prosecution's Account of the Case). Following that story, participants read about the specific pieces of evidence in this case. The order of the evidence was randomized, but the valence of the evidence was held constant. Thus, participants received 3 pieces of incriminating evidence: Eyewitness Identification, Fingerprints, and Phone Records and 3 pieces of exonerating evidence: Traffic Camera Footage, Polygraph Evidence, and Shoe Print Evidence (these evidence vignettes were the same as in Study 1, see Appendix F). After having received that evidence, participants were asked to rate the degree to which each piece of evidence implied David Mitchell was guilty or not guilty on a scale that ranged from -7 (implies not guilty) to 7 (implies guilty). It was expected that after having had the prosecution's story presented as a framework through which to perceive the evidence and after having anchored their opinions by rating the evidence, that participants in this Defense-Delayed condition would perceive the case evidence as strongly inculpatory. Once the ratings were given, participants proceeded to read about the Defense's Account of the case, which focused on a counter-narrative that provided an alternative explanation for the prosecution's

evidence. Specifically, the defense asserted that: this was a random killing in the park, there was no real evidence against Mitchell, and that it could not have been a ‘heated’ crime of passion because there was a several hour cooling-off period between the phone conversation with the victim and the murder (see Appendix Q for the full Defense Story).

In the Defense-Immediate condition, immediately after reading the Prosecution’s Story, participants read the Defense’s Story. Following the two accounts of the case, participants read about the specific evidence.

All participants then proceeded to render a final verdict and then rate the evidence at post-trial. Additionally, participants provided an ‘extent of guilt’ rating on a -7 (not guilty) to 7 (guilty) scale, provided a confidence rating on a 1 (not confident) to 7 (very confident) scale, and then rated all of the specific evidence one more time. Following that, as with the other two studies, participants described their verdicts in an open-ended format, answered believability questions, answered a manipulation check (see Appendix R), provided demographic information, and then finally read a debriefing statement.

## **RESULTS**

### **Evidence Ratings**

The 6 pieces of post-trial evidence were averaged together to create a post-trial evidence rating score, with higher scores representing a stronger perception that the evidence is incriminating. A 2 (verdict) X 2 (timing of defense story: immediate or delayed) ANOVA revealed no significant interaction between verdict and defense story timing on the post-trial ratings of evidence,  $F(1, 272) = .18, p = .68$ . However, there was partial support for Hypothesis 1 in that there was a marginal main effect for the timing of the defense’s story,  $F(1, 272) = 3.03, p = .08$  and a significant main effect of verdict,  $F(1,$

272) = 194.32,  $p < .001$ ,  $\eta_p^2 = 1.00$ . Simple follow-up tests revealed that among convicters, there was no difference between the Defense-Delayed ( $M = 1.45$ ,  $SD = 1.65$ ) and the Defense-Immediate condition ( $M = 1.46$ ,  $SD = 1.58$ ),  $F(1, 100) = .81$ ,  $p = .37$ . However, among acquitters, there was a marginally significant difference such that participants in the Defense-Immediate condition ( $M = -1.66$ ,  $SD = 1.74$ ) rated the evidence as less inculpatory than acquitters in the Defense-Delayed condition ( $M = -1.19$ ,  $SD = 1.87$ ),  $F(1, 172) = 2.93$ ,  $p = .09$ . Across all participants, post-trial evidence ratings were not significantly different on account of whether participants were in the Defense-Immediate condition ( $M = -.47$ ,  $SD = 2.26$ ) or the Defense-Delayed condition ( $M = -.15$ ,  $SD = 2.28$ ),  $F(1, 274) = 1.40$ ,  $p = .24$ .

In terms of evidence integration, the inter-item reliability for the six pieces of evidence can be compared across conditions to evaluate the extent that evidence ‘hangs together’ and represents a construct of guilt or innocence. Among convicters, there was no significant difference between the inter-item correlation of evidence for those jurors in the Defense-Delayed condition (Cronbach’s  $\alpha = .52$ ) compared to the Defense-Immediate condition (Cronbach’s  $\alpha = .58$ ),  $z = .42$ ,  $p = .68$ . Among acquitters, there was once again no significant difference between the inter-item correlations of evidence for jurors in the Defense-Delayed (Cronbach’s  $\alpha = .70$ ) condition versus those in the Defense-Immediate condition (Cronbach’s  $\alpha = .59$ ),  $z = 1.23$ ,  $p$  (one-tailed) = .11. Across all participants, there was no difference in correlations for jurors in the Defense-Delayed condition (Cronbach’s  $\alpha = .78$ ) and the Defense-Immediate condition (Cronbach’s  $\alpha = .77$ ),  $z = .21$ ,  $p = .83$ .

In support of Hypothesis 2, a 2 (verdict) X 2 (evidence ratings interim vs. post-trial) repeated measures ANOVA revealed that there was a significant interaction between verdict and evidence ratings,  $F(1, 130) = 12.39, p = .001, \eta_p^2 = .74$ . Follow-up tests of simple main effects revealed that convicters did not experience coherence shifts from interim ( $M = 1.96, SD = 1.81$ ) to post trial ( $M = 1.74, SD = 1.65$ ), on the basis of the defense's story,  $t(46) = 1.03, p = .30$ . However, acquitters did shift in their evidence ratings from interim ( $M = -.04, SD = 1.71$ ) to post-trial ( $M = -1.19, SD = 1.87$ ),  $t(84) = 7.08, p < .001, d = .61$ .

### **Descriptives for verdict, confidence, and extent of guilt**

**Verdict.** Of the 276 participants, 102 provided guilty verdicts and 174 provided not guilty verdicts. The proportion of verdicts is similar to Study 1. There was no significant difference in verdicts between the Defense-Delayed condition (36% Guilty and 64% Not Guilty) and the Defense-Immediate condition (38% Guilty and 62% Not Guilty),  $\chi^2(1) = .19, p = .71$ .

**Confidence.** The mean confidence-in-verdict across participants was  $M = 4.30, SD = 1.35$ . There was also no difference in confidence between the Defense-Delayed condition ( $M = 4.24, SD = 1.40$ ) and the Defense-Immediate condition ( $M = 4.39, SD = 1.31$ ),  $t(274) = .89, p = .37$ .

**Extent of Guilt.** Participants were also asked to rate the extent of the defendant's guilt, ranging from -7 (not at all guilty) to 7 (very guilty). There was no significant difference between the Defense-Delayed condition ( $M = -.78, SD = 3.89$ ) and the Defense-Immediate condition ( $M = -.36, SD = 3.86$ ),  $t(274) = -.90, p = .37$ . A follow-up analysis can be run on extent of guilt by splitting the output by convicters vs. acquitters.

Among convicts, there was no significant difference between the Defense-Delayed condition ( $M = 3.51$ ,  $SD = 2.29$ ) and the Defense-Immediate condition ( $M = 3.93$ ,  $SD = 1.36$ ),  $t(100) = -1.13$ ,  $p = .26$ . The mirror analysis was also run on acquitters and there was no significant difference between participants in the Defense-Delayed condition ( $M = -3.15$ ,  $SD = 2.17$ ) and the Defense-Immediate condition ( $M = -3.01$ ,  $SD = 2.11$ ),  $t(172) = -.44$ ,  $p = .66$ .

## DISCUSSION

The results indicated that when the prosecution's account of the case was immediately followed by the defense's account of the case, acquitters perceived the post-trial evidence as marginally more exculpatory than when the defense did not provide their account of the case until later. Additionally, there was a marginally higher inter-item correlation among evidence for acquitters who heard the defense's story immediately after the prosecution than those who heard it later on in the trial, suggesting that the defense's story disrupted the extent to which jurors' evaluation of the evidence hung together. In other words, this finding suggests that hearing the defense's account of the evidence reduced coherence effects. However, this effect was only marginally-significant, so must be taken with caution, and it only applied to acquitters. Nonetheless, these results provide preliminary support for the idea that if the defense presents its story to the jury as soon as possible, the jurors will be more likely to perceive the defendant as innocent. Thus, for example, if a trial began in the afternoon and the defense had the option to either follow the prosecution's opening statement that day or present their case the next morning, it seems beneficial to get the counter-narrative to the jurors as soon as possible.

While the idea of counter-narratives is widely supported among lawyers (Scheb & Scheb, 2011), a counter-narrative alone may not be enough to sway jurors. After all, jurors likely develop their own accounts of what happened, beyond the story that the prosecution or the defense attempt to weave. Indeed, even when jurors were asked to write down the flaws, issues, and problems with each piece of evidence in Study 1, coherence shifts still remained unmitigated. Similarly, even if a defense attorney were to do an excellent job pointing out the problems with the prosecution's evidence, that still may not prevent inculcating coherence shifts on the part of jurors. In order to reduce coherence shifts, in addition to getting a counter-narrative out to the jurors as fast as possible, a defense attorney could also make use of debiasing or disconfirmation strategies (Edwards and Smith, 1996). Disconfirmation strategies involve waiting until a decision-maker has reached a final conclusion and then asking him/her to list reasons why that decision could be wrong. Additionally, Simon (2004) found that disconfirmation strategies were successful at reducing coherence effects as well. Future research should examine whether similar effects occur among jurors making decisions in the context of a criminal trial.

Thus, a savvy public defender who is aware of the literature on coherence effects may want to use a multifaceted approach to influencing jury decisions. The public defender might: begin by administering the JBS during voir dire to ensure the pro-defense veniremen are selected, commence their opening story to the jury as soon as possible (as the present study suggests), continue to develop a strong counter-narrative as the trial progresses, and then, during closing statements, end by asking any jurors who are leaning towards guilt to take some time and consider how that conclusion could be



flawed (Edwards and Smith, 1996; Simon, 2004). Given the strong relationship between jurors' scores on the JBS and coherence effects found in Study 1, the voir dire phase may be the most important in swaying the outcome of a trial. Recall that Study 1 found the direction jurors experience coherence shifts in, the magnitude of those shifts, the perceptions of the defendant's overall likelihood of guilt, and so forth, were all significantly related to pro-defense or pro-prosecution attitudes.

A more general, but nonetheless important, recommendation to be derived from coherence effects is that public defenders should not discount the potential impact of a piece of evidence. Because of the inter-related nature of evidence integration, as Study 1 demonstrated, any single piece of evidence may cause a snowball effect that effects the evaluation of all subsequent pieces of evidence.

## **GENERAL DISCUSSION**

### **Coherence effects in jurors**

The results of the Study 1 demonstrated that although jurors begin trials with relatively neutral and uncorrelated perceptions of the evidence, as the trial progress, jurors begin to lean towards either guilty or not guilty. By the time they have reached the end of the process, jurors have polarized in their judgments and are confidently in favor of their selected verdicts. Coherence shifts present a host of problems, in particular how they interact with legal standards, instructions, and policies.

**Polarization of evidence evaluations and the 'reasonable doubt' standard.** It is critical to remember that the probative value of each piece of evidence in Study 1 never changed. However, decision-makers' perceptions of the evidence did change significantly over the course of a one hour study, in accordance with verdict. From a

logical standpoint, the fact that the evidence never changes but the perceptions of the evidence change would constitute a decision-making error. Additionally, that perceptions of the evidence shift away from the abstract rating is logically unsound because the abstract rating constitutes the purest rating of the probative value of that evidence in isolation. Moreover, one cannot blame the shifts observed in the current study on the particularities of the trial or on the inculcating power of any one piece of evidence, because the evidence order and evidence valence were randomized. Each participant essentially had a different trial, and yet we observed the same dramatic coherence effects regardless of the vagaries of the trial.

Furthermore, it is alarming that confidence shifts occur in tandem with shifts in the perception of evidence. In fact, Study 1 found that the more jurors shifted in their opinions of the trial evidence the more their confidence increased. Confidence also shifted from the interim of the trial to the end of the trial as well, further bolstering its effect. Moreover, the present study found that confidence was significantly correlated with jurors' ratings of the defendant's 'likelihood of guilt.' Those results are especially troubling because confidence is a critical measure in legal decision-making as it relates to whether or not the standard of 'beyond a reasonable doubt' will be surpassed (Simon, 2012).

If a juror is at a level of near certitude in his or her verdict, then that surpasses the reasonable doubt standard. Adding to the problem, the standard most jurors use for conviction is much less than near-certitude. On average, uninstructed jurors tend to use a threshold of 78%, even though legal scholars claim the standard should be 90-99% (Simon, 2012). When jurors are instructed how to use the standard, roughly two-thirds of

jurors claim to have understood the instructions, but in reality, only about one-third of jurors understood.

Thus, the fact that coherence effects promote upward shifts in confidence is an alarming problem when one considers the potential for wrongful convictions. Legal standards of proof such as the reasonable doubt standard only protect defendants to the extent that the standard actually functions as it is intended. The relationship between coherence effects and the reasonable doubt standard illustrates a critical difference between how the law is prescriptive but psychology is data-driven. The law establishes what the standard of proof *should* be, but one of the roles of legal psychology is to determine whether the data support the claim that the standard protects defendants. The cases that are clearly lopsided rarely make it past plea bargaining, but the complex and ambiguous cases are the ones that go to trial (Scheb & Scheb, 2011). As such, it is easy to imagine how coherence effects could cause jurors who are on the fence to eventually polarize in their beliefs of the defendant's guilt *and* to become overly confident in their beliefs, to the point where they are prepared to render a verdict of guilty beyond a reasonable doubt. The present study found that coherence effects were hard to reduce, but perhaps future research can take an alternative approach: reducing confidence in jurors and observing the effect of reduced confidence on coherence shifts and the potential to find the defendant guilty beyond a reasonable doubt.

It is worth noting that the jurors in the present study always received three pieces of inculpatory evidence and three pieces of exculpatory evidence. Logically, then, there should almost always have been sufficient doubt in jurors' minds, as half of the evidence in the case pointed to the defendant being not guilty. Even when a case is deliberately

designed to be ambiguous and neutral, as the present study was, over one third of the jurors still found the defendant guilty. One concept that could explain why jurors in this case found the defendant guilty despite there being at least equal-parts exculpating evidence is asymmetrical skepticism.

**Asymmetrical skepticism.** Research is beginning to converge on a trend that suggests it is easier to elicit judgments of guilt than it is to elicit judgments of innocence. For example, Charman, Carbone, Kekessie and Villalba (under review) found that when strongly inculcating evidence was presented prior to ambiguous evidence, the jurors' perception of the ambiguous evidence became more inculcating. The reverse effect was not found; when strongly exculpating evidence was presented prior to ambiguous evidence, the ambiguous evidence was not perceived as more exculpating. Those findings dovetailed with what the present study found.

Study 1 found that jurors who were high in pro-prosecution attitudes, as measured by the JBS, actually experienced a greater magnitude of coherence shifts. In other words, generally pro-prosecution jurors were more readily willing to perceive increasing inculcating value in the evidence over the course of the trial. Furthermore, Study 1 found that the direction of jurors' shifts are predictable, as they are tied to their general pro-prosecution and pro-defense attitudes: A pro-prosecution juror will perceive valenced evidence as more inculcating, ambiguous evidence is more inculcating, and the defendant as more likely of being guilty, than will a pro-defense juror. Additionally, Study 1 found that convicts became more convinced of the defendant's overall likelihood of guilt from interim to post-trial, whereas acquitters did not become more convinced of the defendant's likely innocence.

Those findings are problematic because defendants are supposed to be presumed innocent until proven guilty. However, the data from Study 1 demonstrate that, if anything, it is easier for pro-prosecution jurors to shift more in their opinions than it is for pro-defense jurors. Importantly, the finding that pro-prosecution attitudes predict whether a juror will lean towards guilty and perceive the evidence as inculcating also works against the presumption of innocence. One can imagine a scenario in which a jury that happens to have numerous pro-prosecution jurors comes to convict a defendant even though there is only ambiguous or shoddy evidence against him. Pro-prosecution jurors will polarize in their opinions of the evidence, become highly confident, and find the defendant guilty beyond a reasonable doubt. As Study 1 also found, coherence shifts are very difficult to reduce, as such judicial reminders about the burden of proof being on the prosecution are likely to fall on deaf ears.

Study 3 also found support for asymmetrical skepticism. The interim evidence ratings of convicts in the Defense-Delayed condition did not significantly shift in their perceptions of the evidence at post-trial after hearing the defense's account of the case. Eventual acquitters, though did shift in their ratings from interim to post-trial, such that their perceptions of the evidence became more exculpating. Thus, it seems that convicts were indeed skeptical of the defense's story, unlike acquitters. In other words, beliefs of guilt were more resistant to contrary evidence than were beliefs of innocence.

One last point in support of asymmetrical skepticism is that among switchers in Study 1, the only coherence shift observed was among convicts and took place from interim to post-trial. Furthermore, switchers, who constitute a group of uncertain decision-makers, were more willing to render a guilty verdict (50% voted guilty) than

non-switchers (34.65% voted guilty). This indicates that uncertain, low confidence decision-makers were more readily willing to perceive the case against Mitchell as inculcating than exculpating.

Taken together, the aforementioned series of results point strongly toward jurors having asymmetrical perceptions of the evidence and of Mitchell's guilt. It seems all too easy for jurors to perceive guilt in the evidence, but they seem resistant to equally perceive innocence. Simon (2004) describes coherence effects as a constraint satisfaction mechanism, which could contribute to explaining asymmetrical skepticism. If a juror is presented with hard evidence against a defendant, comes to believe that defendant is guilty, and is then presented with an alibi, that alibi *must* be false. The perception of the alibi is constrained by the beliefs of guilt. However, if a juror is presented with exculpating evidence and comes to believe the defendant is innocent, the alibi could still be false. Jurors could still believe that the defendant was innocent but rationalize that maybe he incorrectly remembered where he was on the evening in question. Inculcating evidence should tend to constrain outcomes more than does exculpating evidence. The focus of the problems with coherence effects has thus far been focused on evidence evaluation, but there are problems in terms of evidence integration as well.

**Evidentiary independence and evidence integration.** To some degree, the general wording in judicial instructions regarding evidence states that evidence is largely independent and that it ought to be weighed in the calculus of rendering of a verdict. However, the results of the present study strongly contradict that assumption and demonstrate more ways in which coherence effects are problematic.

The present study found that the inter-item correlations between the ratings of evidence steadily increased as the trial went on. Indeed, the difference in correlations from abstract to interim and interim to post-trial were all significant. By the end of the trial, the evidence all ‘hung together’ and pointed toward the defendant either being guilty or being not guilty. In that sense, the evidence transformed from separate, distinct facets of a decisions, to a central construct of guilt or innocence.

The interrelationship between pieces of evidence is problematic for a variety of reasons. Logically, each piece of evidence should be a mostly discrete entity. If an eyewitness identifies a perpetrator, that identification is wholly separate event from whether a forensic analyst deems two fingerprints a match. However, as shown in Study 1 (and consistent with similar findings, e.g., Kassin, Dror, & Kukucka, 2013), knowing about one piece of evidence can bias the interpretation of other pieces of evidence. One consequence of the integration and interdependence of evidence is that preliminary leanings towards guilt or innocence can quickly snowball into polarized and confident beliefs of guilt. One way to elaborate on this snowball effect among evidence is to conceptualize it as a bias.

***Biased judgments.*** The present study demonstrated that jurors came to be polarized towards either guilt or innocence, even though they started off mostly neutral in their judgments of the evidence. It is worth repeating that beliefs of guilt or innocence were not instilled or manipulated; instead, participants shifted in one direction or the other of their own accord (as predicted by their pro-prosecution or pro-defense attitudes). When it came time to read and rate the ambiguous evidence at the interim of the trial, participants rated that evidence in accordance with whichever verdict they were leaning

towards. In other words, participants experienced a form of self-generated bias; an artifact of the decision-making process. In this sense, coherence effects are particularly pernicious in that even when there is no outside biasing event to point to, bias could still be present just because of the nature of human cognition. Pondering the extent that evidence is interrelated leads to a variety of follow-up issues regarding the context in which the evidence is presented, for example, the *order* of the evidence.

***Evidence Order.*** The order in which evidence is presented to a decision-maker should not affect the decision's outcome, according to a rationalist point of view. Nevertheless, the results of several studies are beginning to converge on a 'recency effect,' whereby the last piece or pieces of evidence presented has the strongest impact on decisions (Dahl, Brimacombe, & Lindsay, 2009; Charman, Carbone, Kekessie, and Villalba, under review; Price & Dahl, 2013). Charman et al. varied the order in which participants learned about an impactful piece of evidence in relationship to an ambiguous piece of evidence. Specifically, participants either received DNA evidence *first*, before they were presented with an ambiguous alibi, or *second*, after they had learned about the alibi. The result was that when DNA evidence was presented *second*, it had a significantly stronger impact on ratings of the defendant's guilt than when it was presented *first*.

Interestingly, the present Study 1's results coincide with those of Charman et al. (under review). In the present study participants in the control group were significantly more likely to find the defendant guilty if the prosecution's evidence came *second*, after the defense's evidence. Likewise, participants were more likely to find the defendant not guilty when the defense's evidence was present *second*.



Taken together with Charman et al.'s (under review) findings, the results of the present study suggest that the most recent evidence one hears about becomes the most salient in the mind of decision-makers. As Simon (2004) has pointed out, the decision-making process is bi-directional, whereby evidence influences the potential outcome, and potential outcomes feed backward to influence perceptions of the evidence. The recency effect discussed here points to the idea that the most recent piece of evidence one evaluates can *retroactively* alter the overall judgments of the defendant's likelihood of guilt and one's verdict selection.

More research in this area is needed to flesh-out the effects of order. For example, future studies should consider including substantial delays between the presentation of evidence and investigate how that moderates the effect of order. Charman et al.'s (under review) study included a manipulation of a short, five minute delay between the two pieces of evidence and found no effect. However, outside of brief laboratory studies in academic research, real trials can often take days or weeks. If jurors leave the courtroom and return to their homes after a day at trial, the most recent piece of evidence they heard about has the potential to ruminate with them all night and until they are back at trial the next day, which could increase the potency of the recency effect.

Future research should also include more than two sets or pieces of information, to explore order effects in more detail. Additionally, observing the interaction between the order of numerous pieces of evidence and the delay between pieces of evidence could prove fruitful, as that scenario more closely mirrors how real trials proceed. It would be surprising if some stronger effects for primacy were not found, in that the first item has

the potential to taint the perception of the information that follows and the overall decision.

**Lack of awareness in shifts.** Perhaps the most dangerous aspect of coherence shifts is the profound lack of awareness that decision-makers have of their shifting opinions. Similar to Simon's (2004) findings, participants in the present study did not believe their opinions had shifted much. Moreover, when asked to recall their original ratings of the evidence, they failed to do so accurately. There was a significant difference between their abstract ratings of evidence and their memory of the abstract ratings. There was, however, no significant difference between their memories of the abstract evidence and their post-trial ratings of the evidence. Meaning, decision-makers came to believe their most recent attitudes were in fact how they had felt all along. Thus, a juror who perceived shoe-print evidence as neutral at the start of the trial but comes to believe the defendant is guilty, may erroneously think the shoe-print evidence implied the defendant was guilty all along.

One would be hard-pressed to convince a police investigator, a lawyer, or a juror that their opinions had changed because people have an esteem-related drive to think of themselves as consistent (Baumeister & Bushman, 2008). Furthermore, the increased confidence observed in coherence shifts alongside the lack of awareness can account for cases like the Central Park Five, where prosecutors and police detectives were unwilling to abandon their beliefs about the defendants' guilt in the face of DNA evidence to the contrary (Costanzo & Krauss, 2012). If a police investigator confidently swears on the witness stand that the defendant was the most likely suspect all along, the police investigator may believe that is true when he says it. But in reality, the police

investigator may have been seriously considering other suspects at the onset of the investigation. The findings regarding lack of awareness in coherence shifts present a dangerous problem not just on their own, but especially coupled with the fact that coherence shifts are very hard to reduce.

**Countermeasures against Coherence shifts.** One of the goals of the present study was to find a potential policy or procedure that could help reduce coherence shifts. For example, it was predicted that by asking participants to consider the flaws inherent in the evidence, as they were instructed in the Devil's Advocate conditions, that the depth of processing would be increased and the self-generated criticism of evidence would lead to a reduction in coherence effects. Nevertheless, the coherence shifts remained unmitigated. It was originally predicted that telling participants to anticipate having to justify their final decisions would exacerbate coherence effects, because of self-presentation needs and the desire to appear certain and consistent. Yet, it also remained possible that asking participants to justify their decisions may have increased accountability, thus reducing coherence effects. Once again, coherence effects were not reduced by anticipating being accountable for one's decision. Coherence effects appear to be quite resistant to countermeasures.

Future research needs to continue investigating ways to reduce coherence effects, or, at minimum, investigating moderators of coherence effects so as to understand how to prevent them. One potential idea is to keep decision-makers cognitively busy. It is possible that for information to cohere, one must have sufficient cognitive processing available to integrate the evidence into a coherent conclusion. Another potential way to reduce coherence effects could be via an inoculation strategy, whereby the process and

outcome of coherence effects are explained to decision-makers before the decision task begins. If this increases awareness of the potential for coherence effects, participants may be able to prevent them from happening.

***Jury Deliberations as a means of reducing coherence effects.*** An anticipated criticism of the present study is that reducing coherence effects is unnecessary because they are individual phenomena that occur among single jurors, and jury deliberations should negate the impact of coherence shifts. One may assume, in other words, that coherence effects may be ‘washed out’ by the discussion of contradictory opinions. But that does not seem to be the case. Simon (2004) found that deliberations do not reduce coherence effects, and, in some ways, can actually exacerbate them. If a juror comes into deliberations with a pre-deliberation belief of guilt and tries to explain to other jurors why he is right, he will likely become even further entrenched in his already polarized position. This may explain why one study found that there were zero available records of a lone juror swaying a trial (Simon, 2012). More importantly, in 94% of cases, the majority opinion of either guilt or innocence cast in the opening ballot is what ended up being the final verdict. In other words, if seven out of twelve jurors vote to acquit as soon as they sit down in the jury room, then acquittal will almost surely be the result. These findings are easily explained with an understanding of coherence effects.

***Reducing the potential for miscarriages of justice by improving investigations.*** Given that neither jury deliberations nor the manipulations in the present study were able to reduce coherence effects, an alternative approach merits discussion: to reduce problems early on, during the police investigation phase, before unlikely suspects get to trial and end up at the mercy of jurors, in whom coherence effects have been well-

documented. Stopping a snowball from becoming an avalanche is easiest before it has had time to gather momentum. As such, the investigative phase may be the most critical phase of all for preventing biased reasoning, in that it is the first phase in the legal process.

Although coherence effects per se have not yet been examined within the context of police investigators, there is a substantial body of research devoted to investigator tunnel-vision and confirmation bias. For example, Ask and Granhag (2007) found that when police investigators assessed the reliability of different types of evidence, those assessments were biased by their pre-existing beliefs about the suspect's guilt. Similarly, Kassin, Goldstein, & Savitsky (2003) found that when interrogators had beliefs of guilt instilled in them, they interacted more negatively with potential suspects.

Unfortunately, there is a compendium of problems that can arise from investigator tunnel-vision. The way in which beliefs of guilt can taint perceptions of evidence can be referred to as pseudo-corroboration. In other words, if an investigator goes to follow-up on an alibi, if he believes the suspect is guilty, he may look for signs of nervousness in the alibi witnesses. In turn, feeling more assured of his initial suspicions, he may be less thorough in pursuing other potential suspects. It would be during this time that the proverbial 'snowball' begins to gain momentum; beliefs in the suspect's guilt result in more biased evaluations of evidence, further bolstering beliefs in the suspect's guilt. This escalation of error can result in an innocent person going to trial; similar coherence effects may then result in jurors perceiving guilt as well.

A clear example of escalation of error can be seen with the case of *The People v. Beaman* (2006), in which Alan Beaman was accused of murdering of Jennifer Lockmiller

in Illinois. Beaman was convicted on essentially no hard evidence, in fact, the evidence that did exist, such as phone records, store receipts, and odometer mileage, pointed to his innocence. Nevertheless, he was convicted and spent thirteen years in prison until the Center for Wrongful Convictions at Northwest University took up his case and overturned it. The police investigators in this case were affected by tunnel-vision and never bothered to follow-up on other potential suspects. In this case, there were other more likely suspects, such as drug dealers to whom the victim owed money. The initial belief of guilt among investigators led them to perceive more guilt and pseudo-corroboration in flimsy circumstantial evidence, such as neighbors reporting Beaman and Lockmiller arguing. Not only did the escalation of error lead to a wrongful conviction, but subsequently, no one else has been tried for Lockmiller's murder, and the killer remains at large. Beaman's case emphasizes that preventing decision-making errors at their inception is critical.

An important recommendation regarding the investigation phase centers on making it more transparent. For example, more jurisdictions have begun to record interrogations and to keep records of line-ups (Simon, 2012). Those policies are beneficial in terms of reducing the potential for decision-making errors during trial. If a juror hears from a police investigator about the details of an interview, that juror may come to a markedly different conclusion than if the juror watches the raw footage of the interrogation itself. The decision-making ability of jurors is related to the quality and probative value of the evidence. Indeed, whereas jurors return to their regular jobs after a trial, police investigators may have their work-related integrity and self-esteem tied up in their decisions. Thus, erroneous initial opinions could lead to an investigator

exaggerating claims about the interrogation on the witness stand. Second-hand or synthesized evidence, such as an investigator's testimony about the interrogation, is bound to be less pure and subject to more human error than the actual footage of the interrogation is.

Future research should investigate the role that coherence effects play in the decisions of investigators, such as whether or not to arrest a suspect or which suspect to pursue. This phenomenon could be similar to findings among public defenders in Study 2; police investigators may hold strong pro-prosecution attitudes and thus be more resistant to coherence shifts in the direction of innocence. Although Study 1 had a measure of how much pressure jurors felt to come to their decision, it was not significantly related to any variables and that is probably because jurors felt very little pressure ( $M = 1.05$ ;  $SD = 3.76$ ; with maximum pressure being a 7). Police investigators likely experience a great deal of pressure from needing to close cases, answer to the media for highly publicized crimes, prevent repeated offenders from committing more crimes, and so forth. Future studies should examine if pressure moderates coherence effects.

***Coherence shifts in the Appellate process.*** One more reason that improving decision-making early on is critical is that because once the legal process has arrived at the appellate phase, the extent of the evidence integration may be too strong to undo, as a result of coherence effects. In general, for a judge to overturn a lower court's ruling, there must have some aspect of due process that was violated or some manner of severe error (Scheb & Scheb, 2011). Indeed, appeals that cite 'insufficiency of evidence' almost always fail. Sixty nine of the one hundred and sixty five DNA exoneration cases cited

insufficient evidence for conviction, and yet, only *one* of those appeals was granted relief on the basis of lacking evidence (Simon, 2012). That figure is troubling in light of all of the errors in evidence that were made apparent, for example, the use of shoddy forensic science as evidence.

Even if the appeals court were to acknowledge issues with evidence more often, it is unlikely there would be much more relief granted. The appeals process is governed by the doctrine of harmless error, which states that unless an error at trial was so profound as to undermine the entire case, then the end result of the trial would have likely been the same and overturning the verdict is unnecessary (Scheb & Scheb, 2011). In other words, appellate judges tend to assume that removing a single piece of evidence from trial has no impact on how the other evidence was evaluated; it is often assumed to have been a harmless error. The concept of *harmless error* represents a somewhat naïve view of decision-making and it is in direct conflict with the findings of coherence effects. As Study 1 demonstrated with the inter-item correlations between evidence, different pieces of evidence are not evaluated independently of one another, but rather are interdependent. And the correlations between the various pieces of evidence grow stronger as a trial progresses. Thus, if a compelling but invalid piece of evidence would have been removed early on in the trial, it may have been the critical piece of information that constrained the evaluation of other pieces of evidence. Beyond the present study, other research has found that pieces of evidence are not independent and that one piece of information can strongly influence another. For example, Hasel and Kassin (2009) found that when eyewitnesses made an identification from a line-up, but subsequently learned that a different person from the lineup confessed, the majority of the witnesses actually



switched their identification to the lineup member who confessed. That finding speaks to the extent to which pieces of evidence are interdependent, in that learning about the confession was powerful enough to make the witnesses doubt their own memories, thus corrupting the identification procedure. In keeping with previous examples, the removal of a piece of evidence may have stopped the snowball from gaining momentum.

It is critical for people in the legal system to understand how various pieces of evidence are integrated into final judgments, and how preliminary beliefs of guilt can bias the subsequent interpretation of evidence, further increasing beliefs of guilt, resulting in a snowball effect. One reason appeals court judges may be reluctant to see error as ‘harmful,’ is that they have trouble imagining how one piece of evidence can cause jurors’ opinions of other pieces of evidence (and consequently of the guilt of the suspect) to shift so dramatically. Indeed, jurors themselves in Study 1 had trouble conceiving that their own opinions had shifted and changed just over the course of a brief, one hour study.

### **Limitations of the present series of studies**

The present three studies have their limitations and, as always, inferences should be drawn with caution until convergent validity is achieved by future studies replicating the present studies’ results. One important limitation of Study 2 was the low sample size of public defenders able to participate in the study. Future research should continue to investigate coherence effects in attorneys and should provide substantial incentives to encourage attendance and participation.

Another potential limitation concerns the tradeoff between internal validity and external validity. The present studies were very high on internal validity, for example,

evidence order was randomized and there were no differences between the control groups and experimental groups other than the variables that were being manipulated (instruction type in Study 1; timing of defense story in Stud 3). However, the present studies were not as high in external validity, in that the jurors were not in a real courtroom watching a live case, nor was a defendant's liberty truly on the line. The impact of those factors, though, should be minimal in that the cognitive processes behind decision-making, such as coherence shifts, will be present regardless of whether one is making a 'real' decision or fictitious decision. Furthermore, the present series' of studies instilled some gravity to the participants' decisions in that participants were told their results would be used by lawyers to prepare for an upcoming trial (or, in the case of Study 2, public defenders were told their results would be reviewed by the National Bar Association).

### **Overview of decision-making in the legal system**

The United State's Supreme Court has stated, regarding jury decision-making, that one cannot know "The imponderables which cause one to think the way they think" (Simon, 2012, p 192). Yet, it is critical that researchers do investigate the imponderables behind decision-making if wrongful convictions are to be prevented. The DNA exoneration cases brought to light by the Innocence Project (2014) have illuminated some egregious problems within the legal system, ranging from mistaken eyewitness identifications, shoddy forensic science, and false confessions. However, those issues may be reflective of a deeper underlying *cause* of the problem, which is that people in the legal system engage in biased decision-making and exhibit coherence effects. For example, if a forensic examiner is comparing two fingerprints to determine whether or not they are a match, the decision he/she makes is subject to bias on the basis of what the

examiner knows about the suspect or the examiner's motivations (Dror & Cole, 2010). Similarly, biased reasoning could determine whether a police investigator focuses on pursuing one lead or another, whether a juror comes to verdict of guilty or not guilty in an ambiguous case, or whether judge decides to grant relief on an appeal or not. It falls to legal psychologists to systematically research the cognitive processes behind decision-making and to determine practical applications for reducing biased reasoning. Jury trials are an aspect of the legal system that is particularly rife with the potential for coherence effects and biased decision-making.

The institution of trial by jury is considered the pinnacle of justice and lauded as the crowning achievement of the U.S. legal system (Scheb & Scheb, 2011). Indeed, the legal system goes to great lengths to ensure that due process is followed and that defendant's sixth amendment rights to an impartial jury are maintained. Additionally, great faith is put in what jurors can do, ranging from: their ability to understand and follow judicial instructions, to spot shoddy evidence, to discern liars from truth-tellers, to remain impartial, and to weigh evidence independently and come to a reasonable decision. These are no small tasks and the question remains as to whether jurors are up to the challenge.

There is *some* limited evidence that suggests that jurors make relatively good decisions. For example, the strongest predictor of a jury's verdict is the strength of the evidence (Costanzo & Krauss, 2012). However, the correlation between evidence strength and conviction rates is only .50, thus accounting for only 25% of the variance in conviction rates. One must wonder what other factors account for the remaining 75% of the variance. Other research also points to jurors being reasonable decision-makers. For

example, a seminal study by Kalven and Ziesel (1966) revealed the fact that juries and judges agree 75% of the time. However, there is no reason to suspect that the cognitive processes (and the cognitive biases) behind the decision-making of juries and judges are any different. In other words, the same factors that might taint a juror's judgment would be likely to taint a judge's judgment.

In contrast, exonerations proven by DNA evidence (Costanzo & Krauss, 2012) have demonstrated that wrongful convictions and decision-making errors are a genuine concern. In terms of indices of accuracy, DNA exonerations, because of their high accuracy rates, would seem substantially more objective and reliable than jury-judge agreement. Indeed, there is a rather large body of literature suggesting that jurors are poor at a variety of tasks: utilizing judicial instructions (Haney & Lynch, 1994), disregarding prejudicial information from the media (Ruva & McEvoy, 2008), understanding expert witnesses (Cutler, Dexter, & Penrod, 1990), and so forth. But perhaps the most scathing indictment of jury decision-making comes in the form of their susceptibility to coherence shifts (Holyoak & Simon, 1999; Simon, 2004; Simon, 2012; Simon, Pham, Le, & Holyoak, 2001).

### ***Shift in focus***

Coherence shifts represent a serious threat to the legitimacy of the legal system. Coherence shifts are predicted by general pro-prosecution and pro-defense attitudes (which strongly determine the side of the case a juror will favor and how he/she perceive the evidence), occur spontaneously (without the introduction of any external bias), are capable of biasing the perception of subsequent ambiguous evidence, are outside of conscious awareness, happen quickly, and are difficult to mitigate. More research is

needed to discover ways of either reducing or preventing these shifts in order keep innocent people from being wrongfully convicted.

## REFERENCES

- Anderson, C.A., & Sechler, E.S. (1986). Effects of explanation and counterexplanation on the development and use of social theories. *Journal of Personality and Social Psychology*, 50 (1), 24-34.
- Ask, K., & Granhag, P. (2007). Motivational bias in criminal investigators' judgments of witness reliability. *Journal of Applied Social Psychology*, 37, 561-591.
- Ask, K., Rebelius, A., & Granhag, P.A. (2008). The 'elasticity' of criminal evidence: A moderator of investigator bias. *Applied Cognitive Psychology*, 22, 1245-1259. Doi:10.1002/acp.1432.
- Baumeister, R.F., & Bushman, B.J. (2008). *Social psychology and human nature*. Belmont, CA: Wadsworth.
- Carlsmith, J.M., & Aronson, E. (1963). Some hedonic consequences of the confirmation and disconfirmation of expectancies. *Journal of Abnormal and Social Psychology*, 66, 151-156.
- Charman, S.D. (2013). The forensic confirmation bias: A problem of evidence integration, not Just evidence evaluation. *Journal of Applied Research in Memory and Cognition*, 2, 56-58.
- Charman, S.D., Carbone, J.N., Kekessie, S, & Villalba, D. Evidence evaluation and evidence integration in legal decision-making: Order of evidence presentation as a moderator of context effects. Under review.
- Charman, S.D., Gregory, A.H., & Carlucci, M. (2009). Exploring the diagnostic utility of facial composites: Beliefs of guilt can bias perceived similarity between composite and Suspect. *Journal of Experimental Psychology: Applied*, 1-15. Doi:10.1037/a0014682.
- Costanzo, M., & Krauss, D. (2012). *Forensic and legal psychology: Psychological science applied to law*. New York, NY: Worth Publishers.
- Cutler, B.L., Dexter, H.R., & Penrod, S.D. (1990). Expert testimony and jury decision making: An empirical analysis. *Behavioral Sciences and the Law*, 7, 215-225.
- Dahl, L. C., Brimacombe, C. A., & Lindsay, D. S. (2009). Investigating investigators: How presentation order influences participant-investigators' interpretations of eyewitness identification and alibi evidence. *Law and Human Behavior*, 33(5), 368-80. doi:10.1007/s10979-008-9151-y

- Dijksterhuis, A., Bos, M.W., Nordgren, L.F., & Baaren, R.B. (2006). On making the right choice: The deliberation-without-attention effect. *Science*, 311, 1005-1007.
- Dijksterhuis, A. & Nordgren, L. (2006). A theory of unconscious thought. *Social Psychology Program, University of Amsterdam*, 1, 95-109.
- Dror, I.E., & Cole, S.A. (2010). The vision in “blind” justice: expert perception, judgment, and visual cognition in forensic pattern recognition. *Psychonomic Bulletin & Review*, 17(2), 161-167.
- Edwards, K., & Smith, E.E. (1996). A disconfirmation bias in the evaluation of arguments. *Journal of personality and Social Psychology*, 7, 5-24.
- Festinger, L., & Carlsmith, J.M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology*, 58, 203-210.
- Findley, K.A., & Scott, M.S. (2006). The multiple dimensions of tunnel vision in criminal cases. *Wisconsin law review*, 2, 1-108.
- Gilbert, D.T., & Malone, P.S. (1995). The correspondence bias. *Psychological Bulletin*, 117 (1), 21-38.
- Haney, L., & Lynch, M. (1994). Comprehending life and death matters. *Law and Human Behavior*, 18 (4), 411-436.
- Hasel, L.E., & Kassin, S.M. (2009). On the presumption of evidentiary independence: Can confessions corrupt eyewitness identifications? *Psychological Science*, 20 (2), 122-126.
- Hirt, E.R., & Markman, K.D. (1995). Multiple explanation: A consider-an-alternative strategy for debiasing judgments. *Journal of Personality and Social Psychology*, 69(6), 1069-1086.
- Holyoak, K.J., & Simon, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. *Journal of Experimental Psychology: General*, 128 (1), 3-31.
- Huber, O., & Seiser, G. (2001). Accounting and convincing: The effect of two types of justification on the decision process. *Journal of Behavioral Decision Making*, 14 (1), 69-85. Doi:10.1002/1099-0771(200101)14:1<69::AID-BDM366>3.0.C);2-T.
- Innocence Project (2014). Retrieved from Innocenceproject.org. August 15, 2014.
- Kalven, H., & Ziesel, H. (1966). The American jury. *Washington and Lee Law Review*, 158-175.

- Kassin, S.M., Bogart, D., & Kerner, J. (2011). Confessions that corrupt: Evidence from the DNA exoneration case files. *Psychological Science*, 1-5.  
Doi:10.1177/0956797611422918.
- Kassin, S. M., Dror, I. E., & Kukucka, J. (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition*, 2(1), 42-52. doi:10.1016/j.jarmac.2013.01.001
- Kassin, S. M., Goldstein, C. C., & Savitsky, K. (2003). Behavioral confirmation in the interrogation room: On the dangers of presuming guilt. *Law and Human Behavior*, 27(2), 187-203.
- Kassin, S.M., & Wrightsman, L.S. (1983). The construction and validation of a juror bias scale. *Journal of Research in Personality*, 17, 423-442.
- Lerner, J.S., & Tetlock, P.E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, 125(2), 255-275.
- Lord, C.G., Ross, L., & Lepper, M.R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 37, 2098-2109.
- Marksteiner, T., Ask, K., Reinhard, M.A., & Granhag, P.A. (2011). Asymmetrical skepticism towards criminal evidence: The role of goal- and belief-consistency. *Applied Cognitive Psychology*, 25, 541-547. Doi:10.1002/acp.1719.
- Nickerson, R.S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2, 175-220.
- Pennington, N., & Hastie, R. (1992). Explaining the evidence: Tests of the story model for juror decision making. *Journal of personality and Social Psychology*, 62 (2), 189-206.
- Pennington, J., & Schlenker, B.R. (1999). Accountability for consequential decisions: Justifying ethical judgments to audiences. *Personality and Social Psychology Bulletin*, 25 (9), 1067-1081. Doi:10.1177/01461672992512001.
- People v. Beaman, No. 4-05-0610 (4d App., Nov 3, 2006).
- Price, H. L., & Dahl, L. C. (2013). Order and strength matter for evaluation of alibi and eyewitness evidence. *Applied Cognitive Psychology*, 28, 143-150.  
doi:10.1002/acp.2983

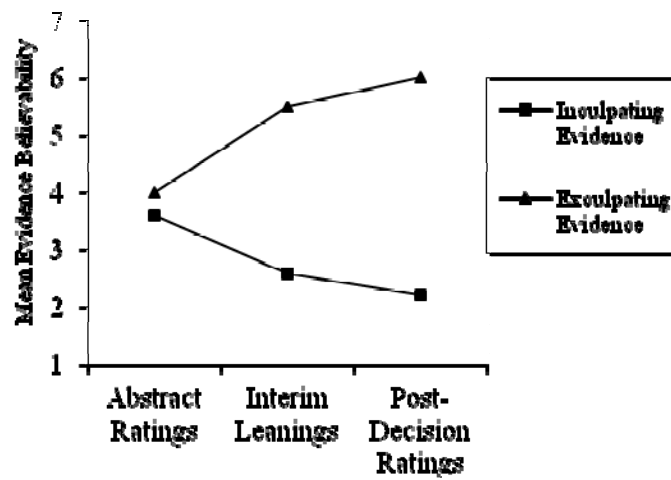


- Ruva, C.L., & McEvoy, C. (2008). Negative and positive pretrial publicity affect juror memory and decision making. *Journal of Experimental Psychology: Applied*, 3, 226-235.
- Scheb, J.M., & Scheb, J.M. (2011). *Criminal law & procedure*. Belmont, CA: Wadsworth.
- Schwartz, B.L. (2011). *Memory: Foundations and applications*. Thousand Oaks, CA: SAGE Publications, Inc.
- Simon, D. (2004). A third view of the black box: Cognitive coherence in legal decision making. *University of Chicago Law Review*, 71, 511-586.
- Simon, D. (2012). *In doubt: The psychology of the criminal justice process*. Cambridge, MA: Harvard University Press.
- Simon, D., Pham, L.B., Le, Q.A., & Holyoak, K.J. (2001). The emergence of coherence over the course of decision making. *Journal of Experimental Psychology*, 27 (5), 1250-1260.
- Swann, W.B. (1985). The self as architect of social reality. In B.R. Schlenker (Ed.), *the self and social life*, 100-125. New York: McGraw Hill.
- Wason, P. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 20, 273-281.
- Wilson, T.D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin*, 116 (1), 117-142.

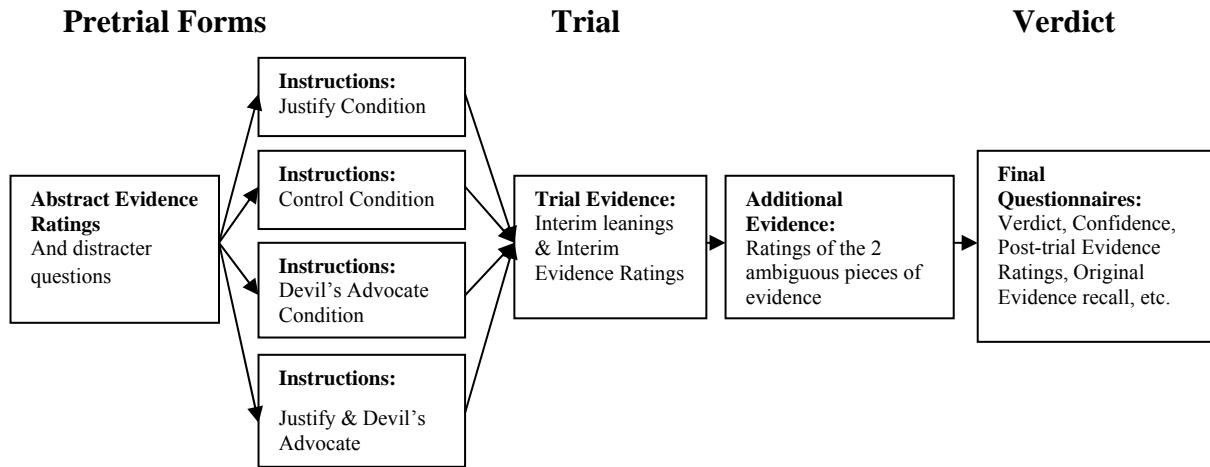
Table 1

Comparison of decision-making standpoints.

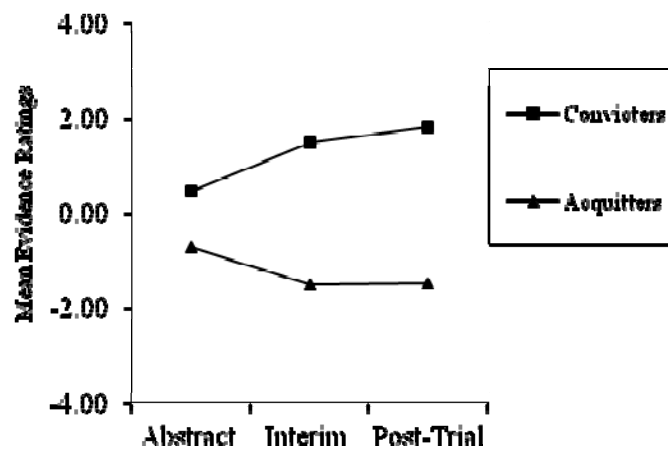
Standpoints	Flow of Logic	Examples
Rationalists	Rationale $\rightarrow$ Conclusion	“Incriminating forensic evidence means he’s guilty”
Critics	Conclusion $\rightarrow$ Post-hoc Rationale	“He’s guilty, due to forensics.” May be due to other factors like race, in reality
Coherence	Rationale $\leftrightarrow$ Conclusion	“Forensic evidence leads me to tentatively conclude he’s guilty; in fact, now that I think about it, his alibi seems flimsy too.”



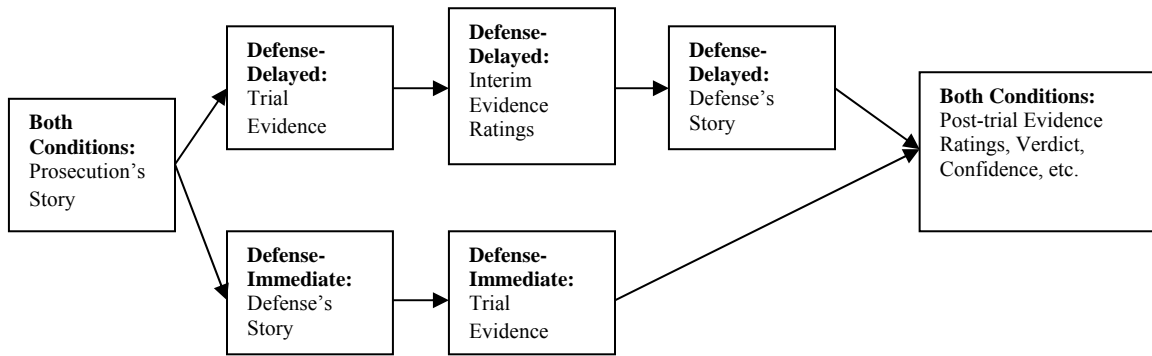
*Figure 1.* Acquitters' views of evidence over the course of the experiment. As the trial progresses, eventual acquitters perceive the exculpating evidence as progressively more believable, whereas they perceive the inculcating evidence is progressively less believable.



*Figure 2.* Procedural Flowchart for Study 1.



*Figure 3.* Jurors' perceptions of the average inculpatory and exculpatory value of evidence over the course of the decision. As the trial progresses, eventual acquitters perceive the evidence as progressively more exculpatory, whereas convicts perceive the evidence as progressively more inculpatory. Note: the graph ranges from -4 to 4 to make the effect apparent, but the scale ranged from -7 to 7.



*Figure 4.* Procedural Flowchart for Study 3.

## **APPENDICES**

### **Appendix A – Evidence Evaluation and Integration paper**

**Evidence evaluation and evidence integration in legal decision-making: Order of  
evidence presentation as a moderator of context effects**

## **Abstract**

Numerous empirical studies show that legal decision-making exhibits context effects: People's initial beliefs influence their evaluation of subsequent evidence. We examine three potential moderators of these context effects: Order of evidence presentation, ability to ruminate, and valence of the initial belief (innocence or guilt). College students ( $n = 382$ ) were presented with DNA evidence (incriminating or exonerating) and an ambiguous alibi in one of two orders (or just the alibi), and then evaluated the alibi's believability and the suspect's likelihood of guilt. Results indicated that alibi evaluation exhibited context effects when (a) initial beliefs were of guilt (but not of innocence), and when (b) evaluating subsequent evidence (but not when retroactively evaluating prior evidence). Rumination failed to moderate any effects. The integration of evidence exhibited recency effects: DNA had a greater impact on participants' beliefs in the suspect's guilt when presented last rather than first.



## **Evidence evaluation and evidence integration in legal decision-making: Order of evidence presentation as a moderator of context effects**

Decision-making within the context of a criminal case is complex. At almost every stage, various players in the legal system – detectives, lawyers, judges, juries – must evaluate disparate and often ambiguous pieces of evidence, integrate them into a coherent narrative, and reach a decision on the basis of that narrative. Unfortunately, as DNA exoneration cases have proven (Innocence Project, 2014), this decision-making is replete with error: Detectives pursue innocent suspects, lawyers pursue suboptimal defense strategies, judges render bad decisions, and juries convict innocent people. An understanding of the faulty decision-making process of people in the legal system is needed: How it is systematically biased, the conditions that promote or inhibit those biases, et cetera. To that end, the current study is aimed at developing this understanding by examining systematic biases and moderators of those biases.

Recent research has begun to uncover a particularly pernicious and ubiquitous bias: The role contextual information has in biasing subsequent legal decision-making. Participant-investigators' and police trainees' evaluations of the reliability of various types of evidence can be biased by their pre-existing beliefs about the suspect (Ask & Granhag, 2007; Ask, Rebelius, & Granhag, 2008). Eyewitnesses' lineup decisions are influenced by their knowledge that a lineup member confessed (Hasel & Kassin, 2009). Mock detectives and jurors perceive more similarity between a suspect and a facial composite of the criminal if they believe that suspect to be guilty (Charman, Gregory, & Carlucci, 2009). Interrogators' behaviors towards, and assessments of, suspects are biased by their initial beliefs of that suspect's guilt (Kassin, Goldstein, & Savitsky, 2003).

In fact, seemingly more objective judgments are also similarly influenced among actual forensic experts (Kassin, Bogart, & Kerner, 2012; Kassin, Dror, & Kuckuka, 2013). Fingerprint examiners' judgments of whether two fingerprints match are biased by their beliefs regarding the suspect (Dror, Charlton, & Peron, 2006; Dror & Cole, 2010; Dror & Rosenthal, 2008). Even the interpretation of DNA evidence – the 'gold standard' of evidence – can be influenced by forensic examiners' beliefs about the guilt of the suspect (Dror & Hampikian, 2011). For the purposes of this manuscript, we refer to any such findings – wherein a preexisting belief affects the subsequent interpretation of evidence – as context effects.

The relatively few psycholegal studies that have studied context effects have explained their findings using a variety of related theoretical approaches, including confirmation biases (Kassin et al., 2013), asymmetric skepticism (Ask & Granhag, 2007; Ask, Rebelius, & Granhag, 2008), and cognitive coherence frameworks (Holyoak & Simon, 1999; Simon, Pham, Le, & Holyoak, 2001). Although these theoretical approaches differ in their details – the scope of their focus, for instance – they are not mutually exclusive; in contrast, they all tend to converge to make a similar basic prediction: Evaluators' preexisting belief in the guilt or innocence of a suspect leads them to evaluate evidence in a manner consistent with their beliefs. The important question we address in the current study is not which of these approaches is correct, as each could be applied to the question of context effects. Rather, our focus concerns the boundary conditions of context effects: When will these effects be more or less likely to occur? What forensically-relevant variables moderate context effects?

## **Presentation order**

It is of course the case that actual criminal investigators and jurors discover and learn about various pieces of evidence gradually over time. Although we know empirically that a piece of evidence can influence evaluations of a subsequently-presented piece of evidence (e.g., Charman et al., 2009; Kassin et al., 2013), it is much less clear whether a piece of evidence can retroactively change the evaluation of an earlier-presented piece of evidence. For example, will an investigator who is presented with an alibi and then subsequently presented with results of a DNA test change his/her evaluation of the alibi? The primary goal of the current study is, therefore, to examine whether presentation order moderates context effects.

Charman (2013) points out two different processes that may be affected by context effects: *evidence evaluation* and *evidence integration*. This distinction is important, as it may be that the mechanism via which context effects occur differs between these two processes. As such, they are discussed separately.

**Evidence evaluation.** Evidence evaluation refers to the process by which a single piece of evidence is evaluated. The pertinent question is whether one's knowledge of one piece of evidence affects the evaluation of a different piece of evidence, and whether any such context effect depends on the order in which those two pieces of evidence are presented to evaluators.

On the one hand, we might expect context effects on evidence evaluation to be primarily unidirectional (i.e., the first piece of evidence will influence the second, but not vice versa), for two reasons. First, the evidence presented first (E1) will tend to lead to the formation of a preliminary conclusion, which then influences how the investigator

evaluates the evidence presented second (E2); this would be a standard context effect. But E2 may be less likely to retroactively influence the evaluation of E1, since E1 has already been evaluated and is thus ‘anchored’ in place. Second, Ask, Granhag, and Rebelius (2011) have shown that when evaluators have a goal to reach a conclusion quickly, they tend to reduce the depth of processing of subsequently presented evidence (and are hence more biased in their evaluation of it). Consider what this means for order effects. Evaluators who initially encounter strong incriminating or exonerating evidence may develop a strong belief concerning the guilt of the suspect, thus increasing their motivation to reach a quick conclusion, and reducing the depth with which the subsequent ambiguous evidence is evaluated, leading to a standard context effect. However, evaluators who initially encounter ambiguous evidence should tend to process it relatively deeply in the absence of a strong belief concerning the guilt of the suspect. The evaluator is only then presented with strong evidence. But since they have already processed the ambiguous evidence relatively deeply, any subsequent evaluations of that ambiguous evidence are less likely to be affected by strong evidence. Consequently, we would expect order effects: E1 should tend to influence the evaluation of E2 more than E2 retroactively influences the evaluation of E1.

On the other hand, some studies have shown that people re-assess the validity of a legal argument over the course of learning about a case (e.g., Holyoak & Simon, 1999). Similarly, it may be that after learning about some particularly strong piece of evidence, people will re-assess evidence evaluated earlier in light of this new knowledge (or that the new knowledge will distort their memories of the quality of the first evidence), thus resulting in a lack of order effects. Importantly, however, studies showing this type of re-

assessment have involved the assessments of the validity of arguments, which are relatively subjective evaluations. In contrast, evaluating the value of a piece of evidence may be perceived by evaluators as a more objective task. Because context effects are weaker to the extent that the information being evaluated is perceived as more objective (Ask, Rebelius, & Granhag, 2008), this re-evaluation of prior evidence may not take place.

Most empirical psycholegal research into context effects has tended to take one of two methodological approaches. The first approach is to instill a belief among participants that a suspect is either guilty or not, and then present them with a piece of evidence to evaluate; these studies tend to show that the initial belief affects the evaluation of the subsequent evidence (e.g., Ask & Granhag, 2007; Charman et al., 2009; Dror, Charlton, & Peron, 2006). A second approach is to present participants simultaneously (or almost simultaneously) with an array of arguments related to a legal case and then have participants rate the extent to which they agree with those arguments at various points in the decision-making process; these studies tend to show that the ratings of the various arguments tend to cohere more and more as participants form conclusions about the case (e.g., Holyoak & Simon, 1999; Simon et al., 2001; Simon, Krawczyk, & Holyoak 2004). Importantly, neither of these paradigms manipulate the order in which the evidence is presented, and consequently it is unknown how order of presentation moderates context effects.

However, two recent studies used a different paradigm from those mentioned above to manipulate evidence order, and thus deserve mention here. Dahl, Brimacombe, and Lindsay (2009) and Price and Dahl (2013) both presented participant-evaluators with

two pieces of evidence (e.g., an alibi and an eyewitness identification decision) in one of two randomly assigned orders, and then obtained a number of judgments, including the perceived credibility of the evidence (a measure of evidence evaluation). Results of both studies suggested that when two pieces of sequentially-presented evidence have strong, opposite valences, the last piece of evidence is perceived as more credible than it would have been in the absence of the other evidence.

There are, however, at least two questions left unanswered by this research. First, this methodology lacks a control group of participants who only receive a single piece of evidence. Although these studies show an order effect, without a control group it is unclear whether this order effect was the result of (a) the first piece of evidence boosting the perceived credibility of the second piece of evidence, or (b) the second piece of evidence suppressing the credibility of the first piece of evidence. In contrast, the methodology used in the current study uses a control group, and thus should be able to clarify this relationship.

Second, both Dahl et al. (2009) and Price and Dahl (2013) only used evidence of clearly opposite valences (one piece incriminating, the other exonerating). It is unclear whether a similar order effect would be found when one of the pieces of evidence is more ambiguous, not clearly pointing towards innocence or guilt. Some research has shown that context effects are greater for more ambiguous pieces of information (Ask et al., 2008), suggesting that the moderating impact of order on context effects may be even more pronounced when one piece of evidence is ambiguous. Accordingly, the current study uses participant evaluations of a weak alibi that neither strongly incriminates nor exonerates the suspect as our key outcome measure.

**Evidence integration.** In contrast to evidence evaluation, evidence integration involves combining multiple pieces of evidence to arrive at a global assessment about the overall guilt of a suspect (Charman, 2013). Detectives and prosecutors may need to combine fingerprint evidence with eyewitness evidence with alibi evidence, for instance, in order to determine whether a particular suspect should be prosecuted. Similarly, jurors must combine multiple pieces of evidence at trial in order to render a guilty/not guilty verdict.

How might order of evidence presentation affect evidence integration? It is crucial to note that, logically speaking, the order in which evidence is presented to an evaluator should not matter in terms of how it is combined: First learning about a suspect's weak alibi and then subsequently about incriminating DNA evidence should lead to the same conclusion as first learning about the incriminating DNA evidence and then subsequently the suspect's weak alibi. We might therefore expect that order should not moderate evidence integration.

However, there are data to suggest that order could matter. In both the Dahl et al. (2009) and the Price and Dahl (2013) studies that manipulated order of evidence presentation, participants assessed the likelihood of the suspect's guilt (a measure of evidence integration). Both studies found that, at least under some conditions, the *last* piece of evidence presented had a greater effect on the ultimate assessment of the suspect's guilt.

Price and Dahl (2013) provided two theoretical mechanisms by which this recency effect might occur. First, they proposed that it may be a result of a contrast effect (Scherer & Lambert, 2009): When the last piece of evidence is of a different valence

from the first piece of evidence, it may appear even stronger as a result of contrasting with the first piece. Second, they proposed a differential accessibility hypothesis: If evaluators make overall guilt assessments based on the accessibility of the various pieces of evidence they have evaluated, then the last piece of evidence presented, having been presented most recently in the past, should tend to be more accessible and hence have a greater impact.

Note however, that in both Dahl et al. (2009) and Price and Dahl (2013), their observed recency effects on evidence integration, like their observed recency effect on evidence evaluation, were only observed when the different pieces of evidence were of strongly opposite valences. Again, the current study includes one piece of evidence that is either incriminating or exonerating, and an ambiguous piece of evidence that is not clearly incriminating or exonerating. This is important, as it has yet to be shown whether similar recency effects exist when evaluating ambiguous evidence. In fact, Price and Dahl's contrast effect explanation for context effects *relies* on the two pieces of evidence being of different valences. Consequently, evidence of a recency effect in the current study would strongly suggest that the contrast effect explanation for recency effects is insufficient.

### **A counter-intuitive prediction**

It is important to note that the prediction that evidence integration will be primarily driven by the last piece of evidence is highly counterintuitive if evidence evaluation is primarily driven by the first piece of evidence (i.e., we see order effects). This is best highlighted with an example. Imagine that an evaluator is presented with two pieces of evidence – strongly incriminating DNA evidence and the suspect's weak alibi –



and that we can assign a numerical value representing the perceived incriminating value of each of these pieces of evidence – for instance, 5 for the strongly incriminating DNA and 1 for the weak alibi. If evaluators simply ‘add up’ the incriminating evidence, then an evaluator who is presented with the alibi followed by the incriminating DNA evidence will obtain a ‘guilt belief score’ of  $1 + 5 = 6$ .

However, an evaluator who is presented with the incriminating DNA evidence followed by the alibi will perceive the alibi as being more incriminating due to context effects (perhaps increasing its incriminating value to, say, 3). Consequently, adding up the incriminating evidence will result in a ‘guilt belief score’ of  $5 + 3 = 8$ .

Thus, under this simple model in which beliefs of guilt are obtained by algebraically adding up individual pieces of evidence, the DNA-first evaluator should exhibit greater beliefs in the suspect’s guilt than the DNA-last evaluator. But this is the exact opposite of what is actually predicted: If overall assessments of a suspect’s guilt are primarily driven by the last piece of evidence presented, then the “incriminating DNA last” evaluators should exhibit greater beliefs in the suspect’s guilt than the “incriminating DNA first” evaluators. This counterintuitive finding would suggest that evidence evaluation and evidence integration are driven by different processes, and that generating an overall assessment of a suspect’s guilt is not simply a matter of “adding up” the various pieces of evidence.

### **Other possible moderators of context effects**

In addition to our main focus of examining presentation order as a moderator of context effects, the current study will also examine two additional possible moderators, given their real-world applicability.

**Initial belief valence.** One such potential moderator is the valence of the initial belief – whether the evaluator initially believes the suspect to be innocent or guilty. Context effects may be thought of as resulting from the constraints imposed by initial information on subsequent decisions (e.g., Holyoak & Simon, 1999). To form a consistent narrative the various pieces of evidence must be coherent; consequently, once the first piece of evidence establishes a belief, the interpretation of subsequent evidence is effectively forced into a narrative that maintains coherence. Importantly, beliefs of guilt should tend to impose stronger constraints on the evaluation of subsequent evidence than beliefs of innocence. Take an alibi, for instance. If an evaluator believes the suspect to be guilty, then the evaluator must believe the suspect's alibi to be weak in order to maintain narrative coherence: Strong constraints are imposed on the evaluation of the alibi, and context effects should thus be observed. But if the evaluator believes the suspect to be innocent, there is not necessarily a similar constraint to judge the suspect's alibi as strong: An innocent person can have a weak alibi without disrupting narrative coherence. Consequently, context effects should be weaker. Indeed, some studies have observed exactly that pattern: Stronger context effects when the initial belief is guilt as opposed to innocence (e.g. Charman et al., 2009; Marksteiner, Ask, Reinhard, & Granhag, 2011).

**Rumination.** Given that multiple pieces of evidence are often obtained sequentially over the course of a criminal investigation, or presented sequentially to a jury over the course of a trial, an additional potential moderator of context effects concerns the opportunity to ruminate over the first piece of evidence. On the one hand, because people naturally tend to engage in confirmation biases, whereby they focus on hypothesis-consistent information and ignore or refute hypothesis-inconsistent

information (Nickerson, 1998), the opportunity to ruminate over the first piece of evidence may tend to further polarize and solidify the emerging conclusion, resulting in stronger context effects. On the other hand, there is some evidence that time delay between learning information and then providing a related judgment may result in weaker context effects (Simon, Krawczyk, Bleicher, & Holyoak, 2008); because rumination requires time, it is possible that context effects will dissipate. It should be noted, however, that participants in Simon et al. were not instructed to ruminate over the prior learned information; it is possible that this dissipation only occurs in the absence of cognitive rehearsal of that information. Finally, it is also possible that the rumination manipulation will fail to moderate context effects, as context effects may not be mediated at all by conscious, deliberative thought. For instance, Simon et al. (2001) showed that context effects occurred equally strongly regardless of participants' goals (e.g., to simply memorize the information vs. to process the information in preparation of receiving new information vs. to process the information in preparation of communicating it to someone else), and even when they were explicitly instructed to delay making a decision, suggesting that these effects may result spontaneously and unconsciously. If so, additional instructions to ruminate may fail to moderate these effects at all.

Thus, the purpose of this paper is to (1) examine context effects within a legal case; (2) examine how the *order* in which evidence is presented moderates these context effects, and whether it moderates evidence evaluation differently from evidence integration; and (3) examine the additional moderators of *opportunity to ruminate* and *belief valence* on context effects. It is expected that the first piece of evidence will lead participants to evaluate the second piece of evidence presented consistently (but not vice

versa), whereas the last piece of evidence presented (c.f. the first piece) will have a greater impact on the integration of evidence into an overall guilt assessment.

## **Method**

### **Design Overview**

Participants were told about a murder that ostensibly occurred recently, and listened to audiorecorded testimony (ostensibly from depositions) that consisted of a discussion of two types of evidence: DNA evidence and alibi evidence. (Control participants only heard about alibi evidence). Non-control participants were randomly assigned (a) to hear about this evidence in one of two orders, (b) to hear that the DNA evidence was either exonerating or incriminating, and (c) to have either an opportunity to ruminate or not after hearing about the first piece of evidence. All participants then evaluated the evidence (their evaluations of the alibi evidence was the primary dependent measure to assess evidence evaluation) and provided their overall beliefs in the guilt of the suspect (to assess evidence integration).

The design of this study was thus a 2 (*order*: DNA evidence first, DNA evidence second) X 2 (*valence* of DNA evidence: inculcating, exculpating) X 2 (*rumination opportunity*: present, absent) + 1 (control: alibi evidence only).

### **Participants**

Three hundred and eighty-two participants (110 males and 272 females; age 15 – 46, mean age = 21 years) were recruited from a research pool of undergraduate students at a large southeastern American university.

## **Materials**

**Case summary.** The fictitious case summary was loosely based on the Michael Blair trial, which involved the murder of a young girl in a park (see Appendix A). According to this case summary, over the course of a two-week investigation, police narrowed in on a suspect – Samuel Scott – who matched a general description given by a witness.

**Alibi testimony.** Because context effects are most likely when evidence is ‘elastic’ (Ask et al., 2008), the alibi was created to be deliberately ambiguous (see Appendix B). Specifically, the suspect claimed to be at his girlfriend’s house playing online poker. Although there were records of someone logging into an online poker site, it could not be verified that that person was the suspect. Furthermore, although the suspect’s girlfriend could state that he was present for part of the night, she claimed that she could not account his whereabouts over the course of the entire night.

**DNA testimony.** The DNA expert testimony was created in collaboration with an actual DNA expert who frequently testifies at trials. This DNA expert also served as the voice actor for the mock DNA expert. Two nearly identical versions of this testimony were created, the only difference between the two being that in one the expert testified that the DNA clearly implicated Mr. Scott, and one in which he testified that it clearly exonerated Mr. Scott (see Appendix C).

**Dependent measures.** Once participants finished learning about the evidence, they were asked to respond to a series of questions concerning their evaluations of the evidence. The primary measure used to examine evidence evaluation was participants’ mean response to the following question: “Consider the alibi evidence that you heard

(alibi evidence meaning: Mr. Scott's description of where he was at the time of the crime). How strongly does the alibi evidence imply the innocence or guilt of Mr. Scott?" Participants responded on a scale from 1 (strongly implies innocence) to 9 (strongly implies guilt). To examine evidence integration, participants were then asked to render a dichotomous verdict of guilty or not guilty and to indicate the overall 'likelihood of the defendant's guilt,' measured on a scale from 1 (not at all likely) to 9 (extremely likely). Participants also responded to questions concerning their confidence in their verdict on a scale from 1 (not at all confident) to 9 (extremely confident), and their perceptions of the reliability of alibi evidence and DNA evidence in general on a scale from 1 (not at all reliable) to 9 (extremely reliable).

## **Procedure**

Participants signed up for the study under the guise that it concerned whether their verdicts would match the verdicts of actual jurors in a controversial Colorado trial that had received substantial pre-trial publicity in Colorado. Upon arrival in the lab, participants were seated at a computer; materials were administered via the computer program Qualtrics.

Participants were given some brief background on the trial and were asked if they had previously heard about it in the media to further bolster the ruse that it was a real trial. They then read the case summary.

Participants were then randomly assigned to one of the nine conditions, and were presented with evidence from the case (DNA and alibi, or just alibi for the control participants). Participants were told that they would hear a recording from the actual depositions of the DNA expert and of the defendant giving his alibi (or just the defendant

giving his alibi for the control participants). They were also able to follow along with the dialogue by reading a transcript on the screen.

**Order manipulation.** Control participants received only alibi evidence. All other participants were randomly assigned to either hear DNA evidence first or alibi evidence first.

**Rumination manipulation.** Non-control participants were also randomly assigned to either a *no rumination* condition (in which the second piece of evidence was presented immediately following the first), or a rumination condition (in which the second piece of evidence was presented five minutes after the first). At the beginning of this five-minute period, rumination participants were told to “Take a few moments and consider the information you have been presented with thus far. You will automatically be taken to the next section after a few minutes.”

**DNA valence manipulation.** Non-control participants were randomly assigned to hear from the DNA expert that the DNA results were either incriminating or exonerating. The testimony was otherwise identical.

After hearing about the various pieces of evidence, participants responded to the questionnaire concerning evidence evaluation and evidence integration, and were then debriefed and excused.

## **Results**

### **Evidence evaluation**

Table 1 displays means and standard deviations for participants’ alibi evaluation as a function of condition. A 2 (order) X 2 (valence) X 2 (rumination) ANOVA revealed that rumination failed to significantly affect participants’ evaluations of the alibi

(rumination:  $M = 4.8$ ,  $SD = 1.8$ ; no rumination:  $M = 4.9$ ,  $SD = 1.9$ ),  $F(1, 341) = .01$ ,  $p = .91$ ,  $\eta_p^2 < .001$ , and also failed to significantly moderate any effects, all  $F$ s  $< 1.3$ , all  $p$ s  $> .25$ . Consequently, all subsequent analyses on evidence evaluation collapse across this factor. Figure 1 displays mean alibi believability ratings as a function of condition, collapsed across rumination.

The ANOVA revealed a significant main effect of order (DNA first:  $M = 5.1$ ,  $SD = 1.8$ ; DNA last:  $M = 4.6$ ,  $SD = 1.7$ ),  $F(1, 341) = 4.88$ ,  $p = .03$ ,  $\eta_p^2 = .014$ , and valence (incriminating DNA:  $M = 5.4$ ,  $SD = 1.9$ ; exonerating DNA:  $M = 4.3$ ,  $SD = 1.5$ ),  $F(1, 341) = 36.51$ ,  $p < .001$ ,  $\eta_p^2 = .097$ . However, these main effects were qualified by a significant order x valence interaction,  $F(1, 341) = 7.91$ ,  $p = .005$ ,  $\eta_p^2 = .023$ . Follow-up analyses revealed that when presented with incriminating DNA, participants evaluated the alibi as being more indicative of guilt if the DNA was presented first rather than last,  $t(180) = 3.35$ ,  $p = .001$ ,  $d = .50$ . However, when the DNA was exonerating, the order in which it was presented had no significant effect,  $t(165) = .56$ ,  $p = .57$ ,  $d = .09$ . Because this analysis does not include a control condition of participants who never received DNA evidence, these results are followed up with targeted pairwise comparisons involving control group participants.

**Incriminating DNA evidence.** Demonstrating general context effects, a priori contrasts revealed that participants who learned about incriminating DNA evidence believed that the suspect's alibi implied significantly more guilt than did control participants who did not learn about DNA evidence,  $t(213) = 2.38$ ,  $p = .02$ ,  $d = .50$ . However, these context effects were moderated by the order in which the evidence was presented: The evaluation of the alibi was only significantly affected by the incriminating



DNA evidence when the DNA evidence was presented *before* the alibi evidence,  $t(121) = 3.53, p = .001, d = .74$ , but not when the incriminating DNA evidence was presented *after* the alibi evidence,  $t(123) = 1.06, p = .29, d = .23$ .

**Exonerating DNA evidence.** No significant context effects emerged when the DNA evidence was exonerating: Participants' ratings of the alibi did not significantly differ between those who received exonerating DNA evidence and control participants, whether collapsed across order,  $t(198) = .87, p = .39, d = .12$ , or whether the DNA evidence was presented *before* the alibi evidence,  $t(106) = 1.04, p = .30, d = .19$ , or whether the DNA evidence was presented *after* the alibi evidence,  $t(123) = .59, p = .55, d = .06$ .

### **Evidence integration**

As expected, participants who learned about incriminating DNA evidence perceived a greater likelihood of the suspect's guilt than control participants who did not learn about DNA evidence,  $t(213) = 13.24, p < .001, d = 2.26$ . Participants who learned about exonerating DNA evidence perceived a lesser likelihood of the suspect's guilt compared to control participants,  $t(198) = 3.99, p < .001, d = .66$ . Similar results were found when looking at verdict as an outcome variable: Participants who learned about incriminating DNA were more likely to provide a guilty verdict than control participants (94.5% vs 24.2%),  $\chi^2(1) = 101.2, p < .001, OR = 3.9$ , and control participants were more likely to provide a guilty verdict than participants who heard about exonerating DNA evidence (4.8%),  $\chi^2(1) = 14.2, p < .001, OR = 5.0$ . For all subsequent analyses dealing with evidence integration control participants are not included, as they did not integrate multiple pieces of evidence into an overall assessment of guilt.

Table 2 displays means and standard deviations for participants' likelihood of guilt ratings as a function of condition. A 2 (order) X 2 (valence) X 2 (rumination) ANOVA revealed that rumination failed to significantly affect likelihood of guilt ratings (rumination:  $M = 5.3$ ,  $SD = 3.0$ ; no rumination:  $M = 5.4$ ,  $SD = 2.9$ ),  $F(1, 341) = .41$ ,  $p = .52$ ,  $\eta_p^2 = .001$ , and also failed to significantly moderate any effects, all  $F$ s  $< 1.1$ , all  $p$ s  $> .31$ . The ANOVA did however reveal a main effect of valence (incriminating:  $M = 7.8$ ,  $SD = 1.4$ ; exonerating:  $M = 2.7$ ,  $SD = 1.6$ ),  $F(1, 341) = 991.4$ ,  $p < .001$ ,  $\eta_p^2 = .744$ , which was qualified by a significant order X valence interaction,  $F(1, 341) = 17.08$ ,  $p < .001$ ,  $\eta_p^2 = .048$ . This interaction is displayed in Figure 2.

Follow-up analyses revealed that the effect of evidence order depended on whether the DNA evidence was incriminating or exonerating. When the DNA evidence incriminated the suspect, participants gave *higher* estimates of the suspect's guilt when it was presented last rather than first,  $t(180) = 2.73$ ,  $p = .007$ ,  $d = .41$ . However, when the DNA evidence exonerated the suspect, participants gave *lower* estimates of the suspect's guilt when it was presented last rather than first,  $t(165) = 3.18$ ,  $p = .002$ ,  $d = .50$ . In other words, regardless of whether the DNA evidence was incriminating or exonerating, the DNA evidence had a stronger impact on the perceived likelihood of the defendant's guilt when it was presented *last*.

### **Verdict confidence**

We have already shown that incriminating DNA evidence has a greater impact on subsequent alibi evaluation than exonerating DNA evidence. There is yet another way to evaluate the relative impact of incriminating vs. exonerating information: If incriminating information produces stronger context effects than exonerating information, we would

expect to see participants' confidence in their guilty verdicts (when they learned about incriminating DNA evidence) to be greater than their confidence in their not guilty verdicts (when they learned about exonerating DNA evidence). This analysis excluded 18 participants whose verdict did not match the valence of the DNA evidence. The hypothesis was supported: Participants who gave a guilty verdict after learning about incriminating DNA evidence were significantly more confident in their verdict ( $M = 7.7$ ,  $SD = 1.2$ ) than participants who gave a not guilty verdict after learning about exonerating DNA evidence ( $M = 6.8$ ,  $SD = 1.9$ ),  $t(329) = 5.62$ ,  $p < .001$ ,  $d = .57$ .

To provide an alternative metric of whether incriminating DNA had a stronger impact than exonerating DNA, a continuous confidence measure was obtained by combining verdict and confidence, resulting in scores that ranged from -8.5 (extreme confidence in innocence) to 8.5 (extreme confidence in guilt). Consistent with the above analysis, incriminating DNA ( $M = 6.6$ ,  $SD = 2.9$ ) shifted confidence (c.f. to control:  $M = -2.4$ ,  $SD = 4.8$ ) more than exonerating DNA did ( $M = -5.7$ ,  $SD = 3.1$ ),  $z = 6.71$ ,  $p < .001$ .

### **General perceptions of evidence**

To examine whether contextual information affected perceptions not only of the specific evidence presented, but also of general perceptions of that class of evidence (c.f., Ask et al., 2008), we also examined the effect of incriminating vs exonerating DNA on participants' perceptions of the reliability of alibi evidence in general and DNA evidence in general. Participants who received incriminating DNA evidence believed that DNA evidence was significantly more reliable ( $M = 8.3$ ,  $SD = .9$ ) than participants who received exonerating DNA evidence ( $M = 8.0$ ,  $SD = 1.5$ ),  $t(347) = 2.32$ ,  $p = .03$ ,  $d = .24$ . However, neither of these differed significantly from control participants' ratings of the

reliability of DNA evidence ( $M = 8.2$ ,  $SD = 1.1$ ); incriminating DNA vs control:  $t(213) = .66$ ,  $p = .51$ ,  $d = .07$ ; exonerating DNA vs control:  $t(198) = .65$ ,  $p = .52$ ,  $d = .07$ .

Participants who received incriminating DNA evidence did not believe that alibi evidence was significantly more reliable ( $M = 4.6$ ,  $SD = 1.8$ ) than participants who received exonerating DNA evidence ( $M = 4.7$ ,  $SD = 1.8$ ),  $t(347) = .49$ ,  $p = .63$ ,  $d = .06$ , and neither of these differed significantly from control participants' ratings of the reliability of alibi evidence ( $M = 4.3$ ,  $SD = 1.8$ ); incriminating vs control:  $t(213) = .88$ ,  $p = .38$ ,  $d = .12$ ; exonerating vs control:  $t(198) = 1.16$ ,  $p = .24$ ,  $d = .17$

## Discussion

Results indicate that context can affect decision-making in two ways: By changing the evaluation of a piece of evidence, and by changing the way different pieces of evidences are integrated together, and that these effects are different.

### Evidence evaluation

Replicating past research (see Kassin et al., 2013), we found a significant context effect on evidence evaluation: Participants' knowledge of a piece of evidence (in this case, DNA) significantly affected their evaluation of a subsequent piece of evidence (i.e., an alibi). Importantly, however, this general context effect was moderated by two variables: Order of presentation of the evidence and the valence of the evidence. Specifically, context effects only emerged when (a) the valenced evidence preceded the ambiguous information, and (b) the valenced evidence was incriminating in nature.

**Moderating effect of order.** Only two previous studies have systematically manipulated the order in which evidence was presented to evaluators to examine whether order of evidence moderates context effects (Dahl et al., 2009; Price & Dahl, 2013). Both

studies found that when the two types of evidence were of strongly opposite valences (one incriminating and one exonerating), the perceived credibility of the evidence depended on the order in which the evidence was presented, with the evidence evaluated being seen as more credible if it was presented last rather than first.

However, in both studies, the lack of a control group that only receives a single piece of information hampered the interpretation of their findings, as it could not be ascertained whether the first piece of evidence influenced evaluation of the last piece of evidence, or whether the last piece of evidence influenced evaluation of the first piece of evidence. Results of the current study help to clarify this relationship: The first piece of evidence influences subsequent evaluations of evidence, but the last piece of evidence does not retroactively influence evaluations of the first piece of evidence.

The current study also expands upon this previous work on order effects in two additional ways. First, both Dahl et al. (2009) and Price and Dahl (2013) demonstrated context effects on evaluations of the credibility of the evidence presented (and in fact, most similar studies look at evidence credibility or reliability: e.g., Ask & Granhag, 2007; Ask et al., 2008). In the current study we demonstrated context effects when participants evaluated the diagnostic value of the evidence: Participants rated the alibi as more indicative of guilt when it followed incriminating DNA than they did otherwise. Thus, we expanded the scope of context effects to include not just the credibility of the person providing the evidence, but also the perceived value of the evidence itself.

Second, both Dahl et al. (2009) and Price and Dahl (2013) showed order effects specifically when the two pieces of information were contradictory to one another – one piece was clearly incriminating and the other clearly exonerating. In fact, one of the

theoretical explanations that Price and Dahl give for their findings is based on contrast effects (Scherer & Lambert, 2009), which predicts order effects *only* when the two pieces of evidence are contradictory. Current results demonstrate similar order effects when the evaluated piece of evidence is not clearly incriminating or exonerating, but rather ambiguous. Thus, although a contrast explanation may be partly responsible for some specific order effects, it fails to provide an explanation for the current findings.

Instead, our interpretation of these order effects is derived from Ask et al. (2011), who showed that activating an efficiency goal (as opposed to a thoroughness goal) led to more biased evaluations of evidence. We can imagine that in the absence of a strong belief of guilt, incoming evidence is evaluated fairly thoroughly. However, if a strong belief in the guilt of a suspect is invoked via particularly damning evidence, this should tend to activate an efficiency goal, resulting in relatively superficial processing of subsequent information and as a consequence, context effects. Note that order effects of the type observed in the current study naturally arise from this account: When strong incriminating evidence is presented first, subsequent evidence is processed superficially. But if the ambiguous evidence is presented first, in the absence of a belief of guilt or innocence, it is processed at a relatively deep level. Thus, even when the evaluator later encounters a strong piece of incriminating evidence, she has already evaluated the ambiguous evidence relatively deeply, and its evaluation is not retrospectively affected by the incriminating information. The evaluator simply accesses the original evaluation of the first ambiguous piece of information, thus precluding context effects.

**Moderating effect of valence.** The majority of studies have also been unable to examine the potential moderating effect of valence of evidence on context effects because

they have either not manipulated the valence of the initial evidence (e.g., Dror et al., 2006; Miller, 1984), or did not include a no-initial-evidence control group (e.g., Ask & Granhag, 2007; Ask et al., 2011; Ask, Rebelius, & Granhag, 2008; Dahl et al., 2009; Price & Dahl, 2013)<sup>1</sup>. Studies that have been able to compare valences have shown a similar pattern to that observed here: Incriminating information has a larger effect on subsequent evidence evaluation than exonerating information (Charman et al., 2009; Marksteiner et al., 2011). This finding suggests that innocent suspects are at a particular disadvantage when they fall under the scrutiny of investigators: Context effects will tend to work to their disadvantage (by making other evidence appear even more incriminating) but never to their advantage (by making other evidence appear even more exonerating).

There are at least two explanations for the finding that incriminating evidence produces stronger context effects than exonerating information. First, context effects may arise when one piece of evidence imposes constraints on the interpretation of another piece of evidence, thus leading it to be evaluated in a manner consistent with the first piece of evidence (e.g., Holyoak & Simon, 1999). In the current study, incriminating evidence would be expected to produce stronger constraints on the interpretation of subsequent ambiguous evidence – and thus produce stronger context effects – than would exonerating evidence. In other words, if a suspect is guilty, then his alibi *must* be inaccurate by definition; the alibi cannot be evaluated any other way within the context of

---

<sup>1</sup> Two additional studies manipulated valence and included a no information control group, but were severely underpowered. Elaad, Ginton, & Ben-Shakar (1994) only included three evaluators in their control group. Dror and Charlton (2006) only provided evidence for three instances in which a previous decision was changed to a different, belief-consistent decision; even these instances may not be an example of context effects, as there were two instances of a previous decision being changed to a different decision in the no information, *control* condition. Conclusions concerning context effects – and the moderating influence of valence in particular – are thus difficult to draw from these studies.

that belief. But if a suspect is innocent, his alibi still could be weak; a weak alibi does not necessarily contradict one's belief that the suspect is innocent. Consequently, we would expect – and we observe – that incriminating rather than exonerating initial evidence should produce stronger context effects on the evaluation of subsequent ambiguous evidence due to the stronger constraints it imposes.

Second, this pattern may arise due to evaluators' cognitive orientation: Evaluators in a criminal case may have a natural tendency to look for evidence indicative of guilt, as opposed to evidence indicative of innocence. If so, then despite our attempts to keep the diagnostic value of incriminating evidence equal (but of opposite valence) to the diagnostic value of exonerating evidence, natural confirmation biases will tend to lead participants to be more inclined to accept at face value incriminating DNA evidence as opposed to exonerating DNA evidence. In turn, the stronger beliefs of guilt among incriminating DNA participants (compared to beliefs of innocence among exonerating DNA participants) should tend to lead to greater context effects when the initial evidence is incriminating vs exonerating.

**Lack of a moderating effect of rumination.** Notably, opportunity to ruminate after the presentation of the first piece of evidence failed to moderate any context effects: Evaluations of the final piece of evidence were equally affected by the initial evidence regardless of whether participants had the opportunity to ruminate on the first piece of evidence or not. This null result may have occurred because the presentation of a single piece of evidence spontaneously and automatically imposes cognitive constraints on subsequent decision-making, unmediated by explicit cognitive rehearsal. This is consistent with Simon et al. (2001), who demonstrated spontaneous context effects



regardless of the explicit goals on the part of participants, and even when they were instructed not to form conclusions. Unfortunately, this suggests that context effects may be particularly pernicious and virtually unavoidable. Nonetheless, one should interpret these null effects with caution as the opportunity for rumination lasted only five minutes. This specific length was chosen semi-arbitrarily based on studies that have shown context effects to last at least 3 minutes (Holyoak & Simon, 1999, Experiment 1), but to wear off by 10-15 minutes (Simon et al., 2008). Nonetheless, it is possible that longer opportunities to ruminate, such as those that may arise over the course of a month-long trial, may in fact moderate context effects, and future studies may wish to examine this possibility.

### **Evidence integration**

One might expect that in order to arrive at an overall assessment of a suspect's likely guilt, an evaluator would 'add up' the evaluations of all of the individual pieces of evidence. But had evaluators done so, we would have expected one of two findings. First, we may have expected to observe a lack of an effect of evidence order; after all, because the evidence is objectively identical regardless of the order in which it was presented, evaluators who add up the individual evaluations of evidence should reach similar conclusions regardless of the order in which it was presented.

Second, we may have expected the first piece of evidence to be more important in determining the final assessment than the second piece of evidence. This is because participants in the incriminating DNA-first condition had their subsequent evaluations of the alibi shifted towards guilt. A simple algebraic model in which participants sum the evaluations of each piece of evidence would predict therefore that incriminating DNA-

first participants should exhibit a greater belief in the suspect's guilt than incriminating DNA-last participants (whose evaluations of the alibi were not affected by the subsequent DNA evidence).

But in fact, we observed neither of these patterns; instead, we found that the last piece of evidence had a greater impact on participants' overall assessment of guilt, indicating that participants are not simply adding up the individual evaluations of evidence. Rather, evidence integration seems to be driven primarily by a recency effect. This is consistent with findings of Dahl et al. (2009) and Price and Dahl (2013), who showed a similar recency effect on participants' assessments of the guilt of a suspect.

This finding expands this prior research, however, in two ways. First, unlike Dahl et al. (2009) and Price and Dahl (2013) who only found a recency effect when the pieces of evidence were strongly contradictory, we found the effect using an ambiguous alibi that implied neither guilt nor innocence, thus expanding the scope of the effect. Second, our results may help distinguish between the two hypotheses – a contrast effect and a differential accessibility effect – put forth by Price and Dahl (2013) to explain recency effects. Specifically, because our key piece of evidence was ambiguous in nature, our results are unlikely to have resulted from a contrast effect<sup>2</sup>. Our results are, however, perfectly consistent with a differential accessibility account: When making guilt assessments, participants determined their current feelings of guilt; if the (relatively strong) DNA evidence was presented last, it was weighed disproportionately in that assessment due to its relatively stronger accessibility in memory. But when the (relatively weak) alibi evidence was presented last, it was weighed disproportionately in the overall

---

<sup>2</sup> That being said, it is possible that an alibi may be perceived as intrinsically exonerating, regardless of its actual content; as such we cannot completely rule out a contrast effect.

assessment of guilt due to its relatively stronger accessibility in memory, diluting the effect of the DNA evidence.

Our overall interpretation for our results is the following: When evaluating individual pieces of evidence, participants process the evidence at a relatively deep level unless they have a preexisting belief of the suspect's guilt that constrains subsequent evaluation, in which case they process subsequent evidence relatively superficially. Because superficial processing tends to result in context effects (Ask et al., 2011), these effects tend to only emerge when the key evidence follows, and not precedes, evidence that establishes belief in the suspect's guilt. And because beliefs of guilt produce stronger constraints in subsequent evidence evaluation than beliefs of innocence, we tend to only observe context effects when the preexisting belief is one of guilt. Evidence integration, however, is driven by a different process. Specifically, to make overall assessments of guilt, participants judge their current feelings<sup>3</sup>; because their current feelings are disproportionately affected by the most recent piece of evidence presented (as a result of differential accessibility in memory), recency effects emerge. Although more research is needed to establish the validity of this interpretation, we note that it is consistent with all of the current findings, as well as the findings of past research that have shown biased evaluations of evidence (e.g., Ask & Granhag, 2007; Charman et al., 2009; Kassin et al., 2013) and recency effects with respect to evidence integration (e.g., Dahl et al., 2009; Price and Dahl, 2013).

## **Implications**

In addition to helping us develop a theoretical understanding of how people

---

<sup>3</sup> For instance, see Schwartz and Clore's (2007) feelings-as-information hypothesis, which more fully elaborates upon this idea

evaluate and integrate evidence, these results also have potential implications for multiple players within the legal system. Our demonstration of context effects (in conjunction with the many other studies that have shown similar effects) suggests that evaluators of evidence should, whenever possible, be blind to the existence of other evidence when making their evaluations. Although this can sometimes be accomplished (e.g., with some forensic experts; Kassin et al., 2013), this ideal cannot always be met; it is often impossible, for instance, for detectives, attorneys, or jurors to evaluate a single piece of evidence in a contextual vacuum. At the very least, a knowledge of the conditions that promote or inhibit context effects is important. When determining whether police officers and detectives may be driven by context effects, for example, it may be important to determine not just the evidence that led them to develop their beliefs, but also the order in which they uncovered that evidence. And attorneys may wish to understand how context effects are moderated by order of presentation when preparing their arguments in court; if jurors' beliefs of a defendant's guilt are affected by context, but only in certain specific situations, then this information will help guide their presentation of evidence at trial.

These results also have implications for our understanding of juror decision-making, as they suggest that jurors may be influenced by factors beyond the probative value of the evidence (e.g., the *order* in which the evidence is presented). Logically, such factors should be irrelevant and unimportant to the mental calculus of jurors. However, context effects may lead jurors to generate overly polarized or extreme opinions about the defendant's guilt (e.g., Holyoak & Simon, 1999). Thus, critical safeguards, such as the reasonable doubt standard, are at risk of being ineffective in light of context effects and their potential to inflate beliefs of guilt beyond the value of the evidence.

Importantly, our results indicate that evidence evaluation and evidence integration are separate processes (see Charman, 2013). Most other research has looked at one or the other in isolation; the current study is one of the few that looked at both processes together, and thus is one of the few to be able to show the disconnect between the two processes. This is important, as psycholegal researchers have used the umbrella term ‘context effects’ (or similar terms) to refer interchangeably to these two processes, and as a consequence exhibit findings that seem to contradict one another. For example, Price and Dahl (2013) claim that their finding of a recency effect contradicts predictions of confirmation biases, which predict a primacy effect. But we believe this apparent contradiction to be illusory: Primacy effects are shown with respect to evidence evaluation; recency effects (like those of Price and Dahl) are shown with respect to evidence integration. Consequently, we encourage future researchers examining context effects to be clear regarding which of these two processes are being examined.

## References

- Ask, K., & Granhag, P. (2007). Motivational bias in criminal investigators' judgments of witness reliability. *Journal of Applied Social Psychology*, 37, 561-591.
- Ask, K., Granhag, P. A., & Rebelius, A. (2011). Investigators under influence: How social norms activate goal-directed processing of criminal evidence. *Applied Cognitive Psychology*, 25(4), 548-553. doi:10.1002/acp.1724
- Ask, K., Rebelius, A., & Granhag, P. A. (2008). The 'elasticity' of criminal evidence: A moderator of investigator bias. *Applied Cognitive Psychology*, 22(9), 1245-1259. doi:10.1002/acp.1432
- Charman, S. D. (2013). The forensic confirmation bias: A problem of evidence integration, not just evidence evaluation. *Journal of Applied Research in Memory and Cognition*, 2, 56-58. doi: 10.1016/j.jarmac.2013.01.010
- Charman, S. D., Gregory, A. H., & Carlucci, M. (2009). Exploring the diagnostic utility of facial composites: Beliefs of guilt can bias perceived similarity between composite and suspect. *Journal of Experimental Psychology: Applied*, 15, 76-90. doi:10.1037/a0014682
- Dahl, L. C., Brimacombe, C. A., & Lindsay, D. S. (2009). Investigating investigators: How presentation order influences participant-investigators' interpretations of eyewitness identification and alibi evidence. *Law and Human Behavior*, 33(5), 368-80. doi:10.1007/s10979-008-9151-y
- Dror, I. E., Charlton, D., & Péron, A. E. (2006). Contextual information renders experts vulnerable to making erroneous identifications. *Forensic Science International*, 156(1), 74-8. doi:10.1016/j.forsciint.2005.10.017
- Dror, I. E., & Cole, S. A. (2010). The vision in "blind" justice: Expert perception, judgment, and visual cognition in forensic pattern recognition. *Psychonomic Bulletin & Review*, 17(2), 161-7. doi:10.3758/PBR.17.2.161
- Dror, I. E., & Hampikian, G. (2011). Subjectivity and bias in forensic DNA mixture interpretation. *Science & Justice : Journal of the Forensic Science Society*, 51(4), 204-8. doi:10.1016/j.scijus.2011.08.004
- Dror, I., & Rosenthal, R. (2008). Meta-analytically quantifying the reliability and biasability of forensic experts. *Journal of Forensic Sciences*, 53(4), 900-3. doi:10.1111/j.1556-4029.2008.00762.x

- Elaad, E., Ginton, A., & Ben-Shakhar, G. (1994). The effects of prior expectations and outcome knowledge on polygraph examiners' decisions. *Journal of Behavioral Decision Making*, 7(4), 279-292.
- Hasel, L. E., & Kassin, S. M. (2009). On the presumption of evidentiary independence: Can confessions corrupt eyewitness identifications? *Psychological Science*, 20(1), 122-6. doi:10.1111/j.1467-9280.2008.02262.x
- Holyoak, K. J., & Simon, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. *Journal of Experimental Psychology: General*, 128(1), 3-31.
- Innocence Project (2014). Retrieved from Innocenceproject.org August 15, 2014.
- Kassin, S. M., Bogart, D., & Kerner, J. (2012). Confessions that corrupt: Evidence from the DNA exoneration case files. *Psychological Science*, 23, 41-45.
- Kassin, S. M., Dror, I. E., & Kukucka, J. (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition*, 2(1), 42-52. doi:10.1016/j.jarmac.2013.01.001
- Kassin, S. M., Goldstein, C. C., & Savitsky, K. (2003). Behavioral confirmation in the interrogation room: On the dangers of presuming guilt. *Law and Human Behavior*, 27(2), 187-203.
- Marksteiner, T., Ask, K., Reinhard, M. A., & Granhag, P. A. (2011). Asymmetrical scepticism towards criminal evidence: The role of goal-and belief-consistency. *Applied Cognitive Psychology*, 25(4), 541-547. doi:10.1002/acp.1719
- Miller, L. S. (1984). Bias among forensic document examiners: A need for procedural changes. *Journal of Police Science and Administration*, 12(4), 407-411.
- Nickerson, R.S. (1998). Confirmation bias: a ubiquitous phenomenon in many guises. *Review of general psychology*, 2, 175-220.
- Price, H. L., & Dahl, L. C. (2013). Order and strength matter for evaluation of alibi and eyewitness evidence. *Applied Cognitive Psychology*, 28, 143-150. doi:10.1002/acp.2983
- Scherer, L. D., & Lambert, A. J. (2009). Contrast effects in priming paradigms: Implications for theory research and implicit attitudes. *Journal of Personality and Social Psychology*, 97, 383-403. doi:10.1037/a0015844
- Schwarz, N., & Clore, G. L. (2007). Feelings and phenomenal experiences. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social Psychology: Handbook of basic principles* (2nd ed., pp. 385-407). New York: Guilford.

Simon, D., Krawczyk, D. C., & Holyoak, K. J. (2004). Construction of preferences by constraint satisfaction. *Psychological Science*, 15(5), 331-336.

Simon, D., Krawczyk, D. C., Bleicher, A., & Holyoak, K. J. (2008). The transience of constructed preferences. *Journal of Behavioral Decision Making*, 21(1), 1-14.  
doi:10.1002/bdm.575

Simon, D., Pham, L. B., Le, Q. A., & Holyoak, K. J. (2001). The emergence of coherence over the course of decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(5), 1250-1260.



Table 1. Means and standard deviations of alibi evaluation ratings as a function of DNA valence, evidence order, and rumination

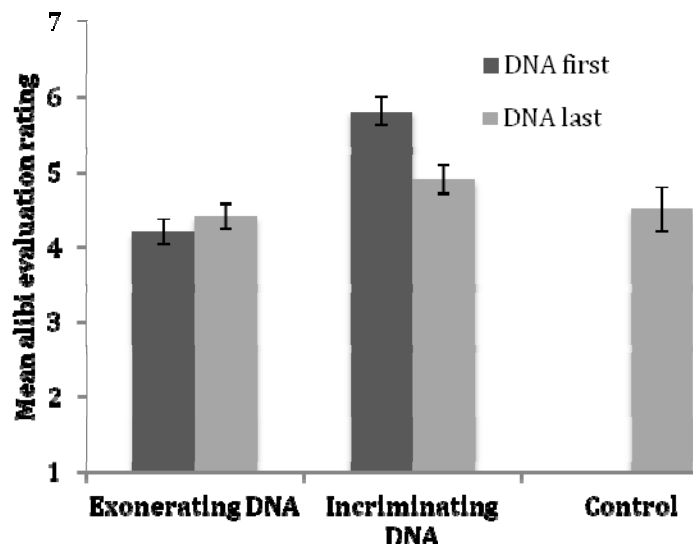
	Incriminating DNA	No DNA	Exonerating DNA
	Rumination		
DNA first	6.0 (1.8)		4.0 (1.4)
DNA last	4.9 (1.8)		4.4 (1.5)
	No rumination		
DNA first	5.7 (1.9)		4.4 (1.4)
DNA last	5.0 (1.9)		4.3 (1.8)
	Collapsed across rumination condition		
DNA first	5.8 (1.8)		4.2 (1.4)
DNA last	4.9 (1.8)		4.3 (1.6)
Total	5.4 (1.9)	4.5 (1.7)	4.3 (1.5)

*Note.* Higher values represent greater beliefs that the alibi implies guilt.

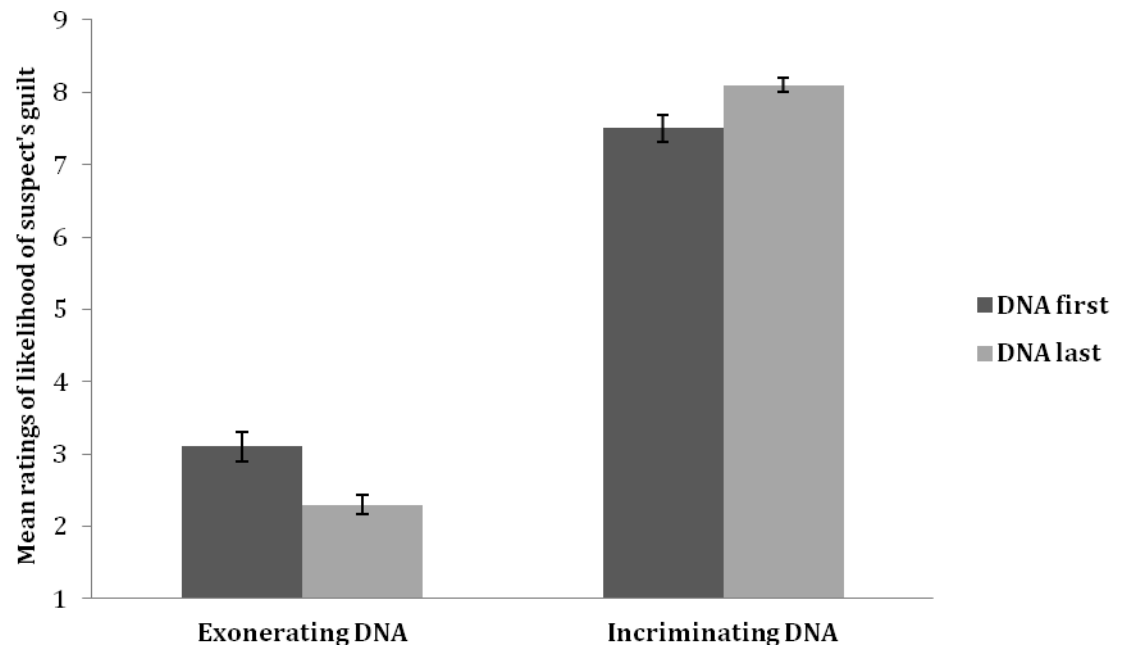
Table 2. Means and standard deviations of belief of suspect's guilt scores as a function of DNA valence, evidence order, and rumination

	Incriminating DNA	No DNA	Exonerating DNA
	Rumination		
DNA first	7.8 (1.6)		3.0 (1.9)
DNA last	8.1 (1.0)		2.3 (1.1)
	No rumination		
DNA first	7.3 (1.9)		3.1 (1.8)
DNA last	8.0 (.9)		2.3 (1.5)
	Collapsed across rumination condition		
DNA first	7.5 (1.8)		3.1 (1.8)
DNA last	8.1 (1.0)		2.3 (1.3)
Total	7.8 (1.4)	3.9 (2.0)	2.7 (1.6)

*Note.* Higher values represent greater beliefs in suspect's guilt.



*Figure 1.* Evidence evaluation: Mean alibi evaluation ratings (error bars represent the standard error of the mean). Higher values represent greater belief that the alibi implies guilt.



*Figure 2.* Evidence integration: Mean ratings of likelihood of suspect's guilt (error bars represent the standard error of the mean)

## **Appendix A: Case summary**

### **The People versus Samuel Scott**

Case Docket: 03801-97844

On June 19<sup>th</sup> 2010, a young girl named Elizabeth Schafman was attending her brother's soccer game at a park in Colorado in the early evening. During a break in the game, she told her family that she was going to 'walk around for a bit.' When she did not return after about 10 minutes, her family grew concerned and her father began to search for her. After the game was over and she hadn't returned or been found, the Schafmans called the police.

The police performed an initial search of the nearby woods and were also unsuccessful at finding her. However, the police returned to search a second time later in the evening and they performed a canine-assisted search of the woods. The dogs led them to a bush in which Elizabeth's body had been thrown. She was declared dead at the scene. It was clear that the cause of death was strangulation and there appeared to have been an intense struggle.

After they had found the body, Detective Harris was placed in charge of the investigation and he began interviewing people who were at the soccer game. Several people who were at the game claimed to have seen a suspicious-looking person wandering around the periphery of the park. The police were able to put together a description of this person, even though the description was quite general. Based on this information, police patrolled the surrounding area over the course of a few days and found several people who matched the general description. The investigation continued to narrow for about 2 weeks, until police settled on a suspect, Samuel Scott.

## Appendix B: Suspect's alibi

**Interviewer:** "Let the record show, this deposition is beginning at 10 a.m. on July 6th, 2010. My name is name is Alisha Benick and I will be interviewing the suspect on behalf of the court. The objective today is to hear the suspect's alibi. Please state your name."

**Mr. Scott:** "Samuel T. Scott."

**Interviewer:** "Mr. Scott, are you aware of the charges being brought against you?"

**Mr. Scott:** "Yes."

**Interviewer:** "Do you deny those allegations?"

**Mr. Scott:** "I do deny them, yes."

**Interviewer:** "Can you tell us where you were and what you were doing on the evening of June 19<sup>th</sup>?"

**Mr. Scott:** "I was at my girlfriend's house. I usually go there to spend time with her after I'm done with work in the afternoon. Ya' know, most days I arrive there around 4 o'clock."

**Interviewer:** "Approximately how far away is your girlfriend's house from the park where the incident took place?"

**Mr. Scott:** "It's about...um...a 5 minute drive."

**Interviewer:** "And do you recall doing anything specific while at your girlfriend's house on the day in question?"

**Mr. Scott:** "Yeah, I mean, I was actually playing online Poker for quite a while that day. While I played, in the living room, my girlfriend mostly is in her bedroom, watching TV."

**Interviewer:** "According to the investigation, the detectives were able to verify that someone was logged in to the online Poker site that day, but not necessarily that it was you. Do you have any way to prove that it was you?"

**Mr. Scott:** "Well, they were able to show *someone* was logged in. But I can't prove that it was me...I guess. Ya' know, I don't know that much about computers."

**Interviewer:** "When your girlfriend was interviewed by the investigators, um, she stated that the two of you had dinner that evening at about 8 o'clock. Is that correct?"

**Mr. Scott:** "That's right, yup."

**Interviewer:** "Additionally, your girlfriend agreed that you been playing online Poker that night. However, she also stated that she thought you had been 'in-and-out' of the house. How do you account for that?"

**Mr. Scott:** “Uh, I, I don’t remember leaving the house, I remember being there the whole time.”

**Interviewer:** “So you believe that your girlfriend may not be remembering that day accurately?”

**Mr. Scott:** “Well, the issue is that it’s just a typical day. It’s hard for anyone to remember what someone else was doing on a certain day, several weeks ago. She might have simply been thinking of a different day.”

**Interviewer:** “Have you ever met or encountered the victim in this case, Elizabeth Schaftman?”

**Mr. Scott:** “No, I have not.”

**Interviewer:** “Well that’s all the questions we have. Thank you for your time, Mr. Scott.”

## **Appendix C: DNA testimony**

### **Italicized phrases were manipulated across the guilty DNA/not guilty DNA conditions**

**Interviewer:** “Let the record show, this deposition is beginning at 11 a.m. on July 8<sup>th</sup>, 2010. My name is Alisha Benick and I will be interviewing the DNA expert on behalf of the court. The objective today is to hear the results of the analyses he conducted. Please state your name.”

**Dr. Taylor:** “My name is Doctor David Taylor.”

**Interviewer:** “And, for the record Doctor, where did you receive your degree and how long have you been working in this field?”

**Dr. Taylor:** “I received my Doctorate in Molecular Biology from Harvard in 1990. I have been working in the field since then, so around 21 years.”

**Interviewer:** “You were asked by the court to analyze and review the DNA evidence in this case. Can you tell us a little bit about the evidence you received?”

**Dr. Taylor:** “That’s correct. My lab received a sample of the perpetrator’s DNA from the crime scene. The incident involved the strangulation of a young girl, Elizabeth Schafman, and there had clearly been a struggle. During this struggle, the victim had dug her fingernails into the flesh of her attacker, presumably in an attempt to get him to let go of her neck. There were remnants of both blood and skin found underneath the victim’s fingernails. That evidence was sent to us by the investigators and we analyzed it.”

**Interviewer:** “Can you tell us, in general, a bit about how DNA testing works?”

**Dr. Taylor:** “Certainly. Every person has a unique DNA signature, commonly referred to as a DNA fingerprint. Now, most of our important DNA sequences are identical among all people; we share 99% percent of our DNA. However, there are certain sites on the human genome that vary widely from one person to another. We can use these sections to distinguish one person from another. So, we compare the sample obtained at the crime scene to the sample we obtained from the suspect, and see if there’s a match.”

**Interviewer:** “How does that process work? How do you find the specific areas of a DNA sequence that makes-up a person’s unique DNA fingerprint?”



**Dr. Taylor:** “Initially, we use what’s called a ‘restriction enzyme’ to cut apart the various sections of the DNA code. Then we employ a technique called “Gel Electrophoresis.” This lets us pull out the certain sections of DNA that we want to examine. It’s basically like using an index to look up something very specific in a massive encyclopedia.”

**Interviewer:** “Well thank you for explaining that, doctor. Now, the critical question is, what did you find when you compared the perpetrator’s DNA to the DNA of the suspect, Mr. Samuel Scott?”

**Dr. Taylor:** “When we compared the two samples, we found that they *were a match/were not a match*. It was clear that the DNA fingerprints from Samuel Scott and the DNA from the fingernail scrapings of Elizabeth Schafman were *identical/quite different*.”

**Interviewer:** “In your opinion, then, did the DNA profile gathered from under Ms. Schafman’s fingernails come from Samuel Scott?”

Dr. Taylor: “Yes, I’m 99.9% sure it must have come from *Mr. Scott/someone else*.”

**Interviewer:** “Now how accurate are the techniques you use to analyze DNA evidence? What is the error rate?”

**Dr. Taylor:** “DNA evidence is widely considered the gold standard in evidence. In fact, DNA testing for legal issues has been going on since the mid 1980’s, so the techniques have been extensively developed. The techniques used are 99.9% accurate.”

**Interviewer:** “So overall, you are highly confident in your conclusions?”

**Dr. Taylor:** “Yes, I am. The story this DNA evidence tells is quite clear.”

**Interviewer:** “Thank you for your time, Dr. Taylor.”

## Appendix B - Recruitment Blurb

### Preparing for Trial:

Come help \_\_\_\_'s (named of university removed) grads prepare for an upcoming trial! The Legal Psychology faculty here at \_\_\_\_ have been asked by local lawyers, who are recent graduates of the \_\_\_\_ law school, to help gather opinions about an important upcoming murder trial. We are looking for students to read through a controversial case and provide us with your feedback and verdict. By providing the lawyers with your feedback, we hope to help them better understand their case. They have worked with us to create a case summary based on the actual evidence that will be presented at trial. The case facts and evidence will be presented to you to help you come to a decision about the defendant's guilt or lack thereof.

## Appendix C – General Legal Opinions

### **Instructions:**

Please provide your opinions about various aspects of the legal system. You will read a detailed description about each item of interest and then you will be asked to rate the extent to which you find that item compelling. For example, you may be asked to rate if you agree or disagree with a legal policy, such as allowing jurors to ask questions. Alternatively, you may be asked if a given piece of evidence, such as polygraph evidence, implies that a suspect is guilty or implies that a suspect is not guilty. Each item that you rate is unrelated and should be considered independently. You may find that you would like more information regarding a topic that we have asked you about, but you must do the best you can with the information provided. You are not expected to have any special expertise in terms of evaluating the evidence. Simply provide your opinion based on your sense of reasonableness, fairness, and good judgment.

### Ballistic Evidence (distracter question)

The shell casings from bullets were found at the scene of a murder. Police investigators gathered the casings and sent them to a forensics lab to be analyzed. The forensic technician compared shell casings to bullets that were found in the garage of the suspect's home. Despite the damage to the bullet casings as a result of being fired, the forensic technician claimed that the casings found at the crime scene did match the type of bullets found in the suspect's garage. The forensic technician noted that the type of bullets found at both the crime scene and in the defendant's garage are a relatively common type of bullet.

- To what extent does the ballistics evidence that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty							Implies Guilty							
-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

#### Defendant's Alibi (ambiguous evidence)

When a detective interviewed the suspect and asked where the suspect was on the evening in question, the suspect told the detective that he went to see a movie with a friend. When the detective followed-up with the movie theater, he found that the theater was indeed able to verify that the suspect's credit card had been used to purchase a ticket that evening. Additionally, the detective interviewed the suspect's friend, who stated that he had met up with the suspect at the movie theater. However, the theater for the particular movie they went to see was almost full and so the suspect and his friend had to sit in different sections of the theater. The friend could not say for sure whether the suspect had been present throughout the entire movie, but the suspect was there when the movie ended and the friend believes the suspect had been in the theater the whole time. The friend did not remember seeing anyone get up and leave the theater. The detective noted that the movie's duration was 2 hours, 50 minutes, and it was located

approximately 10 minutes from the scene of the crime. In terms of time stamps, the detective also noted that the victim's murder did take place during the time in which the suspect was allegedly present at the movie theater.

- To what extent does the alibi that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty							Implies Guilty							
-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

#### Phone Records (ambiguous evidence)

A detective was able to review the call-logs from the cell phones of a victim and a suspect. The detective found that the suspect and the victim did have some correspondence on the day of the incident and throughout the course of that week. The phone logs indicate that the suspect and the victim had an approximately fifteen minute conversation on the day of the murder, a few hours before the murder occurred. When asked about these conversations, the suspect states that the conversation was related to new a project that was being developed at work. The suspect said that it was necessary to ask the victim some questions about the part of the project that the victim was working on, so that the suspect could proceed with the part of the project that the suspect was

working on. The victim's roommate was able to confirm what the phone records had indicated, that a conversation did take place on the day of the murder. The roommate stated that in general, the victim had a good relationship with coworkers and had not mentioned any trouble at work. However, on the day in question, the roommate had overheard some of the phone call between the suspect and victim, and the roommate stated that it was an "intense conversation." The roommate also remembered hearing some references to work-related topics and that there was mention of who would be getting credit for a new project.

- To what extent do the phone records that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty										Implies Guilty						
-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7		
strongly		moderately		weakly		neutral		weakly		moderately		strongly				

#### Death Penalty Policy (distracter question)

The death penalty is a controversial matter of legal policy. Proponents of the death penalty claim that it is a fitting consequence for certain crimes, that it serves as a

deterrent, and that it prevents perpetrators from having the opportunity to reoffend in prison or otherwise. Opponents of the death penalty claim that it is far more financially costly than life imprisonment, that it permanently denies the defendant the opportunity for exonerating evidence to surface, and that it represents an inconsistency whereby society is sanctioning the same act that the perpetrator is guilty of committing.

- To what extent do you support or oppose the death penalty? Circle one number.

Oppose

Support

---

-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

---

#### Juror Question-Asking Policy (distracter question)

The ability for jurors to ask questions during trials is a controversial topic. Proponents of juror question-asking claim that it helps the jurors to have the information they need to arrive at a sound conclusion and that asking questions is critical to understanding any complex situation. Opponents of juror question-asking claim that jurors may become concerned with legally inadmissible topics or that they put undue weight on the response to their questions.

- To what extent do you support or oppose the policy of juror question-asking? Circle one number.

Oppose

Support

---

-7   -6   -5   -4   -3   -2   -1   0   1   2   3   4   5   6   7

strongly   moderately   weakly   neutral   weakly   moderately   strongly

---

Snitch Information (distracter question)

During the process of a murder investigation, a ‘snitch’ claimed to have information about the suspect in the case. The snitch told the police investigators that while he shared a cell with the suspect, the suspect bragged about killing the victim. The snitch knew that the murder weapon was a gun and that the victim was shot in the torso. The snitch is asking for a reduce sentence in light of being helpful. However, most murders involve guns and most lethal wounds are to the torso, which casts doubt on the snitch’s information and statements.

To what extent does the snitch testimony that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7   -6   -5   -4   -3   -2   -1   0   1   2   3   4   5   6   7

strongly   moderately   weakly   neutral   weakly   moderately   strongly

---



### Polygraph Evidence (exculpating)

A suspect in a crime was asked to take a polygraph and he complied. The polygraph machine was hooked up to the suspect to measure his heart rate, breathing, and skin conductance (sweat). The polygraph examiner asked the suspect several questions, some relating to the crime and some unrelated to the crime, for example, a question about the suspect's breakfast. The examiner reported that, in general, the suspect showed elevated levels of physiological arousal throughout the entire exam. The examiner stated that the suspect did not show more arousal for questions related to the crime than for questions unrelated to the crime. It was the examiner's conclusion that suspect was being truthful.

- To what extent does the polygraph evidence that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty							Implies Guilty							
-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

### Polygraph Evidence (inculcating)

A suspect in a crime was asked to take a polygraph and he complied. The polygraph machine was hooked up to the suspect to measure his heart rate, breathing, and skin conductance (sweat). The polygraph examiner asked the suspect several questions, some relating to the crime and some unrelated to the crime, for example, a question about the suspect's breakfast. The examiner reported that, in general, the suspect showed elevated levels of physiological arousal throughout the entire exam. The examiner stated that the suspect showed more arousal for questions related to the crime than for questions unrelated to the crime. It was the examiner's conclusion that suspect was being untruthful about something.

- To what extent does the polygraph evidence that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

---

### Eyewitness Identification (exculpating)

Police interviewed a male witness who was at a park at during the time frame in which a victim was likely murdered. The witness claimed to have seen a suspicious man wandering around the park. The witness was a regular jogger at this park and stated that he had never seen this man before. The police administered a line-up and asked the witness to identify the person that he had seen on the evening in question. The witness

made a selection without much hesitation. The witness did not identify the suspect. The witness selected someone else from the lineup and stated he was pretty confident in his decision.

- To what extent does the eyewitness identification that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty							Implies Guilty							
-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

#### Eyewitness Identification (inculpating)

Police interviewed a male witness who was at a park at during the time frame in which a victim was likely murdered. The witness claimed to have seen a suspicious man wandering around the park. The witness was a regular jogger at this park and stated that he had never seen this man before. The police administered a line-up and asked the witness to identify the person that he had seen on the evening in question. The witness made a selection without much hesitation. The witness identified the suspect. The witness selected the suspect from the lineup and stated he was pretty confident in his decision.

- To what extent does the eyewitness identification that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

---

#### Shoe Print Evidence (exculpating)

Police investigators were able to obtain shoe prints from the dirt surrounding the bushes in which a victim's body has been thrown. The shoe prints from the crime scene were compared to the suspect's shoes. A forensic evidence analyst from the police department found that the shoe prints from the scene of the crime were not similar to the shoes of the suspect. Although some of the shoe print had been smudged, likely by other people walking through the area, the analyst stated that from what he could tell, the shoe prints were not a match. It was the analyst's conclusion that the shoe prints found at the scene of the crime did not come from the suspect.

- To what extent does the shoe print evidence that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

---

### Shoe Print Evidence (inculcating)

Police investigators were able to obtain shoe prints from the dirt surrounding the bushes in which a victim's body has been thrown. The shoe prints from the crime scene were compared to the suspect's shoes. A forensic evidence analyst from the police department found that the shoe prints from the scene of the crime were similar to the shoes of the suspect. Although some of the shoe print had been smudged, likely by other people walking through the area, the analyst stated that from what he could tell, the shoe prints were a match. It was the analyst's conclusion that the shoe prints found at the scene of the crime did come from the suspect.

- To what extent does the shoe print evidence that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty								Implies Guilty						
-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

### Traffic Camera Footage (exculpating)

Footage from a street camera near where a murder took place was reviewed to see if any of the cars that passed through that intersection were similar to the suspect's car. Police investigators reviewed the footage during the approximate hours in which the

murder took place. The investigators found that none of the cars to drive through that intersection were white sedans. No cars from the video footage were similar to the suspect's white Ford Taurus. The police also stated that none of the cars from the footage appeared to be moving faster than one would expect, given the speed limit in that area.

- To what extent does the traffic camera footage evidence you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty							Implies Guilty							
-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

#### Traffic Camera Footage (inculpating)

Footage from a street camera near where a murder took place was reviewed to see if any of the cars that passed through that intersection were similar to the suspect's car. Police investigators reviewed the footage during the approximate hours in which the murder took place. The investigators found that one of the cars to drive through that intersection was a white sedan. The white sedan from the video footage was similar to the suspect's white Ford Taurus. The police also stated that the white sedan from the

footage appeared to be moving faster than one would expect, given the speed limit in that area.

- To what extent does the traffic camera footage evidence you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty							Implies Guilty							
-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

#### Lack of Motive (exculpating)

Coworkers from the same department in which a suspect and his victim worked were interviewed by the police and asked to describe how the suspect and victim got along. The coworkers stated that, overall, the two got along fine. There were apparently no instances in which coworkers saw the two arguing. The head of the firm stated that he was not aware of any friction between his two employees.

- To what extent does motive that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

---

Motive (inculcating)

Coworkers from the same department in which a suspect and his victim had worked were interviewed by the police and asked to describe how suspect and the victim got along. The coworkers stated that, overall, the two did not get along. There were apparently several instances in which coworkers saw the two arguing. The head of the department stated that he was aware of some friction between his two employees.

- To what extent does motive that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

---



### Fingerprint Evidence (exculpating)

A police detective was in charge of gathering evidence from the crime scene of a murder. He found that the victim had been stabbed several times and that the knife was left by her body at the scene of the crime. The detective removed the knife and sent it the lab for fingerprint analysis. The suspect was asked to come to the police station and provide his fingerprints as well. Many of the fingerprints on the knife had been rubbed off and so there were no complete prints on the knife. However, the forensic technician did compare the partial prints that were present and found that they were not a match. It was the tentative conclusion of the analyst was that the suspect's prints did not match those of the partial prints found on the knife.

- To what extent does the fingerprint evidence that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

---

### Fingerprint Evidence (inculpating)

A police detective was in charge of gathering evidence from the crime scene of a murder. He found that the victim had been stabbed several times and that the knife was

left by her body at the scene of the crime. The detective removed the knife and sent it the lab for fingerprint analysis. The suspect was asked to come to the police station and provide his fingerprints as well. Many of the fingerprints on the knife had been rubbed off and so there were no complete prints on the knife. However, the forensic technician did compare the partial prints that were present and found them to be a match. It was the tentative conclusion of the analyst was that suspect's prints did match those of the partial prints found on the knife.

- To what extent does the fingerprint evidence that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty							Implies Guilty							
-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

## Appendix D. – Instructions

### **Instructions**

You will be reading about a Florida Case titled The People v. David Mitchell, case number 03801-97844, which is going to trial shortly in St. Lucie county, Fort Pierce. There are some basic facts of the case that both sides have stipulated, and, as such, you will be reading through a brief case summary of those facts. Following the case summary, the prosecution and the defense will present the evidence that supports their side of the case.

However, when this case goes to trial, jurors will not hear all of the arguments and evidence at once, as trials can often take several days. At the end of each day, judges instruct jurors to withhold making decisions until they have heard all of the evidence. As such, we have split the trial into two segments, Part 1 and Part 2. After you read through Part 1, you will be asked some questions. However, please withhold making a final decision for the case until you have read through Part 2 as well.

Read through all of the information thoroughly and try to imagine that you are a real juror in this trial. Remember that the goal of this activity is to help two graduates of the \_\_\_\_ (name of university removed) law school, Javier Allegria and Michelle Gomez, prepare for an upcoming trial. They will essentially be testing their arguments and the anticipated arguments of the opposing attorney on you, to get your feedback.

You will start the trial by reading through a summary of the crime and the investigation, which both sides of the case have agreed upon.

**Participants in the ‘Justify’ conditions also read the following:**

At the conclusion of the trial you will be asked to justify your decision to the lawyers (in writing) and to the researcher (in person) who is currently running your study.

Be prepared to discuss your eventual decision and to cite the aspects of the case that support your view. When thinking about justifying your decision, it may help you to imagine that you will be trying to convince someone who sees the case differently than you do. This is important because the goal of this activity is to help the lawyers understand how jurors will perceive the case and come to a decision

**Participants in the ‘Devil’s Advocate’ conditions also read the following:**

As you read about the various pieces of evidence presented throughout the trial, you will be asked to take some time to consider any issues with those pieces of evidence. Meaning, even if evidence is presented by one side of the case, either the prosecution or the defense, consider how that piece of evidence would be perceived from the other side of the case. If you were the opposing attorney, what problems, flaws, or issues would YOU find with each piece of evidence?

**Participants in the ‘Justify and Devil’s Advocate’ conditions read both of the preceding paragraphs in addition to the ‘Control’ instructions.**

## Appendix E. – Case Summary

### **The People versus David Mitchell**

Case Docket: 03801-97844

On July 15<sup>th</sup> 2012, a young woman named Justine Alexander was going for a walk in a local park near her apartment in Fort Pierce, in the county of St. Lucy. It was a tradition of hers to go to this park most evenings after work. However, one night her roommates grew concerned when she did not come home by 1 a.m. on a weeknight and was not responding to text messages. The roommates called the police and gave them potential whereabouts for Justine.

The police performed an initial search of the nearby park but were unsuccessful at finding her. However, the police returned the next day to search a second time and they performed a canine-assisted search of the park and its surroundings. The dogs led them to an outcropping of bushes and trees, in which Justine's body had been thrown. She was declared dead at the scene. It was clear that the cause of death was several stab wounds and that there appeared to have been a struggle.

After they had found the body, Detective Thompson was placed in charge of the investigation and he began interviewing people who had been at the park that evening. A few people who were there claimed to have seen a suspicious-looking person wandering around the periphery of the park. The police were able to put together a description of this person, even though the description was quite general. Additionally, the police interviewed Justine's friends and coworkers at the marketing firm that she worked at. Based on this information, police came to believe that David Mitchell, a fellow coworker, was the most likely suspect.

Now that you have a general understanding of this case, you will be presented with the specific evidence. The information you have read so far is a description of what happened during the crime of the police investigation, but it is not to be considered evidence.

## Appendix F. – Case Evidence (Phase 1)

Now that you have read through the case summary, the lawyers in *The People v. David Mitchell* will present you with some of the evidence in this case. Please withhold arriving at a final decision until you have heard of the evidence in this case and until we ask you for a final verdict.

### **Evidence from the Prosecution's Case:**

#### Polygraph Evidence

David Mitchell was asked to take a polygraph and he complied. The polygraph machine was hooked up to Mitchell to measure his heart rate, breathing, and skin conductance (sweat). The polygraph examiner asked him several questions, some relating to the crime and some unrelated to the crime, for example, a question about the Mitchell's breakfast. The examiner reported that, in general, Mitchell showed elevated levels of physiological arousal throughout the entire exam. The examiner stated that he showed more arousal for questions related to the crime than for questions unrelated to the crime. It was the examiner's conclusion that Mitchell was being untruthful about something.

#### Eyewitness Evidence

Police interviewed a witness, Patrick McDougal, who was at the park at during the time frame in which Justine Alexander was likely murdered. The witness claimed to have seen a suspicious man wandering around the park. The witness was a regular jogger at this park and stated that he had never seen this man before. The police administered a line-up and asked the witness to identify the person that he had seen on the evening in question. The witness made a selection without much hesitation. The witness identified

David Mitchell. The witness selected Mitchell from the lineup and stated he was pretty confident in his decision.

#### Shoe Print Analysis

Police investigators were able to obtain shoe prints from the dirt surrounding the bushes in which Justine Alexander's body has been thrown. The shoe prints from the crime scene were compared to the suspect's shoes. A forensic evidence analyst from the police department found that the shoe prints from the scene of the crime were similar to the shoes of the David Mitchell. Although some of the shoe print had been smudged, likely by other people walking through the area, the analyst stated that from what he could tell, the shoe prints were a match. It was the analyst's conclusion that the shoe prints found at the scene of the crime did come from David Mitchell.

#### Traffic Camera Footage

Footage from a street camera near where Justine Alexander's murder took place was reviewed to see if any of the cars that passed through that intersection were similar to the David Mitchell's car. Police investigators reviewed the footage during the approximate hours in which the murder took place. The investigators found that one of the cars to drive through that intersection was a white sedan. The white sedan from the video footage was similar to the David Mitchell's white Ford Taurus. The police also stated that the white sedan from the footage appeared to be moving faster than one would expect, given the speed limit in that area.



## Motive

Coworkers from the same marketing firm in which David Mitchell and Justine Alexander worked were interviewed by the police and asked to describe how David and Justine got along. The coworkers stated that, overall, the two did not get along. There were apparently several instances in which coworkers saw David and Justine arguing. The head of the firm, Dr. James Muligan, stated that he was aware of some friction between his two employees.

## Fingerprint Analysis

Detective Thompson was in charge of gathering evidence from the crime scene of Justine Alexander's murder. He found that Justine had been stabbed several times and that the knife was left by her body at the scene of the crime. Detective Thompson removed the knife and sent it the lab for fingerprint analysis. David Mitchell was asked to come to the police station and provide his fingerprints as well. Many of the fingerprints on the knife had been rubbed off and so there were no complete prints on the knife. However, the forensic technician did compare the partial prints that were present and found them to be a match. It was the tentative conclusion of the analyst was that David Mitchell's prints did match those of the partial prints found on the knife.

## **Evidence from the Defense's Case:**

### Polygraph Evidence

David Mitchell was asked to take a polygraph and he complied. The polygraph machine was hooked up to Mitchell to measure his heart rate, breathing, and skin conductance (sweat). The polygraph examiner asked him several questions, some

relating to the crime and some unrelated to the crime, for example, a question about the Mitchell's breakfast. The examiner reported that, in general, Mitchell showed elevated levels of physiological arousal throughout the entire exam. The examiner stated, though, that he did not show more arousal for questions related to the crime than for questions unrelated to the crime. It was the examiner's conclusion that Mitchell was being truthful.

#### Eyewitness Evidence

Police interviewed a witness, Patrick McDougal, who was at the park at during the time frame in which Justine Alexander was likely murdered. The witness claimed to have seen a suspicious man wandering around the park. The witness was a regular jogger at this park and stated that he had never seen this man before. The police administered a line-up and asked the witness to identify the person that he had seen on the evening in question. The witness made a selection without much hesitation. The witness did not identify David Mitchell. The witness selected someone else from the lineup and stated he was pretty confident in his decision.

#### Shoe Print Analysis

Police investigators were able to obtain shoe prints from the dirt surrounding the bushes in which Justine Alexander's body has been thrown. The shoe prints from the crime scene were compared to the suspect's shoes. A forensic evidence analyst from the police department found that the shoe prints from the scene of the crime were not similar to the shoes of the David Mitchell. Although some of the shoe print had been smudged, likely by other people walking through the area, the analyst stated that from what he could tell, the shoe prints were not a match. It was the analyst's conclusion that the shoe prints found at the scene of the crime did not come from David Mitchell.

### Traffic Camera Footage

Footage from a street camera near where Justine Alexander's murder took place was reviewed to see if any of the cars that passed through that intersection were similar to the David Mitchell's car. Police investigators reviewed the footage during the approximate hours in which the murder took place. The investigators found that none of the cars to drive through that intersection were white sedans. No cars from the video footage were similar to the David Mitchell's white Ford Taurus. The police also stated that none of the cars from the footage appeared to be moving faster than one would expect, given the speed limit in that area.

### Lack of Motive

Coworkers from the same marketing firm in which David Mitchell and Justine Alexander worked were interviewed by the police and asked to describe how David and Justine got along. The coworkers stated that, overall, the two got along fine. There were apparently no instances in which coworkers saw David and Justine two arguing. The head of the firm, Dr. James Muligan, stated that he was not aware of any friction between his two employees.

### Fingerprint Analysis

Detective Thompson was in charge of gathering evidence from the crime scene of Justine Alexander's murder. He found that Justine had been stabbed several times and that the knife was left by her body at the scene of the crime. Detective Thompson removed the knife and sent it the lab for fingerprint analysis. David Mitchell was asked

to come to the police station and provide his fingerprints as well. Many of the fingerprints on the knife had been rubbed off and so there were no complete prints on the knife. However, the forensic technician did compare the partial prints that were present and found that they were not a match. It was the tentative conclusion of the analyst was that David Mitchell's prints did not match those of the partial prints found on the knife.

## Appendix G. – Interim Instructions & Ratings

### Interim Instructions

Shortly, you will begin the second phase of the trial and you will read about some additional evidence. However, before you move on, we would like you to take a moment and indicate your initial opinions about this case. When providing your initial opinions, please refrain from coming to any final decisions at this point in the trial because there is more evidence that will be presented shortly. Remember that you will be asked to justify your decision at the end of the trial, both to the lawyers and the researchers.

- At this point in the trial, are you more inclined to believe that David Mitchell is guilty or not guilty? Circle one:

**Guilty**      **Not Guilty**

- How confident are you that your current beliefs about the defendant's guilt or lack of guilt are accurate?

Not Confident

Confident

---

-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
Very	moderately		weakly		neutral		weakly		moderately		Very			

---

- At this point in the trial, how likely is it that David Mitchell is Guilty? Circle one:

Very Unlikely    Somewhat Unlikely    Neutral    Somewhat Likely    Very Likely

### Initial Evidence Evaluations

Take some time to consider each piece of evidence that you read about. Rate the extent to which each piece of evidence implies the suspect is not guilty or implies the suspect is guilty.

- To what extent does the polygraph evidence that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty	Implies Guilty												
<div style="display: flex; justify-content: space-between; margin: 0;"> <span>-7</span><span>-6</span><span>-5</span><span>-4</span><span>-3</span><span>-2</span><span>-1</span><span>0</span><span>1</span><span>2</span><span>3</span><span>4</span><span>5</span><span>6</span><span>7</span> </div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>strongly</span><span>moderately</span><span>weakly</span><span>neutral</span><span>weakly</span><span>moderately</span><span>strongly</span> </div>													

- To what extent does the motive that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7   -6   -5   -4   -3   -2   -1   0   1   2   3   4   5   6   7  
strongly   moderately   weakly   neutral   weakly   moderately   strongly

---

- To what extent does the shoe print analysis that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7   -6   -5   -4   -3   -2   -1   0   1   2   3   4   5   6   7  
strongly   moderately   weakly   neutral   weakly   moderately   strongly

---

- To what extent does the fingerprint analysis that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7   -6   -5   -4   -3   -2   -1   0   1   2   3   4   5   6   7  
strongly   moderately   weakly   neutral   weakly   moderately   strongly

---

- To what extent does the traffic camera footage you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7   -6   -5   -4   -3   -2   -1   0   1   2   3   4   5   6   7

strongly   moderately   weakly   neutral   weakly   moderately   strongly

---

- To what extent does the eyewitness identification that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7   -6   -5   -4   -3   -2   -1   0   1   2   3   4   5   6   7

strongly   moderately   weakly   neutral   weakly   moderately   strongly

---



## Appendix H. – Case Evidence (Phase 2)

### **Case Evidence (Part 2)**

You heard some of the evidence in this case already, but two more pieces of evidence will be presented to you. These two pieces of evidence have been stipulated, or agreed upon, by both sides of the case.

#### Defendant's Alibi

When Detective Thompson interviewed David Mitchell and asked him where he was on the evening in question, Mitchell told the detective that he went to see a movie with a friend. When Detective Thompson followed-up with the movie theater, he found that the theater was indeed able to verify that Mitchell's credit card had been used to purchase a ticket that evening. Additionally, Thompson interviewed Mitchell's friend, who stated that he had met up with Mitchell at the movie theater. However, the theater for the particular movie they went to see was almost full and so Mitchell and his friend had to sit in different sections of the theater. The friend could not say for sure whether Mitchell had been present throughout the entire movie, but Mitchell was there when the movie ended and the friend believes Mitchell had been in the theater the whole time. The friend did not remember seeing anyone get up and leave the theater. Detective Thompson noted that the movie's duration was 2 hours, 50 minutes, and it was located approximately 10 minutes from the scene of the crime. In terms of time stamps, Detective Thompson also noted that Justine Alexander's murder did take place during the time in which Mitchell was allegedly present at the movie theater.

## Phone Records

Detective Thompson was able to review the call-logs from the cell phones of Justine Alexander, the victim, and David Mitchell, the defendant. The detective found that Mitchell and Alexander did have some correspondence on the day of the incident and throughout the course of that week. The phone logs indicate that Mitchell and the victim had an approximately fifteen minute conversation on the day of the murder, a few hours before the murder occurred. When asked about these conversations, Mitchell states that the conversation was related to new a project that was being developed at work. Mitchell said that it was necessary to ask her some questions about the part of the project that she was working on, so that the Mitchell could proceed with the part of the project that he was working on. The victim's roommate was able to confirm what the phone records had indicated, that a conversation did take place on the day of the murder. The roommate stated that in general, the victim, Alexander, had a good relationship with her coworkers and had not mentioned any trouble at work. However, on the day in question, the roommate had overheard some of the phone call between Mitchell and Alexander, and the roommate stated that it was an "intense conversation." The roommate also remembered hearing some references to work-related topics and that there was mention of who would be getting credit for a new project.

Before moving on to the final phase of the trial, take a moment to consider the evidence you just read and to rate your opinions of that evidence.

- To what extent does the alibi that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty	Implies Guilty
<hr/>	
-7   -6   -5   -4   -3   -2   -1   0   1   2   3   4   5   6   7	
strongly   moderately   weakly   neutral   weakly   moderately   strongly	
<hr/>	

- To what extent do the phone record that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty	Implies Guilty
<hr/>	
-7   -6   -5   -4   -3   -2   -1   0   1   2   3   4   5   6   7	
strongly   moderately   weakly   neutral   weakly   moderately   strongly	
<hr/>	

## Appendix I. – Post-Trial Instructions, Questionnaires, & Manipulation Checks

### Post-Trial Instructions

You have now heard all of the evidence in this case. As such, it is time to render a final verdict and discuss your opinions about this case.

**(Participants in the ‘justify’ and ‘justify + devil’s advocate’ also read:** In a moment, we will have you justify your decision. For now, please provide us with your verdict and your opinions about the case.)

- What is your verdict for David Mitchell? Circle one:

**Guilty      Not Guilty**

- How confident are you that the verdict you reached in this case was accurate?

Not Confident

Confident

---

-7   -6   -5   -4   -3   -2   -1   0   1   2   3   4   5   6   7

Very      moderately      weakly      neutral      weakly      moderately      Very

---

- How likely is it that David Mitchell is Guilty? Circle one:

Very Unlikely    Somewhat Unlikely    Neutral    Somewhat Likely    Very Likely

### **Post-Trial Evidence Evaluations**

Take some time to consider each piece of evidence that you read about. Rate the extent to which each piece of evidence implies the suspect is not guilty or implies the suspect is guilty.

- To what extent does the polygraph evidence that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

---

- To what extent does the motive that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7   -6   -5   -4   -3   -2   -1   0   1   2   3   4   5   6   7  
strongly   moderately   weakly   neutral   weakly   moderately   strongly

---

- To what extent does the shoe print analysis that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7   -6   -5   -4   -3   -2   -1   0   1   2   3   4   5   6   7  
strongly   moderately   weakly   neutral   weakly   moderately   strongly

---

- To what extent does the fingerprint analysis that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7   -6   -5   -4   -3   -2   -1   0   1   2   3   4   5   6   7  
strongly   moderately   weakly   neutral   weakly   moderately   strongly

---

- To what extent does the traffic camera footage you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty							Implies Guilty							
-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

- To what extent does the eyewitness identification that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty							Implies Guilty							
-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

- To what extent does the alibi that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty							Implies Guilty							
-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

- To what extent do the phone record that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

---



### Open-Ended Verdict Discussion

- Please take some time to justify and defend the decision that you came to. Let us (the lawyers) know what factors influenced your decision and why you believe the decision you came to was the correct one. It may help to imagine that you are trying to convince someone who reached the opposite verdict to change their mind.

### Original Evidence Evaluations

Remember how at the beginning of this activity, in the first packet, you rated the various pieces of evidence and matters of legal policy? Try to remember how you rated each piece of evidence and indicate those ratings here. If you cannot remember the exact number, try to guess the number as closely as possible. Do not look back at your first packet to see your original ratings.

- To what extent does the polygraph evidence that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7   -6   -5   -4   -3   -2   -1   0   1   2   3   4   5   6   7

strongly   moderately   weakly   neutral   weakly   moderately   strongly

---

- To what extent does the motive that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty							Implies Guilty							
-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

- To what extent does the shoe print analysis that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty							Implies Guilty							
-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

- To what extent does the fingerprint analysis that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty							Implies Guilty							
-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

- 
- To what extent does the traffic camera footage you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7   -6   -5   -4   -3   -2   -1   0   1   2   3   4   5   6   7  
strongly   moderately   weakly   neutral   weakly   moderately   strongly

---

- To what extent does the eyewitness identification that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7   -6   -5   -4   -3   -2   -1   0   1   2   3   4   5   6   7  
strongly   moderately   weakly   neutral   weakly   moderately   strongly

---

- To what extent does the alibi that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7   -6   -5   -4   -3   -2   -1   0   1   2   3   4   5   6   7  
strongly   moderately   weakly   neutral   weakly   moderately   strongly

- 
- To what extent do the phone record that you read about imply that the suspect is guilty or not guilty? Circle one number.

Implies Not Guilty

Implies Guilty

---

-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

---

## Legal Opinions Survey

This questionnaire is designed to assess people's opinions and attitudes on various legal issues. Please read each item carefully and give as true a picture as possible of your own beliefs by writing in the number 1, 2, 3, 4, or 5, as appropriate.

1. *Strongly agree*
  2. *Somewhat agree*
  3. *Agree & disagree equally*
  4. *Somewhat disagree*
  5. *Strongly disagree*
- 

- \_\_\_ 1. Appointed judges are more competent than elected judges.
- \_\_\_ 2. A suspect who runs from the police most probably committed the crime.
- \_\_\_ 3. A defendant should be found guilty only if 11 out of 12 jurors vote guilty.
- \_\_\_ 4. Most politicians are really as honest as humanly possible.
- \_\_\_ 5. Too often jurors hesitate to convict someone who is guilty out of pure sympathy.
- \_\_\_ 6. In most cases where the accused presents a strong defense, it is only because of a good lawyer.
- \_\_\_ 7. In general, children should be excused for their misbehavior.
- \_\_\_ 8. The death penalty is cruel and inhumane.
- \_\_\_ 9. Out of every 100 people brought to trial, at least 75 are guilty of the crime with which they are charged.
- \_\_\_ 10. For serious crimes like murder, a defendant should be found guilty if there is a 90% chance that he committed the crime.
- \_\_\_ 11. Defense lawyers don't really care about guilty or innocence, they are just in the business to make money.
- \_\_\_ 12. Generally, the police make an arrest only when they are sure about who committed the crime.
- \_\_\_ 13. Circumstantial evidence is too weak to use in court.

- \_\_\_14. Many accident claims filed against insurance companies are phony.
- \_\_\_15. The defendant is often a victim of his own bad reputation.
- \_\_\_16. If a grand jury recommends that a person be brought to trial, then that person probably committed the crime.
- \_\_\_17. Extenuating circumstances should not be considered—if a person commits a crime, then that person should be punished.
- \_\_\_18. Hypocrisy is on the increase in society.
- \_\_\_19. Too many innocent people are wrongfully imprisoned.
- \_\_\_20. If a majority of evidence—but not all of it—suggests that the defendant committed the crime, then the jury should vote *not guilty*.
- \_\_\_21. If someone commits a victimless crime like gambling or possession of marijuana, he should not be convicted.
- \_\_\_22. Some laws are made to be broken.

**General questions related to the material you read:**

Which of the following Instructions were you given at the beginning of the trial?

**Circle one:**

a. At the conclusion of the trial you will be asked to justify your decision to the lawyers (in writing) and to the researcher (in person) who is currently running your study. Be prepared to discuss your eventual decision and to cite the aspects of the case that support your view.

b. As you read about the various pieces of evidence presented throughout the trial, you will be asked to take some time to consider any issues with those pieces of evidence. If you were the opposing attorney, what problems, flaws, or issues would YOU find with each piece of evidence?

c. a. & b. (meaning you were given both the ‘justify’ and the ‘consider the flaws’ instructions).

d. You received ‘none of the above’ instructions.

- If you answered ‘b’ or ‘c’ to the above question, please answer the following two questions. Otherwise, move on to the next page.

- Tell us about your thought process as you considered the issues with each piece of evidence?

- Did considering the issues with each piece of evidence affect your view of the trial overall?

- How much pressure did you feel to provide an accurate verdict, of either guilty or not guilty, in this case?

Lack of pressure to justify

Pressure to justify

-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
very low pressure			moderately low pressure			neutral			moderately high pressure			very high pressure		

Open-Ended Question:

- Were there any similarities between the first packet of materials you read, where you provided your general legal attitudes, and the other packets of materials, such as when you read about the trial? If so, what were they:



Open-Ended Question:

- Do you think the responses you have provided will help the lawyers to prepare for their trial? And if so, why or why not:

- Did your opinions of the evidence you read about change throughout the course of coming to your final verdict of guilty or not guilty?

“My opinions about the evidence (circle one):

“Did not change” “Changed a little” “Changed somewhat” “Changed a lot” “Changed all the time”

### Demographics Questionnaire

1. How old are you? \_\_\_\_\_
2. What is your gender? Check one: \_\_\_\_\_ Male \_\_\_\_\_ Female
3. What is your ethnic origin/race?  
\_\_\_\_\_ White/Caucasian \_\_\_\_\_ American Indian  
\_\_\_\_\_ Hispanic \_\_\_\_\_ Asian  
\_\_\_\_\_ African American \_\_\_\_\_ Other \_\_\_\_\_
4. Indicate your highest educational training as of this date. (check one)  
\_\_\_\_\_ Freshman \_\_\_\_\_ Junior  
\_\_\_\_\_ Sophomore \_\_\_\_\_ Senior  
\_\_\_\_\_ Professional or graduate school \_\_\_\_\_ Other \_\_\_\_\_
5. What is your current work status? Check one:  
\_\_\_\_\_ Employed full time \_\_\_\_\_ Employed part time \_\_\_\_\_ Unemployed
6. What is your occupation? \_\_\_\_\_
7. What is your place of employment? \_\_\_\_\_
8. Is English your primary language? Yes \_\_\_\_\_ No \_\_\_\_\_  
If English is not your primary language, what is your primary language?  
\_\_\_\_\_
9. Do you speak Spanish? Yes \_\_\_\_\_ No \_\_\_\_\_  
If you speak Spanish, how many years have you spoken Spanish? \_\_\_\_\_ Years
10. Have you ever served as a juror before? Yes \_\_\_\_\_ No \_\_\_\_\_  
If yes, was it a civil case or a criminal case? Civil \_\_\_\_\_ Criminal \_\_\_\_\_

## Appendix J. – Recruitment Blurb

“Have you ever wondered why a jury in one of your cases rendered a guilty verdict despite the majority of evidence pointing towards your client's innocence? Or how police officers arrested your client despite there being other more likely suspects?

Come join us for a lunch-time presentation on the pitfalls of human decision-making and how you, as defense attorneys, are in a unique position to address such problems!

The session will begin with a brief demonstration activity that involves reading through a controversial case and providing your opinions. Following that, Dr. \_\_\_\_\_ (name removed) will discuss some serious decision-making biases facing our legal system; issues which should be relevant to your practice.”



(University Logo Removed)

## Study Participation

Your help is needed: The Legal Psychology Department at \_\_\_\_\_ (name of university removed), in conjunction with the National Bar Association (NBA), has been awarded a grant to investigate whether lawyers in different parts of the country are using the same criteria when making decisions related to their cases. Specifically, the Bar seeks to compare whether the opinions of lawyers from various parts of the country match up with the opinions of lawyers from Illinois, concerning a particular case that has already been decided. If you would like to assist, please begin by reading through the general instructions.

## General Instructions

As you proceed through the following materials, please read everything carefully and answer the questions to the best of your ability. Even if you are typically a defense attorney or a prosecutor, we ask that you do not take a perspective in this case. Rather, please strive to remain an impartial evaluator. Additionally, although there is a legal distinction between being innocent and being not guilty, please treat innocence and lack of guilt as synonymous terms for the purposes of this study. Once you have completed or turned-over a page, we ask you not to return to any previous pages in your packet, thus,

some questions may have to be answered from memory. If any question or rating scale does not make sense to you, let us know. Please turn off or silence any electronic devices.

**The Murder of Justine Alexander**

Case Docket: 03801-97844

On July 15<sup>th</sup> 2012, a young woman named Justine Alexander was going for a walk in a local park near her apartment in Rockford, IL, in the county of Winnebago. It was a tradition of hers to go to this park most evenings after work. However, one night her roommates grew concerned when she did not come home by 1 a.m. on a weeknight and was not responding to text messages. The roommates called the police and gave them potential whereabouts for Justine.

The police performed an initial search of the nearby park but were unsuccessful at finding her. However, the police returned the next day to search for a second time and they performed a canine-assisted search of the park and its surroundings. The dogs led them to an outcropping of bushes and trees, in which Justine's body had been thrown. She was declared dead at the scene. It was clear that the cause of death was several stab wounds and that there appeared to have been a struggle.

After they had found the body, Detective Thompson was placed in charge of the investigation and he began interviewing people who had been at the park that evening. A few people who were there claimed to have seen a suspicious-looking person wandering around the periphery of the park. The police were able to put together a description of this person, even though the description was quite general. Additionally, the police interviewed Justine's friends and coworkers at the marketing firm that she worked at. Based on this information, police came to believe that David Mitchell, a fellow coworker,

was the most likely suspect. The investigators learned that Mitchell had anger-management and domestic violence issues in the past.

Now that you have a general understanding of this case, you will be presented with the specific evidence. The information you have read so far is a description of what happened during the crime and the police investigation, but it is not to be considered evidence.

## Appendix M. – Instructions for lawyers regarding Case Evidence

### **Case Evidence (Part 1)**

Now that you have read through the case summary, you will be presented with the majority of the evidence in this case. Some of the evidence is inculpatory and some of the evidence is exculpatory. Please withhold arriving at any final conclusions until you have heard all of the evidence in this case (during Part 2 of Case Evidence) and until we ask you for your final decisions. Remember that your job is to remain as objective as possible when evaluating the evidence, try not to take a side in the case.

### **Case Evidence (Part 2)**

You heard some of the evidence in this case already, but two more pieces of evidence will be presented to you. These two pieces of evidence surfaced later in the investigation and have been stipulated, or agreed upon, by both sides of the case.



Appendix N. – Post-trial Questionnaires for lawyers

**Post-Case Instructions**

You have now heard all of the evidence in this case. Now it is time for you to provide us with your final decisions and opinions.

- As an objective party, are you more inclined to believe that David Mitchell is innocent or guilty? Circle one:

**Innocent**   or   **Guilty**

- How confident are you in your above response?

---

0	10	20	30	40	50	60	70	80	90	100
Not confident at all				Moderately confident				Extremely confident		

---

- David Mitchell is:

---

1	2	3	4	5	6	7
Very Likely Innocent				Very Likely Guilty		

---

- If you were David Mitchell's defense attorney, do you believe it would be in his best interest to negotiate a plea bargain or for the case to go to trial? Circle One:

**Negotiate a Plea Bargain    or    Go to Trial**

- Regardless of your answer to the previous question, please respond to the following question:

If you were to negotiate a plea bargain, what is the maximum numbers of years in prison that you would recommend Mitchell accept in a plea deal?

\_\_\_\_\_ years.

### Post-Case Evidence Evaluations

Take some time to consider each piece of evidence that you read about. Rate the extent to which each piece of evidence implies the suspect is innocent or guilty.

- To what extent does the eyewitness identification that you read about imply that David Mitchell is innocent or guilty? Circle one number.

**Implies Innocence**

**Implies Guilt**

---

-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

---

- To what extent does the polygraph evidence that you read about imply that David Mitchell is innocent or guilty? Circle one number.

**Implies Innocence**

**Implies Guilt**

---

-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

---

- To what extent does the shoe print analysis that you read about imply that David Mitchell is innocent or guilty? Circle one number.

**Implies Innocence**

**Implies Guilt**

---

-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

---

- To what extent does the fingerprint analysis that you read about imply that David Mitchell is innocent or guilty? Circle one number.

**Implies Innocence**

**Implies Guilt**

---

-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

---

- To what extent does the traffic camera footage you read about imply that David Mitchell is innocent or guilty? Circle one number.

**Implies Innocence**

**Implies Guilt**

---

-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

---

- To what extent does the motive that you read about imply that David Mitchell is innocent or guilty? Circle one number.

**Implies Innocence**

**Implies Guilt**

---

-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

---

- To what extent does the alibi that you read about imply that David Mitchell is innocent or guilty? Circle one number.

**Implies Innocence**

**Implies Guilt**

---

-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

---

- To what extent do the phone records that you read about imply that David Mitchell is innocent or guilty? Circle one number.

**Implies Innocence**

**Implies Guilt**

---

-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
strongly		moderately		weakly		neutral		weakly		moderately		strongly		

---

### **Open-Ended Discussion**

- We would like you to justify and defend the conclusions you reached about this case.  
Why do you believe Mitchell is either innocent or guilty and why would you have decided to either go to trial or negotiate a plea deal?

## Legal Opinions Survey

This questionnaire is designed to assess people's opinions and attitudes on various legal issues. Please read each item carefully and give as true a picture as possible of your own beliefs by writing in the number 1, 2, 3, 4, or 5, as appropriate.

1. *Strongly agree*
2. *Somewhat agree*
3. *Agree & disagree equally*
4. *Somewhat disagree*
5. *Strongly disagree*

- 
- \_\_\_ 1. Appointed judges are more competent than elected judges.
  - \_\_\_ 2. A suspect who runs from the police most probably committed the crime.
  - \_\_\_ 3. A defendant should be found guilty only if 11 out of 12 jurors vote guilty.
  - \_\_\_ 4. Most politicians are really as honest as humanly possible.
  - \_\_\_ 5. Too often jurors hesitate to convict someone who is guilty out of pure sympathy.
  - \_\_\_ 6. In most cases where the accused presents a strong defense, it is only because of a good lawyer.
  - \_\_\_ 7. In general, children should be excused for their misbehavior.
  - \_\_\_ 8. The death penalty is cruel and inhumane.
  - \_\_\_ 9. Out of every 100 people brought to trial, at least 75 are guilty of the crime with which they are charged.
  - \_\_\_ 10. For serious crimes like murder, a defendant should be found guilty if there is a 90% chance that he committed the crime.
  - \_\_\_ 11. Defense lawyers don't really care about guilty or innocence, they are just in the business to make money.
  - \_\_\_ 12. Generally, the police make an arrest only when they are sure about who committed the crime.
  - \_\_\_ 13. Circumstantial evidence is too weak to use in court.

- \_\_\_14. Many accident claims filed against insurance companies are phony.
- \_\_\_15. The defendant is often a victim of his own bad reputation.
- \_\_\_16. If a grand jury recommends that a person be brought to trial, then that person probably committed the crime.
- \_\_\_17. Extenuating circumstances should not be considered—if a person commits a crime, then that person should be punished.
- \_\_\_18. Hypocrisy is on the increase in society.
- \_\_\_19. Too many innocent people are wrongfully imprisoned.
- \_\_\_20. If a majority of evidence—but not all of it—suggests that the defendant committed the crime, then the jury should vote *not guilty*.
- \_\_\_21. If someone commits a victimless crime like gambling or possession of marijuana, he should not be convicted.
- \_\_\_22. Some laws are made to be broken.



**General questions related to the material you read:**

- Which of the following Instructions were you given at the beginning of the survey?

**Circle One:**

a. As you progress through this case, we are going to ask you to ‘justify’ your opinions about the case and about the evidence. Which pieces of evidence do YOU find to be the most compelling, probative, and important and why? At the conclusion of this case, you will be asked to justify your decisions (in writing) to the National Bar Association and to the researchers at FIU.

b. As you read about the various pieces of evidence presented throughout the case, you will be asked to take some time to consider any weaknesses that those pieces of evidence have. What problems, flaws, or issues would YOU find with each piece of evidence and why are those flaws important?

c. a. & b. (meaning you were given both the ‘justify’ and the ‘consider the flaws’ instructions).

d. You received none of the above instructions.

- How motivated did you feel to try to help out the Bar and the faculty at FIU by providing an accurate assessment of the suspect’s guilt?

---

1	2	3	4	5	6	7
Not Motivated at all		Somewhat Motivated			Extremely Motivated	

---

Open-Ended:

- Do you have any comments or questions about your experience of reviewing this case?

- Throughout the course of reviewing this case, you heard about each piece of evidence and then eventually came to a final decision about the suspect's guilt or lack thereof.

To what extent did your opinions about the evidence change throughout that process?

---

1	2	3	4	5	6	7
Did Not Change					Changed Completely	

---

### Demographics Questionnaire

1. How old are you? \_\_\_\_\_
5. What is your gender? Check one: \_\_\_\_\_ Male \_\_\_\_\_ Female
6. What is your ethnic origin/race?  
\_\_\_\_\_ White/Caucasian \_\_\_\_\_ American Indian  
\_\_\_\_\_ Hispanic \_\_\_\_\_ Asian  
\_\_\_\_\_ African American \_\_\_\_\_ Other \_\_\_\_\_
4. Is English your primary language? Yes \_\_\_\_\_ No \_\_\_\_\_
5. What is your occupation? \_\_\_\_\_
6. If you're a lawyer, what type of law do you practice? \_\_\_\_\_
7. If you practice criminal law, are you a defense attorney or a prosecutor?  
\_\_\_\_\_

## Appendix O. – Recruitment Blurb for Study 3

Online Study: Help \_\_\_\_\_ (name of university removed) graduates prepare for an upcoming trial

Help \_\_\_\_\_ (name of university removed) grads prepare for an upcoming trial! The Legal Psychology faculty here at \_\_\_\_\_ have been asked by local lawyers, who are recent graduates of the \_\_\_\_\_ law school, to help gather opinions about an important upcoming murder trial. We are looking for students to read through a controversial case and provide us with your feedback and verdict. By providing the lawyers with your feedback, we hope to help them better understand their case. They have worked with us to create a case summary based on the actual evidence that will be presented at trial. The case facts and evidence will be presented to you to help you come to a decision about the defendant's guilt or lack thereof.

## Appendix P. – Prosecution’s Story

### **Prosecution’s Account of the Case:**

Ladies and gentlemen of the jury, the tragic murder of Justine Alexander was a dirty, heat-of-the-moment, and very personal crime. Justine was stabbed numerous times all over her body and it’s clear that there was a struggle. Following that, her body was unceremoniously tossed into a nearby bush, as if the killer thought it was trash. Justine’s death has devastated the victim’s family.

Once you have reviewed the evidence in this case, you’re going to see that this was a clear example of work-related, professional jealousy that lead to a dispute, which escalated into David Mitchell losing his temper and striking out in a rage. This was not a cold and calculated execution from afar with a gun...it was a heinous fight that has all of the tell-tale signs of a crime of passion.

On the evening in question, this incident began with a phone call. Detective Thompson confirmed that the phone records indicate that David Mitchell and Justine Alexander had indeed had correspondence that day. In fact, the records show that there was a 15 minute long conversation between them prior to the murder. Moreover, Justine’s roommate stated that she overheard an “intense conversation” that was related to who would get credit for an important project at work. That is critical. Clearly, there is a record of a dispute here that left Mitchell feeling slighted and set this whole crime in motion. A fifteen minute, heated phone call is not just a simple chat.

It was clear that after the dispute, Mitchell was not happy with the outcome. He went to a nearby park where he must have known Justine was jogging; witnesses stated that someone suspicious had been wandering around the edge of the park. In fact, the witness James McDougal, who was a nightly jogger in that park, said he saw Mitchell at the park that evening and he identified Mitchell from a lineup without any hesitation. It’s as simple as that folks, we have someone who saw Mitchell at the scene of the crime on the evening in question. The defense may try to tell you the street camera from the intersection near the park did not show Mitchell’s car, but that doesn’t mean anything. He could have walked into this large park from any direction.

Once in the park, he stalked Justine until she was in a secluded area...and then he attacked. The detectives found fingerprints on the knife that matched those of Mitchell. Any forensic expert will tell you that murder with a knife represents passionate, personal killing. Mitchell must have tried to wipe off his prints in a hurried attempt to conceal what he had done, but in the heat of the moment, he did a sloppy job of it. The majority of the prints remained.

Following the incident he would have fled the scene. He must have found a place to dispose of his bloodied clothes and his bloodied shoes. After the detectives

interviewed people at the park and people at Justine's work, it didn't take them long to settle on Mitchell as the suspect. The fingerprints matched. And indeed, when Mitchell was given the polygraph, he had elevated arousal throughout the entire examination. The defense may claim he didn't show elevated arousal in response to questions about the crime, but if your arousal is constantly high, it's not possible to see a difference in the arousal based on the questions.

As stated in the beginning, this is a clear-cut case of professional rivalry that led to a bad decision on Mitchell's part. The evidence is damning and it all paints a picture of Mitchell having a motive, being at the scene of the crime according to an eyewitness and fingerprints, and being emotional during subsequent polygraph. We are confident that you will find the appropriate verdict in this case and that justice for Justine's family will be served.

## Appendix Q. – Defense’s Story

### **Defense’s Account of the Case:**

Members of the jury, Justine Alexander’s death was indeed a tragedy. We all mourn the loss of a bright young woman, who, by all accounts, was very friendly and warm. But what the prosecution is trying to do is grasp at straws to weave a tale of ‘worked-related rage.’ The evidence in this case is circumstantial at best and simply does not support the idea of a crime of passion, as we’ll discuss.

It’s an unfortunate fact of life that random killings, muggings, and rapes occur every day. Typically there is no justification needed in those crimes, it’s just a matter of the victim being in the wrong place at the wrong time. Justine’s tragic death was most likely perpetrated by someone she had never met. Think about it: does a white collar professional like Mitchell really fit the profile for someone who is going to go out and stab someone? Let’s start at the beginning.

Yes, there had been a phone call between Mitchell and Alexander that day. But as the records indicate, the phone call was several hours prior to the incident. That is not consistent with a ‘crime of passion.’ He had hours to cool off and it’s absurd to think that he just ran out of the house with a knife. I know I have had heated words with coworkers and bosses over the phone, who hasn’t? It’s a gigantic stretch to assume that there’s a motive here. The detectives never heard anything from Mitchell’s boss about tension or problems at work, or they would have discussed that in their case.

Even if Mitchell was angry, he had no way of knowing how and when Justine would be at the park. There is no record of a friendship or personal communications between the suspect in this case and the victim. Furthermore, the eyewitness testimony of McDougal is problematic for a variety of reasons. Most importantly, the lighting in the park during that time of the evening is very dark. I’ve been there, the park is not well lit, which is probably one of the reasons this crime occurred there. Add on top of that the witness was moving fast due to jogging, and the window of opportunity to view anyone would be slim. It’s important to note that David Mitchell is just your average looking, white guy. He does not stand out in any way. Given the considerable problems that we all know about with eyewitness memory, even in ideal conditions, it would be crazy to put stock in this witness’ account. He probably just wanted to help police and picked someone from the lineup. Furthermore, the police did not provide a video of the lineup procedure, as is policy. They could have biased McDougal’s choice or used a lineup in which McDougal stood out. In that situation, anyone would be likely to identify Mitchell even if they had never seen him. In heinous, high profile crimes such as this, it’s not uncommon for the police to try and force the arrest and prosecution of a suspect due to pressure from the public to take action.

There is no hard evidence that places Mitchell at the park on the evening in question. In fact, as the prosecution alluded to, the traffic camera evidence clearly shows that Mitchell's car never went near the only intersection by the park. Again, the prosecution was grasping at straws. The next nearest place to park without going through that intersection is miles away. If this was really a crime of the moment, as the prosecution suggests, it is not likely that Mitchell parked somewhere far away and walked to the park. On the long walk back to his car, if he was blooded and dirty after attacking someone, we would have had a dozen witnesses that saw him.

The investigation of the crime scene does not support the idea of Mitchell being present. First of all, the shoe print analysis was not a match. The pair of shoes worn by the attacker did not match any shoes that Mitchell has owned or would own, as a professional. The shoe prints did not even seem to be the same size.

Of equal importance, is that the fingerprints on the knife were smudged and hard to read. At best they were only a partial match. For fingerprint evidence to be valid, one must have clear prints to compare. How can someone say the prints were a match when half of the print was missing? Take a look at your fingers, the most distinctive part of the finger is in the middle. If that part is not present, the prints are essentially generic. Furthermore, it was clear that many people had handled this knife and that the prints must have become distorted in the struggle between Alexander and her attacker. For all of these reasons, we should be truly skeptical of the fingerprint results in this particular case.

The last point I'd like to make is that he passed his polygraph examination. It was the examiner's conclusion that Mitchell was being truthful. Those are the examiner's words. Any of us would be nervous in a polygraph and have overall elevated signs of arousal. The point here is that he did NOT show more arousal for the critical questions relating to the crime.

The prosecution's case is flimsy and circumstantial. There is no hard evidence to back it up. This was not a crime of passion, but rather a random and unfortunate killing. Keep in mind the burden of proof is on the prosecution to prove their case 'beyond a reasonable doubt.' There was more than reasonable doubt in this case. I trust you'll find for our defendant.



### Appendix R. – Manipulation Check for Study 3

Immediately after hearing about the prosecution's account of the case, did you then:

- ☐ read about the evidence and rate your initial opinions. Then later you read about the defense's account of the case
- ☐ read about the defense's account of the case. Then later read about the evidence and rated your opinions.

## VITA

### JONATHAN NICHOLAS CARBONE

Born, Rockford, Illinois

2002-2006

B.A., Psychology  
Beloit College  
Beloit, Wisconsin

2006-2009

M.A., Forensic Psychology  
New York University  
New York, New York

Research Laboratories  
John Jay College of Criminal Justice  
New York, New York

2009-2015

Professor and Teaching Assistant  
Classes Taught: Memory and Introduction to Psychology  
Florida International University  
Miami, Florida

2012-2015

Doctoral Candidate  
Florida International University  
Miami, Florida

## PUBLICATIONS AND PRESENTATIONS

Charman, S., Carbone, J., Kekessie, S., & Villalba, D. (2015). Coherence Effects & Order. In-press in *Applied Cognitive Psychology*.

Carbone, J.N. & Winter, R.J. (2009). *Vigilance in the Courtroom: The Struggle for Jurors' Attention*. Carbone presented this paper at the annual American Psychology

Carbone, J.N. & Fisher, R.P. (2009). *The Implications of Inconsistency on the Witness Stand*. Carbone presented this paper at the annual American Psychology and Law Society Conference.