

Choice and Background Knowledge: How do Individuals Evaluate Accumulating Evidence in A Murder Scenario?

Elizabeth Mackenzie¹, Emily Chalmers², Colin A Wastell¹, Piers Duncan³ & Matthew Roberts⁴

¹ Psychology Department, Macquarie University, Australia

² Psychology Department, University of Tasmania, Australia

³ Defence Science and Technology Group, Australian Department of Defence, Australia

⁴ Computing Department, Macquarie University, Australia

Correspondence: Colin A Wastell, Psychology Department, Macquarie University, Australia

Received: February 27, 2018

Accepted: March 20, 2018

Online Published: April 2, 2018

doi:10.5539/ijps.v10n2p1

URL: <https://doi.org/10.5539/ijps.v10n2p1>

Abstract

Can the simple act of selecting a possible suspect of a crime bias the evaluation of the evidence? Does the typicality of the crime impact the assessment of guilt of a suspect? In two experiments, we examine these two questions and find some remarkable results with implications for law enforcement and jury deliberation. Experiment 1 data show that by allowing participants to choose a most-likely-perpetrator, guilt ratings were substantially higher compared to participants who were not allowed to make a choice. This difference persisted after reading a further body of incriminating evidence. In experiment 2 participants were provided with general and specific background information relevant to a suspect, in other words how common was the crime-suspect scenario. When provided with high plausibility compared to low plausibility information, participants gave higher guilt ratings that persisted after further evidence. The results are interpreted in terms of argument theory which provides a parsimonious explanation of the data. These results have implications for the conduct of investigations, for example: putting in place procedures that minimize the effects of suspect prioritization and background information.

Keywords: choice, base rates, argument theory, murder investigation, garden path

1. Introduction

The investigation of crime requires the most stringent application of rigorous methods of reasoning and evidence evaluation. Decades of research into human reasoning and judgment have shown that human beings are subject to biases, and that they utilize heuristics in order to cope with information rich environments in an adaptive manner (Gigerenzer, Hertwig, & Pachur, 2011; Gilovich, Griffin, & Kahneman, 2002). The implication of the use of heuristics and the consequences of biases are that the evaluation of evidence and the judgements made based on it are not always optimal. In the case of the investigation of crime, this may result in the conviction of an innocent person. Criminal investigators and members of juries are subject to the heuristics and biases that are characteristic of the general public, and so must make greater efforts to minimize their impact on the evaluation of evidence when forming a judgment. Resnikoff, Ribaux, Baylon, Jendly, and Rossy (2015) observed “biases induced by a priori knowledge, as well as emotions ... obviously have to be mitigated by complementary mechanisms” (p. 433). In order to be more rigorous, criminal investigators apply systems of evidence evaluation that aim to minimize the negative effects of ordinary human information processing predispositions. However, Police investigators have shown a tendency to decide, very early in an investigation, on the most likely suspects and target their investigation *against* those suspects (Keppens & Schafer, 2006; Sedly, 1993).

Criminal investigations and trials often involve large amounts of information, extended time sequences, multiple motives, and efforts to deceive. These are complex information situations in which an individual is required to evaluate evidence. Investigators are highly trained and experienced, relying on their training and experience to minimize the impact of biases; jurors are guided by the rules of the court, as well as the instruction of the judge. But is training and guidance sufficient? Whilst legal decision makers, once informed of the effects and consequences of biased processing may be more able to recognise these instances, typical investigative approaches tend to produce errors (Alter & Oppenheimer, 2009; Jamieson, 2003). O'Brien (2009) stated that improving

investigatory decision making “requires identifying those factors that bias decision making, and then finding ways to counter that bias” (p. 331). Argument theory may offer a way to re-interpret the evaluation of evidence as a process inherently focuses on defending a position and so identify those aspects of an investigation that require specific attention if the distorting effects of human cognition are to be minimized (Mercier & Sperber, 2017). It was the aim of the following experiments to determine the impact of the participant nominating or not nominating the most likely perpetrator on guilt assessment (Experiment 1) and to examine the impact of background information on guilt assessment (Experiment 2).

1.1 Argument Theory

Argument theory is a coherence-based model of decision making. It suggests that the mind actively works to maintain a state of coherence and avoid inconsistency within mental representations (Lagnado, 2011). Within argument theory, ‘coherence’ refers to the enhancement of elements of information consistent with previously received information, world knowledge and beliefs, as well as the suppression of inconsistent information, in order to maintain stability within mental representations. In several experiments, Simon et al. (2001) suggested that simply trying to understand a legal case would produce a search for coherence among the evidence. One of the most examined biases which enables stability of representations is confirmation bias. Confirmation bias refers to the phenomenon where individuals seek or interpret evidence that align with their prior beliefs, or an existing hypothesis (e.g. Nickerson, 1998). In a study of confirmation bias, O’Brien (2009) noted that “participants, [when] asked early in the case to name a suspect and state why he might be guilty showed a greater tendency to confirm that hypothesis” (p. 328). As a consequence of coherence maintenance, one decision alternative becomes relatively dominant within the representation, allowing the formation of an early decision with a high level of associated confidence, even when there is ambiguous or contradictory evidence present (Simon et al., 2001).

Further, the theory posits that reasoning is a social process and, as such, it is advantageous to treat claims coming from others critically, and to provide the most persuasive argument when communicating ones’ own ideas (Mercier & Sperber, 2008, 2011, 2017). This position is supported by findings that people are generally poor at producing arguments against, and falsifications of, hypotheses that they generate themselves (Poletick, 1996). However, when an argument is presented as coming from someone else, people are more likely to attempt to falsify it and more readily abandon the hypothesis, if the evidence suggests it is necessary to do so (Cowley & Byrne, 2005). Changing beliefs is effortful and takes time; often, it may be easier to ignore alternate perspectives and arguments rather than modify existing beliefs (Hernandez & Preston, 2013).

1.2 Choice and Background Knowledge

Argument theory predicts that once a person has made a choice they will be more likely to defend it and less likely to change their position, even in the face of contradictory information (Mercier & Sperber, 2008). Within reasoning literature, and more specifically in the criminal-legal context, the cognitive aspects of ‘commitment to a choice’ is an area of significant interest (Ariely & Norton, 2008). Ask and Granhag (2007) demonstrated that, once investigators had provided the name of the suspect they believed to be guilty, they evaluated witnesses who provided statements consistent with their choice as more reliable and credible than witnesses who provided statements implicating an alternate suspect.

Mock investigation studies have provided evidence supporting the idea that individuals develop causal hypotheses in response to complex evidence; particularly once commitment to a choice has occurred, this influences not only search for information but also the processing and evaluation of information that is received (O’Brien, 2009; Simon et al., 2004; Weeks, Wastell, Taylor, Wearing, & Duncan, 2012). Furthermore, interrogator expectations have been shown to influence perceptions of guilt (Kassin, Goldstein, & Savitsky, 2009). Additionally, it has been shown that before a final decision had been made, evaluations of evidence shifted in line with the decision that would be produced (Simon et al., 2004). As commitment to a choice has been shown to influence evaluations of criminal evidence (Ask & Granhag, 2005), further exploration of the effect of choice on decision making is an area of extreme importance to the criminal-legal field.

When processing complex evidence, individuals also reportedly combine and coordinate new information with general knowledge, past experiences, and expectations to generate one or more coherent mental representations (Bex, Van Koppen, Prakken, & Verheij, 2010). In the case of criminal investigators, their training and experience along with their organisation’s standard operating procedures provide guidance for the conduct of investigations. There is a strong tendency for individuals to contextualise problems in relation to their prior knowledge, and for responses to be impacted by ‘strongly held’ prior beliefs (Evans, 2006). Additionally, people’s background knowledge may impact upon their expectations surrounding the decision problem (Medin, Coley, Storms, & Hayes, 2003). Background knowledge is very important to legal decision making. Most investigators use

available evidence in combination with their own experience and knowledge to inform the best course of action for their investigation and decision making (Resnikoff et al., 2015). This, however, highlights that variations in knowledge may also lead to variations in individual's expectations as to what is relevant for consideration in their decision making process (Medin et al., 2003).

Situations often present decision makers with two types of background information, these are general, base rate information which depicts "how things usually are in such situations" and specific indicator information which depicts "how things appear to be in the particular situation" (Bar-Hillel, 1980, p. 1). It is proposed that the general plausibility of an occurrence is determined by the base rate of its occurrence (as understood by the decision maker); an occurrence with a high base rate would be judged as more generally plausible. For example, in the case of a murder, a generally accepted 'fact' would be that following the murder of a cheating spouse, their partner would be the most likely suspect, compared to someone else. Determining the specific plausibility of an occurrence is informed by individuating information relevant to that particular situation, for example: the spouses' knowledge of the infidelity.

It is evident that individuals are capable of producing rationalisations in order to avoid acknowledging information that is inconsistent with their beliefs (Tetlock, 2005). Furthermore, they may be particularly motivated to defend their beliefs once they have publicly committed to them (O'Brien, 2009). It was the aim of the following experiments to examine the effects that perpetrator choice and background information have on the evaluation of evidence and judgement, in the form of guilt ratings. Determining the baseline impact of choice and background knowledge in a mock investigation provides a foundation for the further examination of these variables with the aim of countering biases in real-world investigative situations.

1.3 Fictional Crime Vignettes and Garden Path Methodology

The use of fictional case files containing a description of a crime, a police report and a set of evidence implicating one or more suspects is an established method for examining decision making in the context of complex evidence (Ask & Granhag, 2007; O'Brien, 2009; Wastell, Weeks, Wearing, & Duncan, 2012b). This method allows the examination of how various factors may influence the final judgment produced, such as evidence presented, time allowed for decision production, and commitment to a decision option (Ask & Granhag, 2007; O'Brien, 2009; Simon et al., 2004). Furthermore, 'guilt ratings', or judgments made about a suspect's level of perceived 'guilt', have been utilised as outcome measures in fictional crime experiments (e.g. Lagnado & Harvey, 2008). Similarly, Wastell, Feeney, Coley, and Weeks (2012a) utilised crime vignettes and determined that guilt ratings increased following the presentation of a body of incriminating evidence. In our experiments, a slightly modified case file method was used in order to examine firstly, the impact of the participant nominating or not nominating the most likely perpetrator on guilt assessment (Experiment 1) and secondly, to examine the impact of background information on guilt assessment (Experiment 2).

Feeney, Coley, and Crisp (2010) utilized sequential presentation of information in order to examine the impact of individual pieces of information throughout the process of decision making. The present study used the same "garden path" format of information presentation for the final piece of evidence (see Feeney et al., 2010; Wastell et al., 2012a) to examine the effect of incriminating compared to non-incriminating evidence on guilt assessment in both choice and background knowledge experimental conditions (See Figure 1). The term "garden path" refers to the concept that individuals are led into an expectation which may or may not be upheld by subsequent information. In the present experiments, this involved presenting a piece of evidence, either incriminating or non-incriminating, after the participants had been provided with a body of consistently incriminating evidence. Participants presented with the non-incriminating evidence were 'led down the garden path' before being provided with a piece of evidence that was contrary to their expectations. Previous research has demonstrated that the presentation of information inconsistent with suspect guilt, and inconsistent with those previously presented, significantly reduced subsequent guilt ratings (Wastell et al., 2012a).

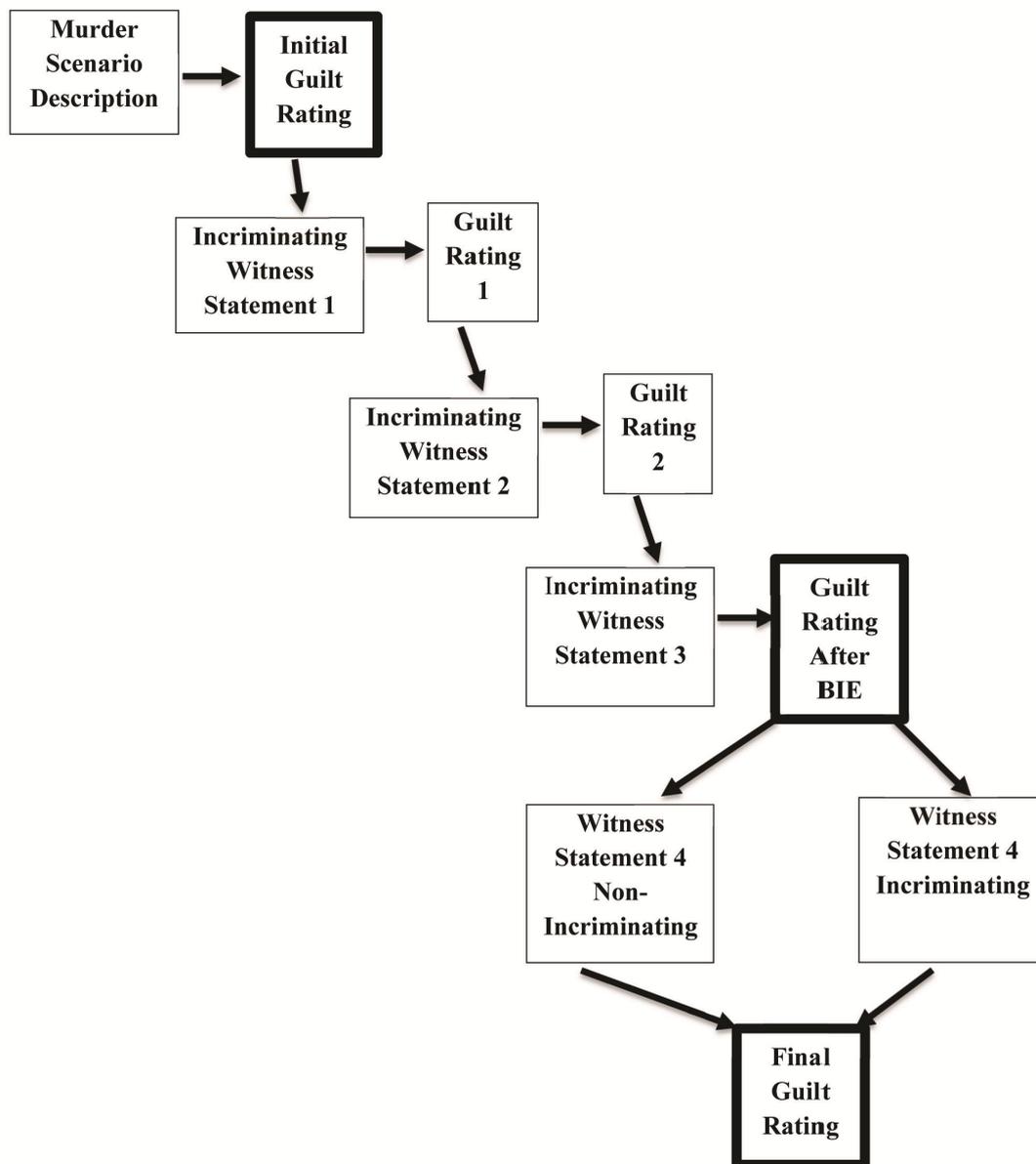


Figure 1. Experimental sequence for both experiments showing garden path. Following the body of incriminating evidence (BIE) guilt rating participants either read a further incriminating witness statement or a non-incriminating witness statement. The boxes in heavy outline correspond to the data points in the graphs in figures 2 and 3.

2. General Method

2.1 Participants

Participants were Australian undergraduate university students who received course credit for their participation. For Experiment 1, there were 110 participants, including 91 females and 19 males, whose ages ranged between 18 and 50 years ($M=20.84$, $SD=4.44$). For Experiment 2, there were 221 participants, including 160 females, 60 males and one other, whose ages ranged from 17 years to 48 years ($M = 20.43$, $SD = 4.36$).

2.2 Materials and Apparatus

Ethics approval for both experiments was obtained from the Macquarie University Human Research Ethics Committee. Materials for this study were presented to participants online, through the survey software *Qualtrics* and Analysis Simulation Project (ASP; adapted from Weeks et al., 2012). ASP recorded the information accessing activity of participants, including guilt ratings and time spent viewing each piece of evidence.

Crime Scenarios and Witness Statements. The materials for both experiments included two crime scenarios, a murder and a theft, adapted from Wastell et al. (2012a; See Appendix A), which included a short description of a fictional case. Accompanying the scenarios, were four additional pieces of evidence, presented sequentially, in the

form of witness statements (taken from Wastell et al., 2012a; See Appendix B). The first three pieces of evidence were consistent with the prime suspect being responsible for the crime; for example: (e.g. Witness: Mary Ladle, Status: Sarah's co-worker, "Sarah is always yelling at someone on the phone. I guess it must be her husband. This morning, yeah, it sounded pretty bad."). The final piece of evidence (FPE) either further incriminated the prime suspect (FPE-I) or was non-incriminating (FPE-N; implicating an alternate suspect). For each crime scenario, participants provided six guilt ratings on an 11-point Likert scale in response to the statement: "Given the information so far, please rate the likelihood that [the indicated suspect] is guilty on a scale from -5 (most probably not guilty) to 5 (most probably guilty)". A high, positive guilt rating indicated a higher level of perceived guilt.

Plausibility Manipulation Check. The plausibility manipulation check was comprised of five true or false questions related to the general and specific plausibility aspects of each scenario (See Appendix C). An example of a true or false question addressing the general plausibility component was: "In similar circumstances, it is more common that the spouse is responsible for the murder". A higher score indicated a higher level of accuracy in participants' ability to explicitly report on the plausibility aspects of each crime scenario.

2.3 General Procedure

The study was advertised through an online research participation portal. Some participants were required to participate internally and, in groups of up to six, these participants were seated at individual computers that were shielded from one another and their start times were staggered. The remaining participants completed the study externally. All of the materials were presented to participants online. After completing an information and consent form and responding to the demographics questionnaire, participants were presented with the first crime scenario; the order of crime presentation was randomized. Participants provided guilt ratings following the presentation of the scenario and following each of the sequentially presented pieces of evidence, in the form of witness statements. Participants were also given the opportunity to revisit each piece of evidence before providing a final guilt rating.

Experiment 1

Hypotheses: Choice of perpetrator

1. In line with argument theory, it was predicted that participants allocated to the choice condition (i.e. able to choose the most likely perpetrator of a crime) would provide significantly higher initial guilt ratings compared to participants allocated to the no-choice condition.
2. In line with previous research (Wastell et al., 2012a) it was predicted that after viewing a **body of incriminating evidence** (BIE), guilt ratings would increase significantly in both conditions. It was also predicted that those in the choice condition would provide significantly higher guilt ratings following a BIE compared to those in the no-choice condition, following from the initial difference in guilt ratings.
3. In an exploratory comparison with previous research we included a third independent variable comparing non-incriminating evidence with further incriminating evidence. It was predicted that only participants in the no choice condition would provide significantly lower guilt ratings when the **final piece of evidence** (FPE) was **not incriminating** (FPE-N) compared when it was **incriminating** (FPE-I). This prediction follows from the assertion that, once a person has made a choice they will be more likely to defend it and less likely to change their position, even in the face of contradictory information (Mercier & Sperber, 2008).

Method

Procedure

In manipulating the independent variable of choice, participants were randomly allocated to either the 'choice' or 'no choice' condition. After reading the crime scenario, participants in the choice condition were presented with the question: "Which suspect do you think is most likely to be guilty?", and then they selected one of the two suspects named in the crime vignette and provided a guilt rating for their chosen suspect. After reading the crime scenario, participants allocated to the no-choice condition were presented with a statement from a third party that nominated the suspect (e.g. "The victims neighbour Mrs Potts believes Ashley to be the prime suspect") and then provided an initial guilt rating for the named suspect.

Participants were presented with four sequential pieces of evidence and provided a guilt rating after each piece. Participants, were shown a BIE, consisting of three witness statements, one at a time. Participants provided a guilt rating after each statement. The final piece of evidence presented to participants was either incriminating (FPE-I) or non-incriminating (FPE-N) and, following this, participants were again asked to provide a guilt rating. After the fourth piece of evidence participants were able to revisit each witness statement and they then provided a final guilt rating.

Results

Inclusion criteria

Guilt ratings greater than two standard deviations from the mean initially, following a BIE, the FPE-I or FPE-N were excluded from analysis (Field, 2013). Analysis was conducted using 92 participants, whose ages ranged from 18 to 50 years ($M = 20.79$, $SD = 4.362$). Sixteen were males and 76 were females.

Data Analysis

A between subjects ANOVA determined there was a significant effect of choice on initial guilt rating, $F(1, 90) = 7.013$, $p = .010$, partial $\eta^2 = .072$. Participants in the choice condition's initial guilt rating was 0.86 points higher compared to those in the no-choice condition ($p = .002$, $d = 0.72$; Figure 2). Similarly, a between subjects ANOVA determined there was a significant effect of choice on guilt rating following a BIE, $F(1, 90) = 9.783$, $p = .022$, partial $\eta^2 = .057$. Participants in the choice condition's guilt rating following a BIE was 0.98 points higher compared to those in the no-choice condition ($p < .0005$, $d = 0.82$). Furthermore, a repeated measures ANOVA determined that there was a significant change from initial to guilt rating following a BIE, $F(1) = 148.886$, $p < .0005$, partial $\eta^2 = .623$; planned pairwise contrast revealed that there was a significant increase of 1.48 and a significant increase of 1.57 for choice ($p < .0005$, $d = 1.32$) and no-choice ($p < .0005$, $d = 1.27$) respectively (Figure 2).

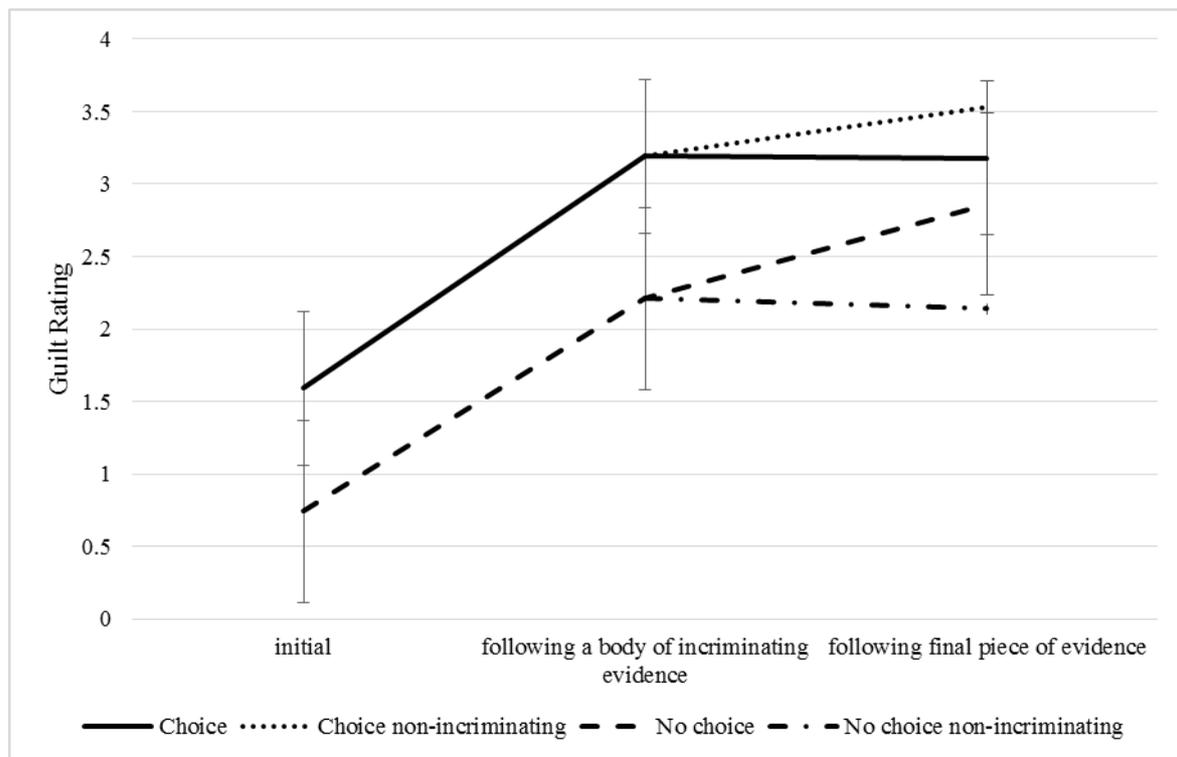


Figure 2. Guilt rating by choice and consistency for Murder. Guilt ratings were averaged across incrimination for the initial, following a body of incriminating evidence.

A between subjects ANOVA was conducted to determine the effect of FPE level of incrimination on guilt rating following the FPE. There was no effect of FPE level of incrimination on the guilt rating following the FPE, $F(1, 88) = 0.066$, $p = .797$, partial $\eta^2 = .037$ (Figure 2).

Summary Experiment 1

In this sample, when participants were able to choose the most likely suspect of a murder their initial and subsequent guilt ratings were substantially and statistically higher than when they are not given that choice. However, when presented with either further incriminating evidence or non-incriminating evidence the differences in guilt ratings were not significantly different in either condition. The data for the final evidence is somewhat unexpected for the choice condition where non-incriminating evidence is associated with an increase, albeit non-significant, in guilt rating.

Experiment 2

Hypotheses –Background Knowledge condition.

1. Previous research has found that individuals use their background knowledge to inform their decision making (Resnikoff et al., 2015). It was predicted that participants in the high plausibility background knowledge condition would report higher initial guilt ratings compared to those in the low plausibility condition. It was also predicted that those in the mixed plausibility conditions would provide significantly lower and significantly higher guilt ratings compared to the high and low plausibility conditions, respectively.
2. Furthermore, it was predicted that participants in the high plausibility condition would report higher guilt ratings following a BIE, compared to those in the low plausibility condition. It was also predicted that those in the mixed plausibility conditions would have significantly lower and significantly higher guilt ratings compared to the high and low plausibility conditions, respectively. Furthermore, in line with previous research (Wastell et al., 2012a) it was predicted that after viewing a BIE, guilt ratings would increase significantly in each of the plausibility conditions.
3. The strong tendency for individuals to contextualise decision situations in relation to their prior knowledge is important, as individuals' background knowledge may impact upon their expectations surrounding the decision problem (Evans, 2006; Medin et al., 2003). Similar to the final prediction in the choice experiment, it was also predicted that only participants in the low plausibility condition would report significantly lower guilt ratings when the final piece of evidence (FPE) was non-incriminating (FPE-N) compared when it was incriminating (FPE-I).

Procedure

For Experiment 2, participants were randomly allocated to one of the four plausibility conditions (high general and specific plausibility; low general and specific plausibility; high general and low specific plausibility; low general and high specific plausibility). Plausibility was manipulated in the scenarios, with respect to the general plausibility and specific plausibility that the designated prime suspect was responsible for the crime. General plausibility was manipulated by indicating that the prime suspect belonged, or did not belong, to a 'category of person' more often responsible for that type of crime, under similar circumstances. For example: "Australian Bureau of Crime Statistics data suggests that, in cases of murder, with similar circumstances, most often (someone other than) the victim's spouse is responsible". The low general plausibility component is included in parentheses. Specific plausibility was manipulated by suggesting that the prime suspect possessed, or did not possess, a specific characteristic that made them more likely to be responsible for the murder.

Similar to Experiment 1, participants were presented with four pieces of evidence, in sequence, and provided a guilt rating after each piece. Additionally, immediately after reading the crime scenario, participants completed the corresponding 'Plausibility Manipulation Check' (See Appendix C). Participants were then shown a BIE, consisting of three witness statements, sequentially, and provided a guilt rating after each of piece of evidence. The last piece of evidence presented to participants was either incriminating (FPE-I) or non-incriminating (FPE-N) and, following this, participants were again asked to provide a guilt rating. As with Experiment 1, following the FPE, participants were able to revisit each witness statement and provided a final guilt rating.

Results

Inclusion criteria

Guilt ratings greater than two standard deviations from the mean, initially, following a BIE, or following the FPE-I or FPE-N, were excluded from analysis (Field, 2013). Participants were also screened for their understanding of the plausibility manipulation (Table 1), however, no additional participants were excluded based upon their responses. Thus, for Experiment 2, analysis was conducted using 210 participants; ages ranged from 17 to 48 years ($M = 20.46$, $SD = 4.457$), 152 were females, 57 were males, and one other.

Table 1. Plausibility Manipulation Check

Crime	Plausibility	N	M	SD
Murder	High	55	4.31	0.982
	Low General, High Specific	55	4.09	1.221
	High General, Low Specific	54	4.54	0.884
	Low	49	4.06	1.248

Note. Average scores for each condition neared the maximum possible score of five indicating that participants were mostly able to accurately explicitly report on the plausibility aspects of the crime scenarios.

Data Analysis

There was a significant effect of plausibility on initial guilt rating, averaged across FPE level of incrimination, $F(3, 206) = 12.717, p < .0005$, partial $\eta^2 = .151$. Planned pairwise contrasts determined that initial guilt rating for high general and specific plausibility was significantly higher compared to each of the other plausibility conditions; whereby, their initial guilt rating was 0.92 points higher compared to low general, high specific plausibility ($p = .005, d = 0.63$), 1.26 points higher compared to high general, low specific plausibility ($p < .0005, d = 0.69$), and 2.26 points higher compared to low general and specific plausibility ($p < .0005, d = 1.22$; Figure 3). Additionally, initial guilt ratings for the low general and specific plausibility condition was significantly lower compared to each of the mixed plausibility conditions; their initial guilt rating was 1.34 points lower compared to low general, high specific plausibility ($p < .0005, d = 0.72$), and 0.99 points lower compared to high general, low specific plausibility ($p = .003, d = 0.46$; Figure 3). However, when comparing the mixed plausibility conditions there was no significant difference in initial guilt rating.

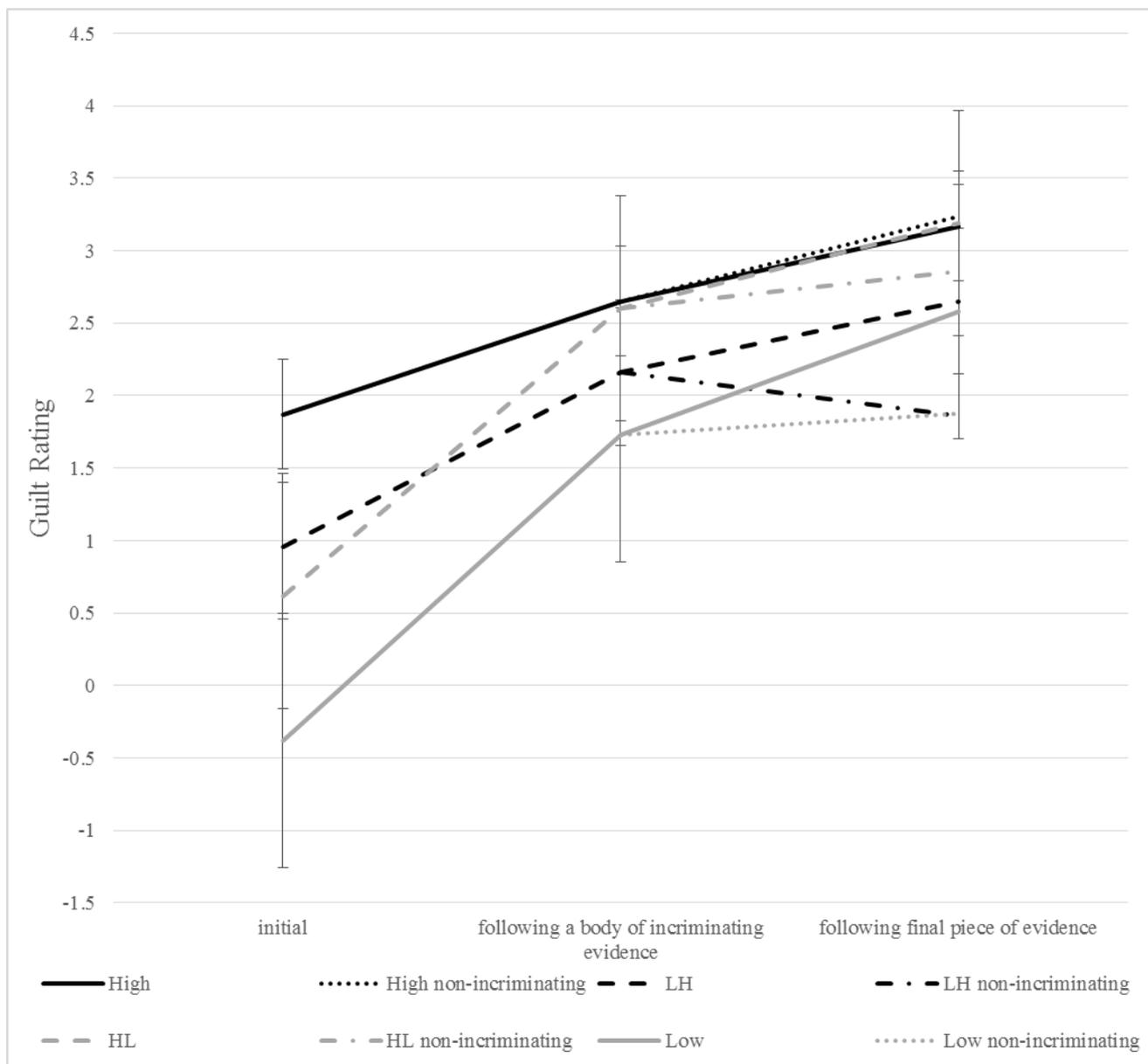


Figure 3. Guilt rating by plausibility and consistency. Guilt ratings were averaged across consistency for the initial and following consistent body of evidence. High = high general and specific plausibility, LH = low general, high specific plausibility, HL = high general, low specific plausibility, Low = low general and specific plausibility.

A between-subjects ANOVA determined there was no significant effect of FPE level of incrimination for the guilt rating following a BIE, $F(1, 201) = 0.941, p = .333, \text{partial } \eta^2 = .005$. There was a significant effect of plausibility on guilt rating following a BIE, averaged across FPE level of incrimination, $F(3, 205) = 3.07, p = .029, \text{partial } \eta^2 = .043$. Planned pairwise contrasts determined that participants in the high plausibility condition's guilt ratings were 0.92 points higher compared to the low plausibility condition ($p = .008, d = 0.51$); however there was no difference when compared to either of the mixed plausibility conditions. Participants in the low plausibility condition also reported significantly lower guilt ratings, by 0.87 points, compared to high general, low specific plausibility ($p = .010, d = 0.44$), however, there was no significant difference compared to low general, high specific plausibility (Figure 3). There was also no significant difference between the mixed plausibility conditions.

A repeated measures ANOVA determined that there was a significant change from initial guilt rating to guilt rating following a BIE, $F(1) = 194.102, p < .0005, \text{partial } \eta^2 = .486$. Planned pairwise contrasts determined that from initial to following a BIE, guilt rating increased 0.77 for high general and specific plausibility ($p = .001, d = 0.59$), 1.20 for low general, high specific plausibility ($p < .0005, d = 0.76$), 1.98 for high general, low specific plausibility ($p < .0005, d = 1.03$) and 2.10 for low general and specific plausibility participants ($p < .0005, d = 0.96$; Figure 3). Finally, a between subjects ANOVA was conducted to determine the effect of FPE level of incrimination on guilt rating following the FPE. There was no effect of FPE level of incrimination on the guilt rating following the FPE, $F(1,199) = 3.001, p = .085, \text{partial } \eta^2 = .015$.

Summary Experiment 2

When primed with high general and specific plausibility background information or a combination of either high or low general and specific plausibility background information, participants in Experiment 2 gave substantially and significantly higher guilt ratings than when primed low general and specific plausibility information. Furthermore, participants primed with high general and specific plausibility information gave the highest guilt ratings compared to all of the other plausibility conditions. These findings suggest that the combination, rather than the individual components, of plausibility had the greatest effect on increasing the perceived guilt of the suspect. After reading a BIE, participants across the plausibility conditions' guilt ratings significantly increased compared to their initial guilt ratings; however, guilt ratings also appeared to converge, whereby the initial differences between plausibility conditions reduced. Additionally, when presented with either further incriminating evidence or non-incriminating evidence the differences in guilt ratings were not significantly different for any of the plausibility conditions.

3. General Discussion

As predicted, our two experiments show that when participants were either given the opportunity to choose the suspect or were provided with background information, in the form of plausibility enhancing information relevant to a nominated suspect, they produced substantially higher guilt ratings. When participants were not given the choice, or were given low plausibility information, their initial guilt ratings were at the mid-point of the scale which indicated a tendency toward neutrality (neither a guilty nor not-guilty rating). These findings indicate that both choice and background information impacted initial decision making. Furthermore, as predicted, after a BIE, guilt ratings in all conditions of choice and plausibility increased substantially. These findings are in line with those of Wastell et al. (2012a) and those that suggest individuals incorporate new information and evidence when forming a coherent mental representation (Bex et al., 2010).

Finally, when comparing the guilt ratings following the FPE-I to FPE-N, there were no significant differences in guilt ratings. In part, this finding is in accordance with Mercier and Sperber's (2008) assertion that individuals are less likely to change their judgment, even in the face of contradictory evidence; however, surprisingly, this occurred whether or not participants made a choice, and occurred regardless of the level of plausibility of the background information. Largely, these findings are consistent with confirmation bias and with the idea that new information is evaluated depending on whether or not it is consistent with the emerging decision and that inconsistent evidence is deliberately suppressed in order to maintain a state of coherence (Simon et al., 2004). The mind actively works to maintain a state of coherence and avoid inconsistency within mental representations (Lagnado, 2011).

The present findings contribute to an existing literature demonstrating inaccuracies in belief updating in response to new information, and is consistent with the assertion that individuals are able to produce cognitive justifications for ignoring or discounting information that is not in line with their existing beliefs (Tetlock, 2005; Tetlock & Gardner, 2015). Argument theory points to the use of a growing, or emerging, coherence as an underpinning of arguments that could be used to defend and justify decisions (Lagnado, 2011). Our results are largely consistent with the assertions of argument theory. In both experiments, early evaluations of guilt varied, dependent upon

either choice or background information. Allowing choice or providing high plausibility background information facilitated a commitment to a suspect, as shown by the initial guilt ratings.

4. Limitations

The current experiments are limited in a number of ways. Firstly, the participants were university students, not criminal investigators. However, university students are generally intelligent and so make a good baseline for the general impact of choice and background information on evidence evaluation. In future research, conducting these experiments with practitioners is required to assess the impact of training and experience on decision making; however, assessing the impact of expertise was not the aim of the present experiments.

Secondly the participants did not choose the evidence, it was presented in a predetermined sequence. This limitation meant that any individual preference for facts and possibly different stories was not able to be tested. However, Wastell et al. (2012b) did facilitate such individual evidence accumulation and found that a majority of participants were more inclined to select information consistent with their initial preference than to suspend judgment for any length of time. Participants in Wastell et al.'s experiment also exhibited information selection behaviour consistent with argument theory.

A third limitation is the relative amount of incriminating evidence (three witness statements) compared to the single piece of non-incriminating evidence. The minimal effects seen among the participants in the non-incriminating evidence condition could simply be a result of the amount of evidence. The fourth limitation is the order of non-incriminating evidence presentation. All participants were provided with incriminating evidence immediately after making their initial guilt rating. We did not test order of evidence effects. Future research should address both the third and fourth limitations by providing both variations in amount and order of evidence and so ascertain if these aspects of the situation enhance or diminish the effects detected in our experiments.

5. Conclusion

We set out to establish a baseline for the impact of choosing the perpetrator and of general and specific background information on evidence evaluation and the judgment of guilt in a mock murder investigation. Our results indicate that when given a choice or provided with high plausibility background information implicating a suspect, people were more inclined to provide higher initial guilt ratings that were not substantially diminished by subsequent non-incriminating evidence. This is a concern, as neither of these manipulations, choice nor plausibility, are based on the content of the evidence presented, but are the result of happenstance situational factors. These findings raise concern as to the conduct of investigations and should provide motivation for developing procedures to minimize their effects.

Mercier and Sperber (2017, p. 270) assert that reasoning takes place best in the context of robust “back-and-forth of conversation, when people can exchange arguments and counterarguments”. In terms of countering the effects of perpetrator selection and background information, it would seem that such robust exchanges should directly challenge both of these aspects of an investigation in order to minimize their potentially distorting impact on the evaluation of evidence. Ask and Granhag (2005) agree that premature or unconsidered identification of a suspect can distort a police investigation in pervasive ways. Even though in some countries a rational trial process may be expected to acquit wrongly identified suspects, a trial cannot investigate or even indicate which other people should be suspected; the investigative trail may have gone cold by that time. In jurisdictions where conviction rates are always very high, one might wonder if flawed investigation decisions are ever corrected by trials – and the importance of good investigative decision-making is even more crucial.

References

- Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*, *13*, 219-235. <https://doi.org/10.1177/1088868309341564>
- Ariely, D., & Norton, M. I. (2008). How actions create-not just reveal-preferences. *Trends in Cognitive Sciences*, *12*, 13-16. <https://doi.org/10.1016/j.tics.2007.10.008>
- Ask, K., & Granhag, P. A. (2005). Motivational sources of confirmation bias in criminal investigations: The need for cognitive closure. *Journal of Investigative Psychology and Offender Profiling*, *2*, 43-63. <https://doi.org/10.1002/jip.19>
- Ask, K., & Granhag, P. A. (2007). Motivational bias in criminal investigators' judgments of witness reliability 1. *Journal of Applied Social Psychology*, *37*, 561-591. <https://doi.org/10.1111/j.1559-1816.2007.00175.x>
- Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica*, *44*, 211-233. [https://doi.org/10.1016/0001-6918\(80\)90046-3](https://doi.org/10.1016/0001-6918(80)90046-3)

- Bex, F. J., Van Koppen, P. J., Prakken, H., & Verheij, B. (2010). A hybrid formal theory of arguments, stories and criminal evidence. *Artificial Intelligence and Law*, 18, 123-152. <https://doi.org/10.1007/s10506-010-9092-x>
- Cowley, M., & Byrne, R. M. J. (2005). When falsification is the only path to truth. Paper presented at the Twenty-Seventh Annual Conference of the Cognitive Science Society, Stresa, Italy.
- Evans, J. ST. B. T. (2006). The heuristic-analytic theory of reasoning: extension and evaluation. *Psychonomic Bulletin & Review*, 13(3), 378-395. <https://doi.org/10.3758/BF03193858>
- Feeney, A., Coley, J. D., & Crisp, A. (2010). The relevance framework for category-based induction: Evidence from garden-path arguments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(4), 906-911. <https://doi.org/10.1037/a0019762>
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). LA: Sage.
- Gigerenzer, G., Hertwig, R & Pachur, T. (2011). *Heuristics: The foundations of adaptive behaviour*, New York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199744282.001.0001>
- Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*. New York: Cambridge University Press. <https://doi.org/10.1017/CBO9780511808098>
- Hernandez, I. & Preston, J. L. (2013). Disfluency disrupts the confirmation bias. *Journal of Experimental Social Psychology*, 49, 178-182. <https://doi.org/10.1016/j.jesp.2012.08.010>
- Jamieson, A. (2003). A rational approach to the principles and practice of crime scene investigation: i. principles. *Science & Justice*, 44, 3-7. [https://doi.org/10.1016/S1355-0306\(04\)71678-0](https://doi.org/10.1016/S1355-0306(04)71678-0)
- 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, 187-192. <https://doi.org/10.1023/A:1022599230598>
- Keppens, J., & Schafer, B. (2006). Knowledge based crime scenario modelling. *Expert Systems with Applications*, 30, 203-222. <https://doi.org/10.1016/j.eswa.2005.07.011>
- Lagnado, D. A. (2011). Thinking about evidence. *Proceedings of the British Academy*, 171, 183-223. <https://doi.org/10.5871/bacad/9780197264843.003.0007>
- Lagnado, D. A., & Harvey, N. (2008). The impact of discredited evidence. *Psychonomic Bulletin & Review*, 15, 1166-1173. <https://doi.org/10.3758/PBR.15.6.1166>
- Medin, D. L., Coley, J. D., Storms, G., & Hayes, B. L. (2003). A relevance theory of induction. *Psychonomic Bulletin & Review*, 10, 517-532. <https://doi.org/10.3758/BF03196515>
- Mercier, H., & Sperber, D. (2008). Intuitive and reflective inferences. In *Two Minds: Dual Processes and Beyond*. Oxford: Oxford University Press.
- Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, 34, 57-74. <https://doi.org/10.1017/S0140525X10000968>
- Mercier, H., & Sperber, D. (2017). *The enigma of reason*. Cambridge MA: Harvard University Press.
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, 2(2), 175. <https://doi.org/10.1037/1089-2680.2.2.175>
- O'Brien, B. (2009). Prime suspect: An examination of factors that aggravate and counteract confirmation bias in criminal investigations. *Psychology, Public Policy, and Law*, 15, 315-324. <https://doi.org/10.1037/a0017881>
- Poletiek, F. H. (1996). Paradoxes of falsification. *Quarterly Journal of Experimental Psychology*, 49, 447-462. <https://doi.org/10.1080/713755628>
- Resnikoff, T., Ribaux, O., Baylon, A., Jendly, M., & Rossy, Q. (2015). The polymorphism of crime scene investigation: An exploratory analysis of the influence of crime and forensic intelligence on decisions made by crime scene examiners. *Forensic Science International*, 257, 425-434. <https://doi.org/10.1016/j.forsciint.2015.10.022>
- Sedley, S. (1993). Whose Justice? In *London Review of Books*, page 6, 23 June 1993.
- 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, 1250-1260. <https://doi.org/10.1037/0278-7393.27.5.1250>
- Simon, D., Snow, C. J., & Read, S. J. (2004). The redux of cognitive consistency theories: evidence judgments by

constraint satisfaction. *Journal of Personality and Social Psychology*, 86, 814-819. <https://doi.org/10.1037/0022-3514.86.6.814>

Tetlock, P. E. (2005). *Expert political judgment*. Princeton: Princeton University Press.

Tetlock, P. E., & Gardner, D. (2015). *Superforecasting: The art and science of prediction*. New York: Random House.

Wastell, C., Feeney, A., Coley, J. & Weeks, N. (2012a). *Processing disfluency relates to belief revision via a process of evidence reappraisal*. Paper presented at the International Conference on Thinking, London.

Wastell, C. A., Weeks, N., Wearing, A., & Duncan, P. (2012b). Identifying hypothesis confirmation behaviors in a simulated murder investigation: Implications for practice. *Journal of Investigative Psychology and Offender Profiling*, 9, 184-198. <https://doi.org/10.1002/jip.1362>

Weeks, N. J., Wastell, C. A., Taylor, A. J., Wearing, A. J., & Duncan, P. (2012). Tracing decision processes in complex, ambiguous, information-rich environments. *International Journal of Psychological Studies*, 4, 158-173. <https://doi.org/10.5539/ijps.v4n1p158>

Appendices

Appendix A

Experiment 1.

Murder Scenario

As a private investigator, you have been commissioned to investigate the murder of James Smith. Listed below are some of the case details.

Date: Saturday the 10th of February 2018

Estimated time of death: Between 2:00pm to 4:00pm

Location: 4 High St, Rydel

Cause of death report: James Smith suffered a fatal gunshot wound to the head. Crime scene investigators reported finding gunshot residue on the victims face suggesting that the shooting occurred at close range. The bullet was extracted and examined, yet experts were unable to specify the exact type of weapon the bullet came from.

Initial reports indicate that James Smith was having an affair with Ashley Buddle, a close neighbour of the Smiths. When questioned, Sarah Smith (James' wife) admitted to being aware of the infidelity.

The following pages will present witness statements collected prior to your arrival on scene.

[The victims neighbour Mrs Potts believes Sarah/Ashley to be the prime suspect.¹]

{Which suspect do you think is most likely to be guilty

Ashley

Sarah²}

¹ Included in the no-choice condition

² Included in the choice condition

Experiment 2.

Murder Scenario

As a private investigator, you have been commissioned to investigate the murder of James Smith. Listed below are some of the case details.

Date: Saturday the 10th of February 2018

Estimated time of death: Between 2:00pm to 4:00pm

Location: 4 High St, Rydel

Cause of death report: James Smith suffered a fatal gunshot wound to the head. Crime scene investigators reported finding gunshot residue on the victims face suggesting that the shooting occurred at close range. The bullet was extracted and examined, yet experts were unable to specify the exact type of weapon the bullet came from.

Initial reports indicate that James Smith was having an affair with Ashley Buddle, a close neighbour of the Smiths. When questioned, Sarah Smith (James' wife) admitted to being aware of the infidelity. Australian Bureau of Crime Statistics data suggests that, in cases with similar circumstances, most often [someone other than] the victim's spouse is responsible. The victim's spouse is [not] known to have a history of physically violent tendencies, a factor associated with murder.

At present the prime suspect is *Sarah Smith*.

The following pages will present witness statements collected prior to your arrival on scene. You will also be required to indicate to what extent you believe *Sarah* to be guilty.

Note. Low probability and likelihood aspects are included in parentheses.

Appendix B**Evidence: Witness Statements.**

Murder

Witness: Bob Newton

Status: Sarah's/Ashley's close friend

Yeah, Sarah/Ashley and I go to the gun range all the time. We're avid enthusiasts. This morning, sure, we always go to the range on Saturday

Witness: Mary Ladle

Status: Sarah's/Ashley's co-worker

Sarah/Ashley is always yelling at someone on the phone. I guess it must be her husband. This morning, yeah, it sounded pretty bad.

Witness: Hillary Thompson

Status: Sarah's/Ashley's boss

Sarah/Ashley has had some problems at work lately. I had to give her a warning the other day. Afterwards, others heard her yelling at her husband on the phone

Witness: Jane Morslee

Status: Neighbour

Sarah/Ashley was definitely at the house around that time. I was out for a walk and saw her go inside. What time, I guess it was around 2pm.

Appendix C**Plausibility Manipulation Check.**

Murder Manipulation Check

Based on the information presented to you in the previous section please answer the following true or false questions:

1. In similar circumstances, it is more common that the spouse is responsible for the murder
2. In similar circumstances, it is more common that someone other than the spouse is responsible for the murder
3. People who are convicted of murder are more likely to have physically violent tendencies
4. Emily has a history of physically violent tendencies
5. Emily has no history of physically violent tendencies

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).