The effect of timing of DNA and judicial instruction on coherence-shifts during mock-juror evaluation of DNA evidence.

Meg Feeney

This thesis is submitted in partial fulfillment of the Honours Degree of Bachelor of Psychological Science

School of Psychology

The University of Adelaide

October 2018

Word Count: 8,966

Table of Contents

List of Tables 5

List of Figures 6

Abstract 7

Declaration 8

Acknowledgements 9

Chapter 1: Introduction 10

1.1. The problem with DNA evidence 10

1.2. Not what you say, but when you say it? 11

The Story Matters. 15

Coherence-based reasoning (CBR). 15

Methodological limitations with CBR in the literature. 17

Confidence. 19

Debiasing attempts. 19

1.3. The problem with judicial instructions – a little too late? 20

1.4. The current study. 23

Chapter 2: Method 25

2.1 Ethics 25

2.2 Participants 25

2.3 Materials 26

2.4 Procedure 27

Design statement. 28

Chapter 3: Results 30

3.1 Data screening and quality control 30

3.2. A priori comparisons 30

Hypothesis 1 30

Hypothesis 2 34

Hypothesis 3 37

Hypothesis 4 41

Hypothesis 5 43

3.3. Post-hoc exploratory analysis 44

Chapter 4: Discussion 52

4.1 Overview 52

4.2 What effect does timing of DNA evidence in a criminal trial have on mock-juror evaluation of evidence? 52

4.3 Can a pre-trial instruction mitigate coherence-shifts better than a post-trial instruction? 53

4.4 Strengths 55

4.5 Limitations 56

4.6 Implications for the literature 57

4.7 Implications for the legal system 59

4.8 Directions for future research 59

4.9 Concluding remarks 60

References.....62

Appendix.....75

Appendix A – Ethics Approval.....76

Appendix B – Poster.....77

Appendix C – Survey link to randomize.....78

Appendix D – Participant preamble page and informed consent.....79

Appendix E – Survey.....80

Appendix F – Tests for normality hypothesis 1.....108

Appendix G – Tests for normality Hypothesis 4.....109

Appendix H – Tests for normality for Hypothesis 2.....110

Appendix I – Preliminary results email sent to participants.....111

List of Tables

Table 1: *Aims and hypothesis*.....24

Table 2: *A priori comparisons testing prosecution DNA inflation of subsequent evidence using Mann Whitney U tests* 35

Table 3: *A priori comparisons testing defence DNA inflation of subsequent evidence using Mann Whitney U* 36

Table 4: *A priori comparisons for the effect of pre vs post-trial instruction on initial ratings of DNA evidence using Mann Whitney U*.....38

Table 5: *The effect of pre vs. post-trial instruction on coherence-shift of DNA evidence using Mann Whitney U tests*: 40

Table 6: *A priori comparisons using Mann-Whitney U tests for confidence ratings across conditions*.....42

Table 7: *Distribution of verdict across condition*.....43

Table 8: *Cross tabulation between condition and response*.....46

Table 9: *Demographic information including gender and age across conditions*.....47

Table 10: *Comparisons for second ratings of prosecution DNA evidence* 50

Table 11: *Comparisons for second ratings of defence DNA evidence* 51

List of Figures

Figure 1.Theoretical frameworks 12

Figure 2. Pathways of testimony evaluation under explanatory coherence (Thaggard, 2005) 14

Figure 4. Experimental conditions 29

Figure 5. Boxplot showing initial rating of prosecution DNA evidence across conditions. 32

Figure 6. Boxplot showing prosecution DNA coherence-shift across conditions 33

Figure 7. Boxplot showing initial ratings of defence DNA across groups. 34

Figure 8. Boxplot displaying confidence ratings in verdict across conditions. 41

Figure 9. Plot of probability of guilt and verdict. 44

Figure 10. Boxplot showing distribution of confidence ratings by verdict. 45

Figure 11. Boxplot showing secondary ratings of prosecution DNA evidence 48

Figure 12. Boxplot showing secondary ratings of defence DNA evidence 49

Figure 13. Boxplot showing total coherence-shift by verdict 49

Abstract

The integrity of a trial rests on the jury's ability to objectively and accurately evaluate the facts, yet a large body of research indicates jurors do not reason objectively. The Story Model (Pennington & Hastie, 1988) Explanatory Coherence (Thaggard, 1989) and Coherence-Based Reasoning (Simon, 2004) were used as the theoretical basis for the current study. Using a 2*2*2 experimental design 257 participants, acting as mock-jurors read a criminal trial transcript and rated the evidence twice. Here we show that DNA presented early by the prosecution is associated with greater coherence-shifting. A recency effect was found for initial ratings of DNA evidence. This supports the literature that jurors do not reason objectively; however, the two cognitive effects appear to work in opposite directions, resulting in homogenous ratings of that same evidence by deliberation. DNA timing did not impact confidence ratings. Further, DNA did not inflate subsequent evidence, failing to replicate the recent assimilation hypothesis (Rassin, 2017). Pre-trial instruction was associated with better understanding of DNA evidence and lower ratings of prosecution DNA. Inconsistencies in findings indicate that the cognitive processes may be different for the prosecution and the defence. Limitations, including a lack of power, and implications the theoretical frameworks and the legal system are discussed.

Keywords: *coherence-shift; DNA evidence; juror reasoning; judicial instruction*

Word count: 205

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma in any University, and, to the best of my knowledge, this thesis contains no materials previously published except where due reference is made. I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

October 2018

Acknowledgements

This thesis is dedicated to the individuals who have been, and continue to be, wrongfully prejudiced by misapplication of DNA evidence. From the inception of this study, the case of Sally Clark – one of many cases of wrongful conviction - stuck with me. Convicted soley on the basis of an erroneously calculated statistic, it continues to highlight the importance of striving to understand the process by which we decide to take away someone's freedom and label them a criminal, so that miscarriages of justices can be prevented. Thank you to Doctor Carolyn Semmler (specifically for teaching me the importance of hypotheses and transparent research!) - your passion and knowledge are inspiring - and the rest of the teaching staff at The University of Adelaide for making this the most challenging but rewarding year of my tertiary studies. To my family for your love and support. Lastly, to my friends – of which the 2018 honours cohort now includes.

"There is no greater tyrant than that which is perpetuated under the shield of law and in the name of justice."

-Charles de Montesquieu

Sally Clark the day her conviction was quashed.

Chapter 1: Introduction

1.1. The problem with DNA evidence

Although perceived to be reliable and heralded as the 'gold standard' (Lieberman, Carrell Miethe, & Krauss, 2008), DNA is subjective due to inherent human perceptual elements guised in the form of numerical beliefs (Dror & Hampikian, 2011). It is highly prejudicial (Appleby & Kassin, 2016; Dartnall & Goodman-Delahunty, 2006) with up to 93% of people finding DNA the most persuasive evidence of guilt (Lieberman, et al., 2008).

Judges and jurors are unable to identify faulty scientific testimony (McAuliff, Kovera & Nunez, 2009) resulting in untested and flawed DNA testimony being declared admissible (Kovera & McAuliff, 2000). The US National Academy of Science Report on Forensic Science highlighted the potential for faulty, untested forensic science to be given undue weight and result in wrongful convictions (Strengthening Forensic Science in the United States: A Path Forward, 2009). Following this, the PCAST Report raised alarms by citing studies showing 11% of microscopically "matched" hair samples came from different sources and that in 3,000 cases involving microscopic hair analysis, 95% contained invalid testimony inculpating the defendant (President's Council of Advisors on Science and Technology, 2016).

An individual's guilt or innocence should be evaluated objectively on the facts. Yet despite the fallibility of DNA evidence, jurors demonstrate a poor ability to accurately evaluate DNA evidence by overweighting (Faigman & Baglioni, 1988; Lieberman et al., 2008) and underweighting DNA evidence (Nance & Morris, 2005). Further, its statistical expression provides ripe conditions for heuristic and fallacious reasoning, including; 'the prosecutors fallacy', 'the defence-attorney's fallacy' and 'the weak evidence effect' (Thompson 1987; McQuiston-Surrett, & Saks, 2009; Martire, Kemp, Sayle, & Newell, 2014).

This potential threat to the justice system cannot be understated or limited to the United States jurisdiction; 71 wrongful convictions have been identified in Australia between 1922-

2015 (Dioso-Villa, 2015). This thesis explores how prejudice may arise simply by virtue of the order of DNA presentation, and how pre-trial instructions may be more effective than post-trial instructions at combating fallacious reasoning.

1.2. Not *what* you say, but *when* you say it?

Jurors receive evidence piece by piece over time. The following frameworks (see Figure 1) will set the foundation for investigating the effect of receiving DNA evidence at different time points on assignments of probative weight. Further, alongside empirical support, these frameworks will form the basis for examining if pre-trial instruction can mitigate prejudice more than post-trial instruction.

The literature has yet to unify one theory of juror reasoning, despite many of the theoretical frameworks being compatible (Byrne, 1995). Combined, the Story Model (Pennington & Hastie, 1986; 1988; 1992; 1993), Explanatory Coherence (Thaggard, 1989; 2005) and Coherence-Based Reasoning (Simon, 2004) provide a cohesive framework, unaccounted for by heuristics alone. They posit that evidence does not exist in isolation but interacts with surrounding evidence and results in distorted judgements that do not align with optimal judgements as defined by theories of rationality (Plous, 1993).

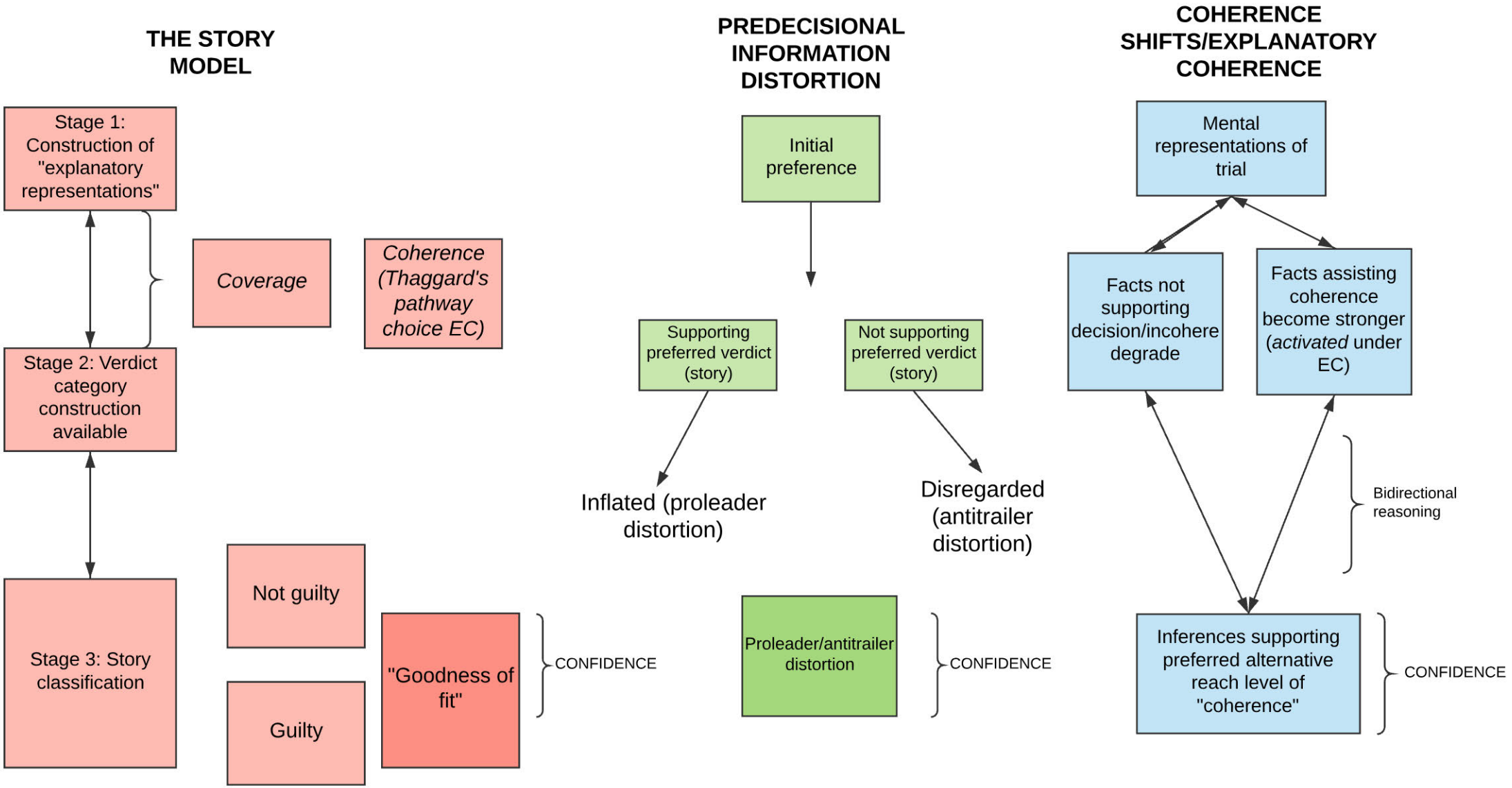


Figure 1.Conceptualization of the theoretical frameworks in the current study

The Story Model (Pennington & Hastie, 1986; 1988; 1992; 1993)

Predecisional Information Distortion (Carlson & Russo, 2001; Blanchard, Carlson & Meloy, 2014)

Explanatory Coherence/Coherence-shifts: (Thaggard, 1989; Simon 2004)

Explanatory coherence. Explanatory coherence posits that the *acceptability* of a verdict depends on the 'set of propositions' it belongs with. Propositions depends on the elements around them and if the elements do not explain a proposition - overall skepticism is increased (Thaggard, 1989). Reasoning is bidirectional (known in explanatory coherence as 'symmetry'), meaning evidence impacts verdict and verdict impacts evidence. One verdict needs more facts supporting it to be 'activated' (receiving reciprocal support) *at a certain level* (beyond reasonable doubt) than the non-chosen verdict (Thaggard, 1989). Facts with no support 'degrade'.

This 'degrading' of facts without support is echoed by empirical research. Knowledge that a cause was involved in an event has been shown to trigger confirmatory search strategies (Shaklee & Fischhoff, 1982). In addition, focusing on single explanations causes competing explanations to reduce credibility as the alternative explanation *inhibits* construction of other explanations and triggers a confirmatory search for information (Sloman, 1994). Given that in a criminal trial, propositions are often not mutually exclusive (i.e. defendant may be at the scene, but did not commit the crime), this inhibition of other explanations would be concerning.

Given this, there are two pathways of assessment for testimony by explanatory coherence (see Figure 2) (Thaggard, 2005). This framework supports that DNA evidence may be more readily accepted as 'consistent' when presented earlier in the trial, as competing explanations have not yet been put forward to obstruct its accessibility.

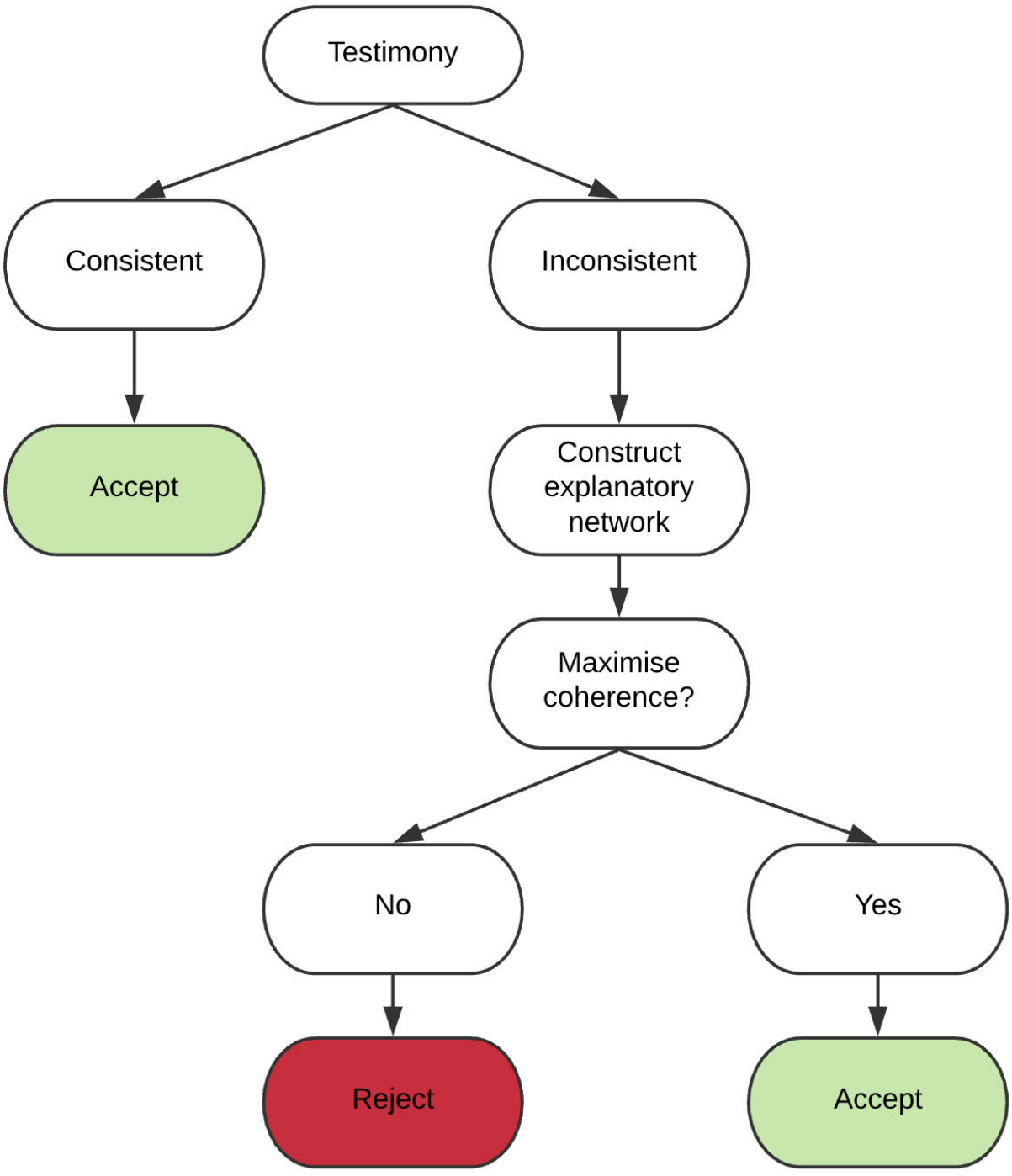


Figure 2. Pathways of testimony evaluation under explanatory coherence (Thaggard, 2005)

The Story Matters. The Story Model posits that jurors chose between competing stories (Pennington & Hastie, 1986; 1988; 1992; 1993). Pennington and Hastie (1988) highlighted the importance of the *order* of evidence presentation; when the prosecution presented their case in story-order, compared to the defence presenting in witness-order, 78% of participants returned a guilty verdict. When this was reversed (prosecution items in witness-order; defence items in story-order) the conviction rate among participants dropped to 31%. This indicates that it is not so much *what* is presented at trial, but *how* it is presented.

Aligning with explanatory coherence and coherence-based reasoning, the Story Model theorizes that 'explanatory representations' form a guiding structure for evidence integration which determines the ultimate verdict. The probative implications for verdict cannot be evaluated alone and are dependent on the interrelatedness of other evidence (Pennington & Hastie, 1993). 'Coverage' refers to how much of the evidence is explained by the story and 'coherence' is the degree to which the story is not contradictory. The more coherent a story, the more plausible and the more the plausible, the higher the confidence.

Finally, there is a body of research under the predecisional information distortion framework supporting that preexisting preferences lead to distortions of new information in favour of the initial preference (Blanchard, Carlson, & Meloy, 2014; Carlson & Russo, 2001; Russo, Medvec, & Meloy, 1996). Pretrial bias has been found to predict verdict, supporting distortion of facts (Estrada-Reynolds, Gray, & Nunez, 2015)

Therefore, inculpating DNA evidence when presented at the beginning of a trial may trigger one story representation to take the lead, and subsequent evidence to be considered in comparison, making it more prejudicial than DNA presented at the end.

Coherence-based reasoning (CBR).[i] CBR is theoretically founded in explanatory coherence and constraint satisfaction mechanisms (see Glöckner & Betsch, 2008; Simon, 2004;

Glöckner & Engel, 2013; Glöckner, Betsch, & Schindler, 2010; Holyoak & Simon 1999; Simon, Pham, Le, & Holyoak, 2001).

CBR begins with 'mental representations of a trial,' forming a multi-directional scaffold of possible connections (Simon, 2004). Empirical evidence demonstrates that reasoning is *bidirectional,* (Holyoak & Simon, 1999; Simon, 2004; Lundberg, 2004) subconscious, precedes the decision, and that *coherence-shifts* occur (Simon, Snow & Read, 2004; Simon, 2004).

Coherence-shifts have proven to be very robust in the literature (Lundberg, 2004; 2011; Holyoak & Simon, 1999; Simon et al., 2001; Simon, 2004). They refer to the process where facts with no support from the scaffold degrade and those with bidirectional support become stronger as they aid coherence. This process polarizes the evidence, resulting in a twisted representation where facts no longer reflect their actual probative weight as they have been transformed to aid coherence. Initial leanings to convict have shown stronger coherence than initial acquit leanings (Charman, Gregory, & Carlucci, 2009). This is particularly concerning when DNA forms part of a circumstantial case, as coherence-shifts may enable DNA to 'cinch' a verdict and trigger a 'spreading coherence'. This may impact memory recall by inhibiting alternative explanations, causing the threshold of 'beyond a reasonable doubt' to be reached (Holyoak & Simon, 1999; Simon, Snow, & Read, 2004).

Outside the mock-juror setting, empirical research indicates only a developing preference is needed for this distortion to emerge (Russo, Medvec, & Meloy, 1996). Committing to a verdict is a driver of distortion (Polman & Russo, 2012) and sequential presentation of information is a driver of commitment (Jonas, Schulz-Hardt, Frey, & Thelen, 2001). Blanchard, Carlson, and Guha (2011) have shown that not only is the preferred decision over weighted (proleader distortion), but the alternative decision is pushed down (antitrailer distortion). This gap increases confidence (Carlson & Russo, 2001). This is concerning as coherence-shifts may

begin to emerge before the trial even begins, particularly in the case of exposure to negative trial media (Hope, Memon, & McGeorge, 2004).

Methodological limitations with CBR in the literature. Backed by the transient nature of coherence-shifts, the literature to date has used a 'stepwise evolution' type design, whereby participants' ratings of evidence are measured at pre-test, leaning and post-test. In the pre-test 'seemingly unrelated' vignettes have been used as the dependent measure whereby this vignette is incorporated into the trial stimulus and rated again (Holyoak & Simon, 1999; Simon, 2004; Greenspan & Scurich, 2016). This assumes that the transient nature of coherence-shifts will remove any priming effects. Further, the vignettes may produce an increase in accessibility of arguments which are consistent with the anchor. Anchoring effects are robust (Furnham & Boo, 2011) and have been found even in legal experts (e.g. judges) using subtle manipulations, where defence sentencing recommendations assimilate towards those made by the prosecution (Englich, Mussweiler, & Strack, 2006). Therefore, although the literature has placed a time delay (of 2-3 weeks) between the two ratings and has successfully replicated coherence-shifts, this thesis will not use pre-test vignettes. Instead, it will examine the ratings during the trial as evidence arises, and then again before rendering a verdict to see if coherence-shifts are replicated within a trial and if they are greater for those who receive DNA evidence first.

Some studies allocated participants to one side or as arbitrators or changed facts of the stimulus materials to be either exculpating or inculpating to prompt context effects. For example, Holyoak and Simon (1999) changed evidence about the defendant's character to successfully prompt coherence-shifts and Greenspan and Scurich (2016) changed the strength of confession evidence. Fawcett (2016) studied the effects of an "ambush" alibi, finding that when the alibi is given later in trial to ambush the other party, it is perceived as less credible. These studies are not generalisable to the current study given that changing evidence adds additional inferences to the decision. For example, an 'ambush' alibi may infer that one party

was lying, whereas DNA evidence being introduced to the trial at different times does not add additional inferences. Additionally, jurors are never explicitly instructed to favour a particular side. Finally, although these studies have stimulated coherence-shifts they have not examined how the temporal process of information itself may influence coherence-shifts.

Constabile and Klein (2010) found support for recency effects, but analysis indicated that memory recall was a strong mediator for this effect. Furnham (1986) changed the order of defence and prosecution arguments and found recency effects whereby when the prosecution presented last, there was an increase in convictions, but this is not ecologically valid as it is enshrined in statute that the prosecution must set out their case first. Two studies examined the order effects of alibi and eyewitness identification and found the last piece of evidence had greater impact on assessment of guilt (Dahl, Brimacombe, & Lindsay, 2009; Price & Dahl, 2013). However, both of these studies were in the context of a mock-police investigation, not a mock-juror setting. Also, the evidence presented either strongly implicated or exculpated the defendant, making the effects hard to unwind in terms of potential contrast effects. Charman, Carbone, Kekessie, and Villalba (2016) corrected this by examining whether ambiguous evidence could retrospectively change the evaluation of previous evidence and considered rumination and order effects as mediators. While rumination had no effect on verdict, they did observe recency effects. However, probability of guilt and dichotomous verdict were used as the dependent measure to distinguish evaluation from integration.

The current study improves the methodology employed above by taking a second rating of the manipulated evidence as a dependent measure compared to probability of guilt and verdict. Therefore, the impact on DNA evidence at different time points can be distinguished from global assessments of guilt.

Davis, Tindale, Nagao, Hinsz, and Robertson (1984) examined order effects on deliberation of three different charges – reckless homicide, aggravated battery and criminal

damage to property. Participants received the information in different orders with criminal damage to property remaining in the middle. Therefore, it was either preceded by a more serious or less serious charge. Conviction of earlier charges inflated later conviction. This was explained by the contrast between charges, but the results could also be explained under CBR where prejudicial evidence given early has more impact. This finding would also be compatible with primacy effects postulated by Pennington and Hastie (1982).

Finally, the assimilation hypothesis has recently received empirical support in a study demonstrating initial knowledge of DNA evidence inflates subsequent evidence (Rassin, 2017). However, the stimulus materials used only comprised short vignettes of evidence and lack ecological validity. This is consistent with the finding that confession evidence inflates ratings of handwriting evidence (Kukucka & Kassin, 2014).

Therefore, CBR supports that early presentation of strong inculpating evidence may initiate beliefs of guilt, and trigger spreading coherence. Ambiguous evidence may be bolstered to cohere with early beliefs. Further, all subsequent evidence will be compared to the evidence which entered the mental representation early on. If this early evidence is strong, it may afford 'relevance' to facts supporting it thereby inhibiting other explanations. This fits with explanatory coherence as discussed under 1.1.

Confidence. Confidence in verdict is often unrelated to the accuracy of DNA knowledge or comprehension (Goodman-Delahunty & Hewson, 2010). Studies have found proleader and antitrailer distortion increased confidence regardless of verdict (Simon, Snow, & Read, 2004). The Story Model postulates the fit between story and verdict determines confidence. Therefore, if one story has been bolstered it will fit better with that verdict category.

Debiasing attempts. In the mock-juror setting, explicitly warning people upfront about the 'continued influence effect' concerning inadmissible evidence has reduced juror reliance on misinformation (Ecker, Lewandowsky, & Tang, 2010). However, the assertion that hypothesis

disconfirmation mitigates coherence-shifts (Simon, 2004) has been withdrawn (Simon, 2012). No effective way to debias coherence-shifts in the juror realm have been successful.

Attempts to debias coherence-shifts outside the mock-juror setting have had mixed results. Raising expectations of accountability in front of others has been found to enhance integrative thinking and debias facts (Anderson & Sechler, 1986; Colombo, 2018). Delaying the formation of an emerging conclusion (Charman, 2013), and consideration of 'component lists' prior to choice, has mitigated coherence-shifting and over-confidence (Carlson & Pearo, 2004). Studies using hypothesis disconfirmation have been mostly unsuccessful with some eliciting a detrimental response (Colombo, 2018; Walmsley & Gilbey, 2017).

1.3. The problem with judicial instructions – a little too *late*?

Instructions are typically given post-trial. They are directions from the judge to the jury about how the jury should evaluate a case according to the law. Instructions aim to ensure jurors do not reason in a prejudicial way and will always be given when DNA forms part of a circumstantial case. The judiciary's faith in instructions as gate-keeping mechanisms capable of mitigating the prejudice of DNA was reaffirmed by the majority in the High Court case of *Aytugrul v The Queen* [2012] HCA 15 (Heydon J).

Yet, psychological research has long indicated that post-trial instructions are ineffective (Kassin & Wrightsman, 1979; Ogloff & Rose, 2005) even when aimed at combating misleading testimony (McQuiston-Surrett & Saks, 2009) or target comprehension of DNA (Dartnall & Goodman-Delahunty, 2006). One study cited negligible difference between a group receiving no instruction and a post-trial instruction (63% and 59% conviction rate respectively) compared to pre-trial instruction (37% conviction rate) (Kassin & Wrightsman, 1979). Even judicial instruction clearly stating the prosecution expert was unjustified did not impact verdict, and surprisingly increased conviction rate compared to a control receiving no instruction

(Eastwood & Caldwell, 2015). Further, jurors report confidence where objective ability to follow instructions is inadequate (Ogloff & Rose, 2005).

In contrast, pre-trial instructions significantly improve juror's ability to remember and distinguish probative from non-probative evidence (ForsterLee, Horowitz, & Bourgeois, 1993), by helping to provide jurors with a legally appropriate scaffold to evaluate and integrate evidence throughout the trial. Trabasso and Sperry (1985) examined importance of story events, finding that if an event was in the causal chain from the beginning of the story until the end it was a predictor of judgement of importance. This fits with the Story Model as story structure provide an 'index' for the importance of different pieces of evidence (Pennington & Hastie, 1993). Pre-trial instruction can set the story structure according to the law and not jurors' preconceived notions, potentially improving juror evaluation of cases.

Pre-trial instructions also improve juror's ability to apply the law to the facts even while simultaneously using "exemplars" of crime in their evaluation and reduce prejudice against the defendant's characteristics (Smith, 1991; Heuer & Penrod, 1989). Studies showing no effect on 'quality of verdict', (Smith, 1991) may be due to the fact that assessing jurors' verdicts as 'right' or 'wrong' objectively is hard and requires methodological changes to do so (ForsterLee, Horowitz & Bourgeois, 1993), which has been neglected for the most part in the literature. Elwork, Sales, and Alfini (1977) found moderate improvement in memory recall of probative evidence when instruction was given pre *and* post-trial. Other studies, despite stating 'trends towards significance' do not provide significant statistical support for the effectiveness of pre-trial instruction on improving assessments of probative weight (Wheatcroft & Keogan, 2017; Rowe, 1997). Although Rowe (1997) did not find a significant effect for participants given an instruction, the study was underpowered. Further, Wheatcroft and Keogan's (2017) instruction only consisted of a verbal warning by the researcher to "be aware of the danger of placing disproportionate weight on the evidence, as opposed to your general impression and assessment

of it". This takes away the cloak (i.e. the judge) which an instruction typically wears and may have a different impact than if the instruction is embodied as part of the trial stimulus materials.

Some studies have explained the failure of instruction to have an impact in terms of increased cognitive load (Wheatcroft & Keogan, 2017) but this seems overly simplistic. The trial process itself is cognitively demanding, and it does not explain the failure of even simple instructions to have an effect, nor the success of pre-trial instruction. Others have speculated that jurors perceive a lack of 'specialty' on behalf of the judge (Eastwood & Caldwell, 2015). A large portion of the literature has suggested post-trial instructions inhibit memory of trial facts compared to pre-instructed/non-instructed jurors (Kassin & Wrightsman, 1979), but these results are mixed, with some finding no difference in recall between groups (Hastie, Penrod, & Pennington, 1983; Elwork, Sales & Alfini, 1977; Heuer & Penrod, 1989).

Elwork et al., (1977) suggested that pre-trial instruction attuned jurors to probative pieces of evidence and found pre-instructed jurors were more likely to remember evidence in favour of the defence. Others have suggested that pre-trial instruction raises skepticism from the outset, changing the evaluation of evidence (Kassin & Wrightsman, 1979). This idea of increased skepticism fits with the theory of explanatory coherence (Thaggard, 1989). Increased skepticism triggered by pre-trial instruction would also act to delay commitment to a decision – a driver of distortion in the literature – and enable better evaluation of DNA evidence (Smith, 1991; Polman & Russo, 2012).

However, studies examining instructions have not tracked the ratings of evidence so there is no way to know if post-trial instruction *inhibits* memory, or if jurors are not attuned to legally relevant facts, or if that information is simply disregarded/downgraded at the time of evaluation when it does not maximize coherence.

It is likely jurors are unable to re-evaluate polarized evidence retrospectively when post-trial instruction is given, as at this point they enter into story classification (Pennington & Hastie, 1993).

Studies have focused on simplifying instructions to increase comprehension (Baguley, McKimmie, & Masser, 2017; Elwork, Sales, & Alfini, 1977), instruction on standards of proof, right to silence, witness reliability and other specifics about applying law to the facts, but this is wasted if the instruction is delivered too late. Previously the effect of instruction on verdict has been the focus; in contrast, this thesis will focus on the weight attributed to the DNA evidence.

Finally, this study will include a manipulation check with the instruction, as it is not clear from past studies if manipulation checks directly targeting the instruction have been included (Eastwood & Caldwell, 2015; Kassin & Wrightsman, 1979).

1.4. The current study.

Research should aim to help jurors objectively evaluate trial facts. Distinguishing probative from non-probative facts and appropriate attribution of weight to evidence is central to a fair trial. Therefore, this study will focus on the weightings of DNA evidence and its impact on the surrounding evidence. If jurors are influenced by temporal effects, it is important to ensure one side is not prejudiced by virtue of presentation order. Prejudicial evidence (like DNA) could be presented at a later point in the trial, which may enable more accurate assignments of probative weight.

Finally, debiasing attempts by opposing expert testimony and cross-examination have had mixed success. If effective, pre-trial instruction has the capacity to be easily adopted (Heuer & Penrod, 1989), and adapted to other evidence at no cost to either party.

Table 1

Aims and hypothesis

Research questions	Hypotheses
One: What effect does timing of DNA evidence in a criminal trial have on mock-juror evaluation of evidence?	**1.** DNA given at the beginning will be attributed more weight and will gain further weight throughout the trial than DNA given at the end of the trial. **2.** DNA evidence at the beginning of the trial will inflate subsequent evidence. **3.** Participants will be more confident in their decision when the DNA is presented at the beginning of the trial.
Two: Can a judicial instruction given earlier in the trial mitigate coherence-shifts better than a judicial instruction given at the end of the trial?	**4.** When the judicial instruction is given at the beginning of the trial, there will be less coherence-shifting, and the DNA will not be weighted as favourably as when the judicial instruction is given at the end of the trial.
Aims one and two	**5.** Participants in the condition where the DNA is presented first, and the judicial instruction is presented last, will be more likely to vote guilty.

Note. Preregistered on the Open Science Framework at DOI 10.17605/OSF.IO/J3FBK.

Chapter 2: Method

2.1 Ethics

The study was approved by the *University of Adelaide Human Research Ethics Subcommittee,* approval number *18/16* (Appendix A). Participation was voluntary and anonymous. Informed consent was obtained from participants prior to their commencement in the study, indicated by ticking 'yes' at the bottom of the information preamble page (Appendix D). Identification numbers used for students participating in exchange for course credit assured anonymity. Participants were informed that their data would be made available on the Open Science Framework in a de-identified, quantitative form.

2.2 Participants

A total of N = 257 participants completed the study, after 50 responses were removed for failing manipulation checks. A recent study has provided support for the use of student populations in mock-juror research (Bornstein et al., 2017). Therefore, level one psychology students at the University of Adelaide (N=99) were recruited via the Research Participation Pool and completed the study in exchange for course credit. The study also comprised members of the public recruited by posters on the University of Adelaide campus and word of mouth. Members of the public were able to go in the draw to win 1 of 3 $50 David Jones vouchers in exchange for their participation. All participants were current Australian residents.

This study did not conform to the eligibility criteria for jurors in SA (set out in the *Juries Act* 1927). 2 participants were aged under the age of 18. 2 participants were aged 70 or over. Only 2 participants (law clerk + judicial assistant) identified an occupation within the legal practice excluded under *the Juries Act* 1927. Participants were not screened for past criminal convictions. If participants completed the survey and passed all manipulation checks it was assumed that they had the requisite fluency of the English language.

This inclusion criteria were due to the time and resource constrains on the study. Further, while some studies indicate judges are more conservative in their reasoning (McQuiston-Surrett & Saks, 2009), research indicates judges and lawyers fall prey to the same types of fallacious reasoning as lay people (Guthrie, Rachlinski & Wistrick, 2001).

An *a priori* power analysis indicated 84 participants were needed to obtain 80% power and a medium effect size of .3 at an alpha level of .05. Given the mixed results in the literature, different methodology in the current study and unknown effect size, the study aimed for 30 participants in each group (n = 240 total).

2.3 Materials

The study was completed online via Survey Monkey™. Participants first opened a survey which then randomly assigned them to 1 of 8 experimental conditions (see Appendix C). The study survey consisted of a transcript written by the researcher (see Appendix E) of approximately 5,000 words. The stimulus materials used are available on the Open Science Framework. The facts of the transcript (titled *Bayar v The Queen* for the purposes of the experiment) were taken from the Australian case of *Aytugrul v The Queen* [2012] HCA 15. Certain facts were omitted/changed to fit the time limitations of the study. A criminal case was chosen for this study, as it is typically criminal cases where juries are employed in Australia. A 10-point Likert scale was used to rate the evidence at both time points. A Likert scale was used as it has been rated as easy to use by participants, has acceptable validity and reliability, and is consistent with previous literature (Preston & Colman, 2000; Holyoak & Simon, 1999).

Dependent measures included ratings of all evidence at two time points, dichotomous verdict, probability of guilt (0-100) and confidence on a 10-point Likert scale. In addition, a multiple-choice question assessed participants understanding of the DNA evidence (specifically we were interested to see if the judicial instruction had been effective in combating the prosecutor's fallacy) at the end of the trial (see question 42 Appendix E). Participants were

invited to make optional comments about their reasoning at the end of the study. There were four manipulation checks with the study; one accompanied the judicial instruction and another three questions concerning the beginning, middle and end of the trial.

2.4 Procedure

Participants accessed the study via a Survey Monkey™ web address and were randomly allocated to one of eight conditions, which then linked them to a second survey containing the study.

Participants read the participant preamble and gave informed consent. Participants were blinded and did not know the purpose of the study; rather they were informed broadly that it was aimed to assess the ways in which jurors decide during a complex trial. Participants first answered demographic questions, including age, gender, occupation and educational history before reading the transcript. The stimulus materials were organized so each piece of evidence was considered and rated separately. Throughout the transcript participants rated each piece of evidence on a 10-point Likert scale from -5 (highly indicative of innocence) to +5 (highly indicative of guilt). There were four manipulation checks imbedded in the study. The questions were multiple-choice and concerned the beginning, middle and end of the transcript. A further manipulation check was included with the judicial instruction to ensure participants active engagement with the instruction; where participants were asked to list three reasons why DNA may not indicate guilt. Participants then re-rated the same pieces of evidence on the same scale before rendering a dichotomous verdict, rating their confidence in their decision and giving a probability of guilt. A final optional question invited participants to write something about their reasoning.

Participants had the option to receive a summary statement at the culmination of data collection, provided that an email address had been supplied (see Appendix H) and to go in the draw to win a voucher.

Design statement. This study used a between subjects experimental 2 (pre-trial and post-trial instruction) * 2 (prosecution DNA beginning/end of prosecution case) *2 (defence DNA beginning/end of defence case) design. The resulting experimental conditions are outlined in Figure 3 below.

There is no control, as all comparisons are made between groups. This is justified as the design considers important ecological and external validity points. Firstly, a judge always instructs a jury on DNA evidence and the focus of the current study is the difference between pre vs. post-trial instruction. Therefore, a condition where participants are not instructed would be inapplicable. Secondly, DNA evidence is central to the research question, concerning the *temporal* effect of the DNA, not the effect of DNA itself. Therefore, the research questions and hypotheses do not require a baseline as it would be unsurprising to find a lower conviction rate in a circumstantial case without DNA evidence.

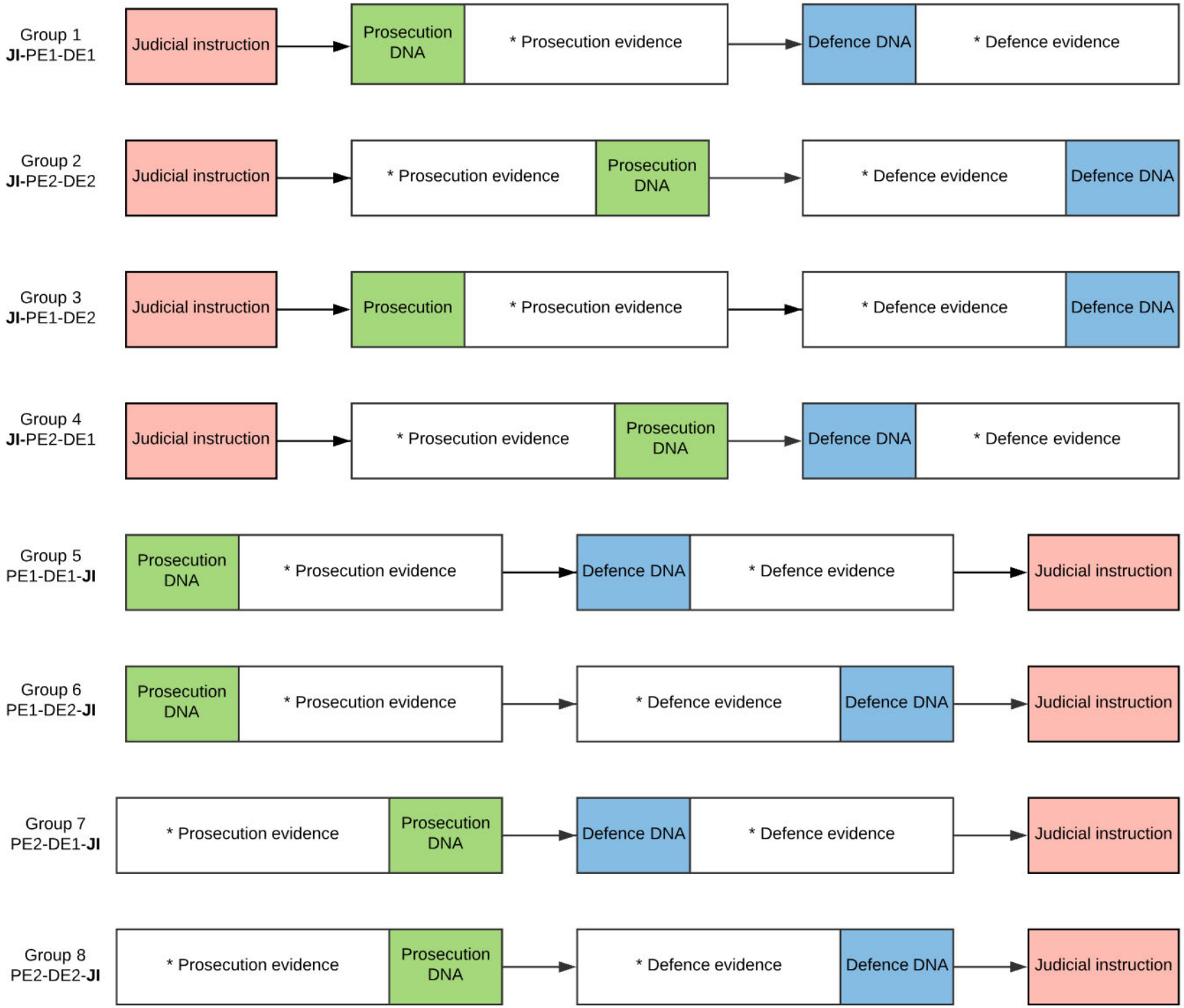


Figure 3. Experimental conditions

Chapter 3: Results

3.1 Data screening and quality control

The experiment was conducted online, and it was necessary to ensure participants engaged with the materials. As outlined in Chapter 2, four manipulation checks were used. On inspection of the data, it was apparent that despite engaging with the task (evidenced by participants' open-ended response) 19 participants failed only the first multiple-choice manipulation (see question in Appendix E). This question asked participants to distinguish between the statistical expression of the DNA match given by the prosecution and the defence. Participants struggled to understand/discern between the two opposing expert testimonies despite engaging with the task. We decided to follow through with a strict exclusion criterion as stipulated in pre-registration. Exclusion of this data also ensures that any effect in the data is not due to misunderstanding of DNA evidence. 257 respondents remained after a further 31 responses were excluded due to attrition and failing manipulation checks.

3.2. *A priori* comparisons

A priori comparisons were pre-registered on the Open Science Framework prior to data collection; however, their use was dependent upon the data being normally distributed. The coherence-shift statistic was calculated by taking the difference between the two ratings (second rating before rendering a verdict – initial rating of the evidence) to give a change score for the two pieces of DNA evidence. A negative score on coherence-shift means that the participants' rating shifted to favour the defence, and a positive change score means that the participants' rating shifted to favour the prosecution. A score of 0 means that the rating did not change.

Hypothesis 1. This hypothesis predicted that DNA given at the beginning would be attributed more weight and would gain further weight throughout the trial than DNA given at the end of the trial. The different aspects of this hypothesis are considered in turn below.

Histograms and QQ plots confirmed that the data was not normal, but homogeneity of variance was acceptable (see Appendix F). Therefore, we used non-parametric tests to carry out *a priori* comparisons. As hypotheses are directional, the one-tailed exact statistic is reported.

Initial ratings of prosecution DNA evidence

The hypothesis that DNA given at the beginning would be attributed more weight was unsupported. Group 2 (*Mdn*=37.78) who received the DNA last had significantly higher ratings than Group 3 (*Mdn*=28.95), receiving the DNA first *Z*=-1.932, *p*=.027. Likewise, Group 8 (*Mdn*=38.18) who received the DNA evidence last had significantly higher ratings than Group 6 (*Mdn*=27.32), receiving DNA evidence first, *Z*=-2.410, *p*=.008. Group 7 (*Mdn*=38.45), also received the DNA last and had higher ratings than Group 5 (*Mdn*=26.91), Z=-2.605, *p*=.004. There was no significant difference between Group 1 (*Mdn*=30.93) and Group 4 (*Mdn*=32.20), Z=-.297, *p*=.386.

A boxplot in Figure 4 gives a visual illustration of the initial rating across groups confirming the trend for those groups receiving the prosecution DNA evidence last (indicated by 'P2') to have higher ratings.[ii]

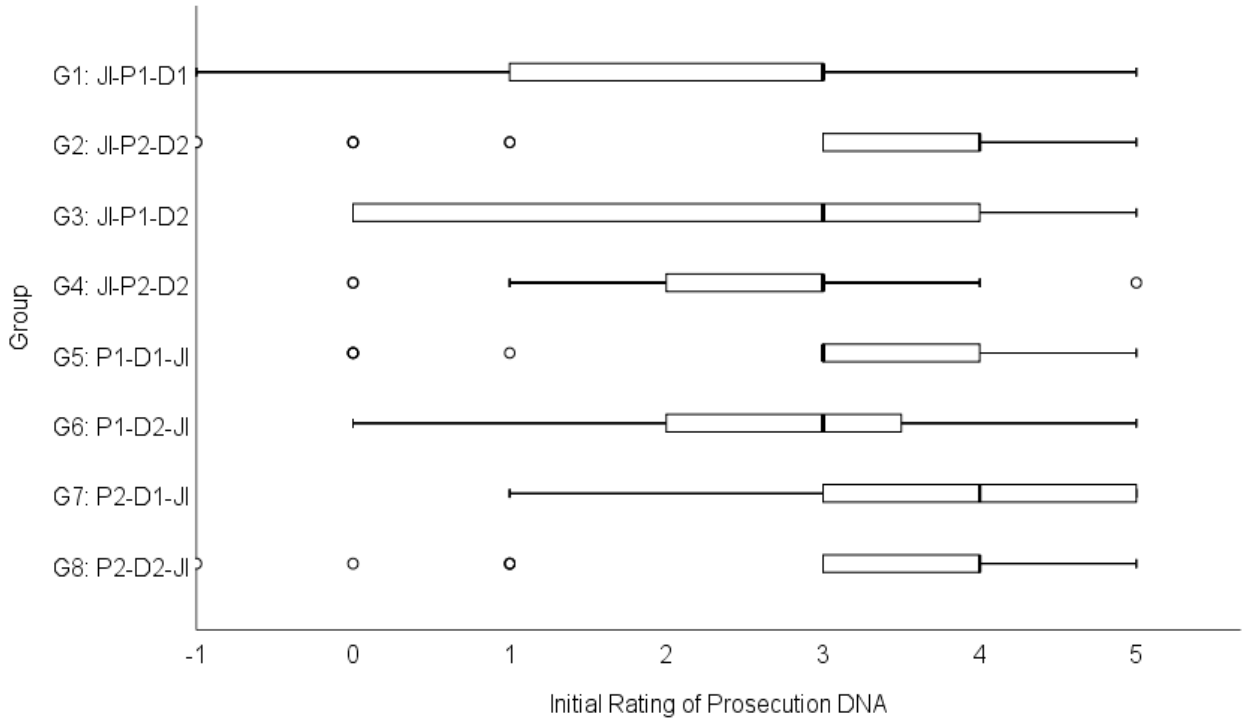


Figure 4. Boxplot showing initial rating of prosecution DNA evidence across conditions.

-5 = highly indicative of innocence, 0 = Neutral, 5 = highly indicative of guilt.

Coherence-shift of prosecution DNA evidence

The hypothesis that DNA given at the beginning would *gain* more weight throughout the trial than DNA given last was partially supported. As predicted Group 1 (*Mdn*=36.28) who received the DNA first had a greater coherence-shift than Group 4 (*Mdn*=25.70), *Z*=-2.464, *p*=.007. Similarly, Group 6 (*Mdn*=37.35) - who received the DNA first - shifted more than Group 8 (*Mdn*=29.03) who received the DNA last, Z=-1.892, *p*=.029 and Group 5 (*Mdn*=37.88) – who received the DNA first - had a greater shift than Group 7 (*Mdn*=26.77) receiving the DNA last, *Z*=-2.463, *p*=.007. There was no difference between Group 2 (*Mdn*=36.44) and Group 3 (*Mdn*=30.28), *Z*=-1.356, *p*=.089. Therefore, these results provide partial support for the Hypothesis 1 that DNA given earlier in the trial would shift more in weight than DNA given at the end.

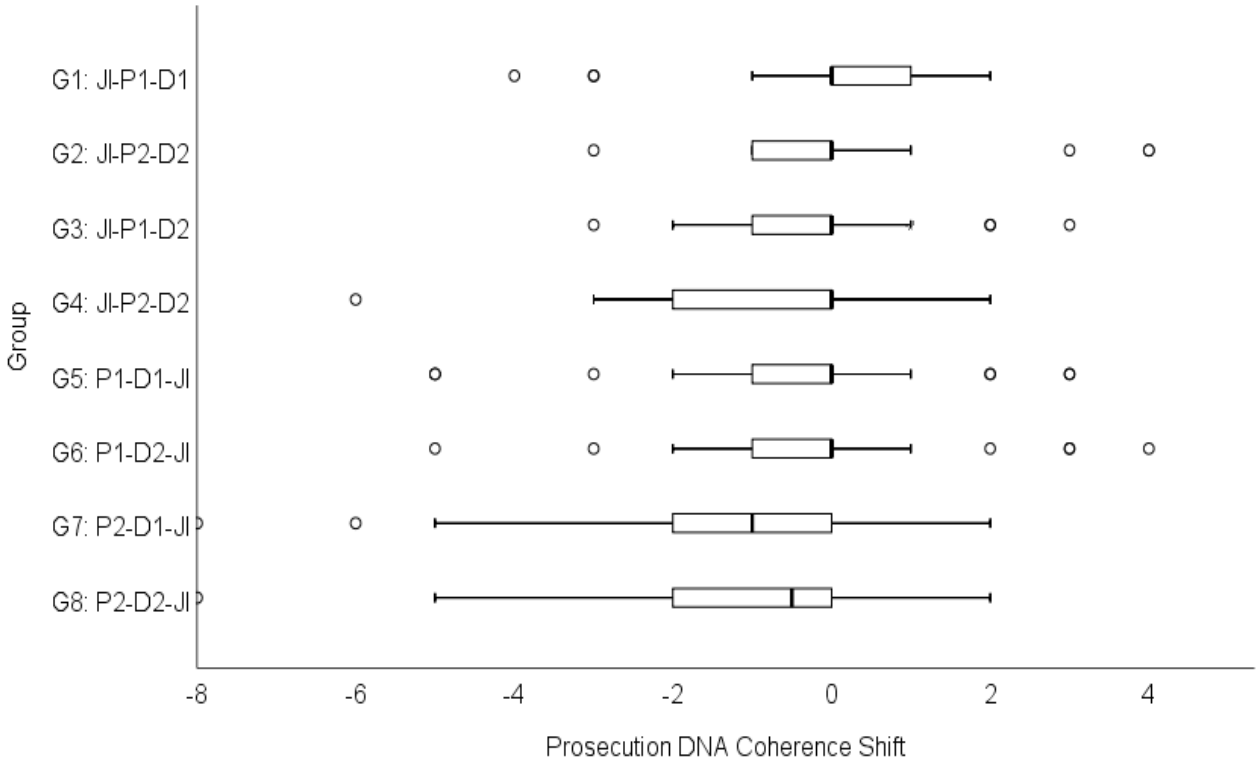


Figure 5. Boxplot showing prosecution DNA coherence-shift across conditions

Initial ratings of defence DNA evidence

It was predicted that the DNA evidence would be attributed more *probative* weight; therefore, the measure in context of the defence DNA, means lower ratings were more probative as they shifted towards "highly indicative of innocence." Comparisons on initial ratings of defence DNA did not support Hypothesis 1.

Group 2 (*Mdn*= 27.12) - who received the DNA last - rated the DNA evidence more persuasive than Group 4 receiving it first (*Mdn*=36.82), *Z*=-2.136, *p*=.016. Similarly, Group 8 (*Mdn*=28.94), receiving the DNA second rated it higher than Group 7 (*Mdn*=37.45) receiving the DNA first, *Z*=-1.856, *p*=.032. There was no significant difference between Group 5 (*Mdn*=31.15) and Group 6 (*Mdn*=33.94), *Z*=-.611, *p*=.27 or Group 1 (*Mdn*=35.24) and Group 3 (*Mdn*=31.66) *Z*=-.776, *p*=.222. The boxplot below illustrates the trend for DNA evidence given last to be rated more highly.

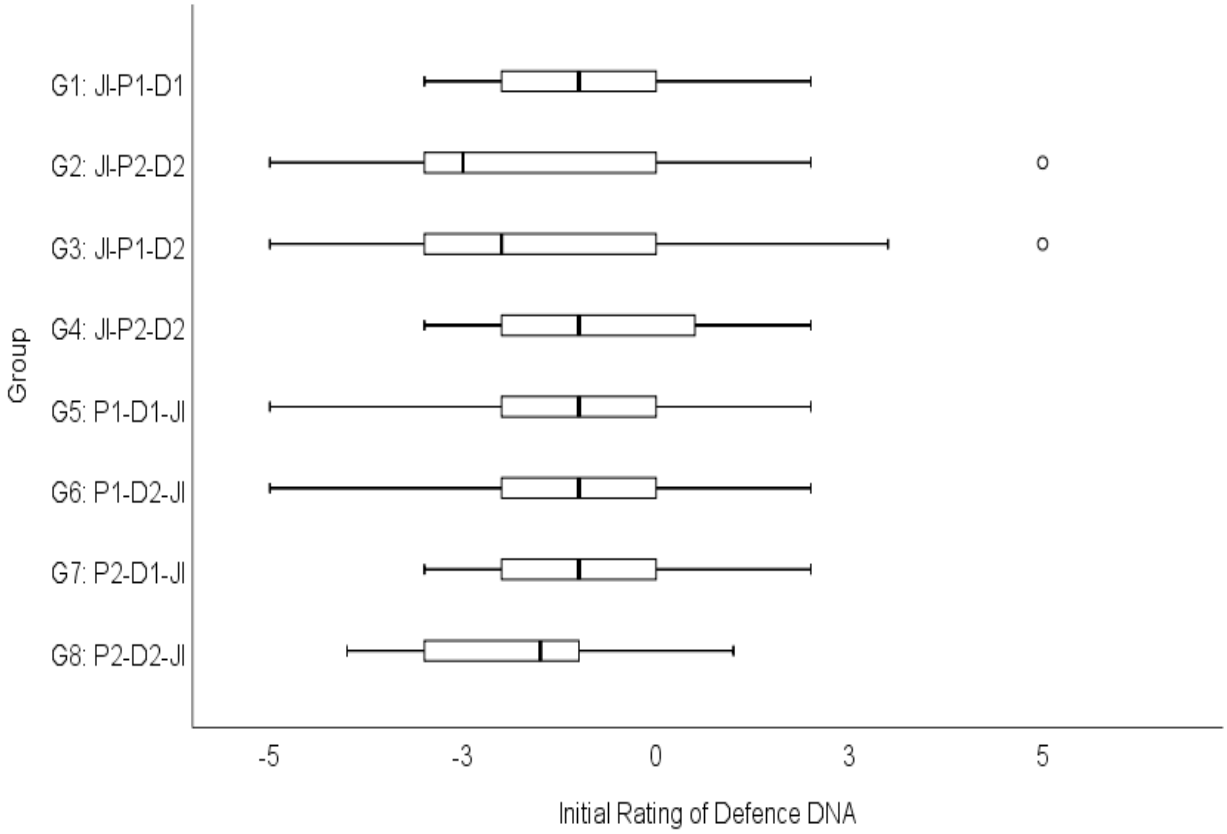


Figure 6. Boxplot showing initial ratings of defence DNA across groups.

– 5 = highly indicative of innocence, 0 = neutral, +5 = highly indicative of guilt.

Coherence-shift of defence DNA evidence

It was predicted that early presentation of defence DNA would increase coherence-shift. This was not supported. There was no difference for comparisons on timing and coherence-shift. There was no significant difference between Group 2 (*Mdn*=32.69) and Group 4 (*Mdn*=30.05), *Z*=-.608, *p*=.274, Group 1 (*Mdn*=31.99) and Group 3 (*Mdn*=35.11), *Z*=-.698, p=.245, Group 5 (*Mdn*=31.55) and Group 6 (*Mdn*=33.73), *Z*=-.524, *p*=.303 or Group 7 (*Mdn*=29.47) and Group 8 (*Mdn*=36.22), *Z*=-1.539, *p*=.063.

Hypothesis 2. The hypothesis that DNA evidence would inflate subsequent evidence was not supported. The data was not normally distributed however, homogeneity of variance was not violated (Appendix H). We used non-parametric tests to carry out *a priori* comparisons.

Prosecution inflation of the "neighbour's testimony"

To examine the effect of prosecution DNA evidence, comparisons were made on the neighbor's testimony, which was either the first piece of evidence if the DNA was presented last or the second piece of evidence if it followed DNA evidence. Comparisons showed no significant difference between ratings of testimony following or not following DNA. Results of the Mann Whitney U tests used compare conditions are presented in Table 2 below.

Table 2

A priori comparisons testing prosecution DNA inflation of subsequent evidence using Mann Whitney U tests

Comparison[iii]	Groups compared	"Neighbours testimony" ratings
2	2(JI-**P2**-D2)	2(*Mdn*=34.29) and 3(*Mdn*=32.66)
	3(JI-**P1**-D2)	*Z*=-515, *p*=.368
3	1(JI-**P1**-D1)	1(*Mdn*=31.94) and 4(*Mdn*=30.96)
	4(JI-**P2**-D1)	*Z*=-.308, *p*=.357
6	6(**P1**-D2-JI)	6(*Mdn*=34.94) and 8(*Mdn*=31.24)
	8(**P2**-D2-JI)	*Z*=-1.042, *p*=.140
7	5(**P1**-D1-JI)	5(*Mdn*=33.21) and 7(*Mdn*=31.74)
	7(**P2**-D1-JI)	*Z*=-.416, *p*=.369

Note. See Appendix E for full stimulus materials. * in Figure 3 Chapter 2 indicates neighbours' testimony. Bold font = manipulation compared. P1 = DNA first, proceeding neighbour's testimony. P2 = DNA last, not proceeding neighbour's testimony.

Defence inflation of the "mechanics testimony"

Comparisons were made on the mechanics testimony, which was the first piece of evidence presented by the defence, or the second piece of evidence following the DNA. Comparisons showed no significant difference between ratings of mechanics testimony following or not following defence DNA. Results of Mann Whiney U tests are presented in Table 3 below.

Table 3

A priori comparisons testing defence DNA inflation of subsequent evidence using Mann Whitney U

Comparison	Groups compared	"Neighbours testimony" ratings
1	2(JI-P2-**D2**)	2(*Mdn*=30.56)
	4(JI-P2-**D1**)	4 (*Mdn*=32.64)
		Z=-.469, *p*=.323
4	1(JI-P1-**D1**)	1(*Mdn*=34.01)
	3(JI-P1-**D2**)	3(*Mdn*=32.95)
		Z=-.250, *p*=.396
5	7(P2-**D1**-JI)	7(*Mdn*=34.55)
	8(P2-**D2**-JI)	8(*Mdn*=31.59)
		Z=-.726, *p*=.236
8	5(P1-**D1**-JI)	5(29.92)
	6(P1-**D2**-JI)	6(35.24)
		Z=-1.246, *p*=.109

Note. Bold font = manipulation. D1 = DNA first, proceeding mechanics testimony. P2 = DNA last, not proceeding mechanics testimony.

Hypothesis 3. This hypothesis predicted that pre-trial instruction would reduce coherence-shifting and lower initial ratings of DNA evidence compared to groups receiving post-trial instruction. The different aspects of the hypothesis are considered in turn below.

Initial ratings of prosecution DNA

Comparisons on initial ratings for prosecution DNA evidence and pre-trial vs. post-trial instruction provided partial support for Hypothesis 3. Group 1 (*Mdn*=30.09) – who received a pre-trial instruction - had lower ratings than Group 5 (*Mdn*=38.03), *Z*=-1.769, *p*=.039. Group 4 (*Mdn*=20.88) with pre-trial instruction also had lower ratings than comparison Group 7 (*Mdn*=38.24), *Z*=-4.088, *p*=.000. Despite this, there was no significant difference between Group 2 (*Mdn*=32.79), and Group 8 (*Mdn*=36.21), *Z*=-.742, *p*=.231 or Group 3 (*Mdn*=30.75) and Group 6 (*Mdn*=33.29), *Z*=-.573, *p*=.287. Figure 5 mentioned above under Hypothesis 1 shows the shift in initial ratings across conditions.

Initial ratings of defence DNA

There was no significant difference for pre vs. post-trial instruction on initial ratings of defence DNA evidence. Results are presented in Table 4 below.

Table 4

A priori comparisons for the effect of pre vs. post-trial instruction on initial ratings of DNA evidence using Mann Whitney U tests

Comparison	Groups compared	Initial ratings	
		Defence	Prosecution
9	1 (**JI**-PE1-DE1)	1(*Mdn*=34.56) and 5(*Mdn*=33.42)	1(*Mdn*=30.09) and 5(*Mdn*=38.03)
	5* (PE1-DE1-**JI**)	*Z*=-.245, *p*=.405	*Z*=-1.769, *p*=.039
10	2 (**JI**-PE2-DE2)	2(*Mdn*=32.15) and 8(*Mdn*=36.85)	2(*Mdn*=32.79) and 8(*Mdn*=36.21)
	8* (PE2-DE2-**JI**)	*Z*=-1.000, *p*=.160	*Z*=-.742, *p*=.231
11	3(**JI**-PE1-DE2)	3(*Mdn*=30.38) and 6(*Mdn*=33.68)	3(*Mdn*=30.75) and 6(*Mdn*=33.29)
	6* (PE1-DE2-**JI**)	*Z*=-.729, *p*=.235	*Z*=-.573, *p*=.287
12	4 (**JI**-PE2-DE1)	4(*Mdn*=29.54) and 7(*Mdn*=30.42)	4(*Mdn*=20.88) and 7(*Mdn*=38.24)
	7* (PE2-DE1-**JI**)	*Z*=-.201, *p*=.423	*Z*=-4.088, *p*=.000

Note. * indicates the group receiving post-trial instruction, expected to have higher initial ratings. As outlined above, the measure in context of defence DNA means lower ratings were more probative as they shifted towards “highly indicative of innocence.”

Prosecution DNA coherence-shift

Comparisons on pre vs. post-trial instruction on coherence-shift for prosecution DNA did not provide support for Hypothesis 3. When the instruction was given pre-trial in Group 2 (*Mdn*=39.76), ratings of the prosecution DNA shifted more than ratings in Group 8 (*Mdn*= 29.24) receiving post-trial instruction, *Z*=-2.333, *p*=.010. Likewise, Group 4 (*Mdn*=33.84) who received the pre-trial instruction, shifted more than Group 7 (*Mdn*=26.53), *Z*=-1.701, *p*=.045. There was no significant difference between Group 1 (*Mdn*=36.99) and Group 5 (*Mdn*=30.92), *Z*=-1.345, *p*=.091 or between Group 3 (*Mdn*=29.92) and Group 6 (*Mdn*=34.15), *Z*=-.968, *p*=.169. These comparisons indicate a trend for pre-trial instruction to cause a greater coherence-shift.

Defence DNA coherence-shift

There were no significant differences in coherence-shift on the defence DNA evidence between pre-trial conditions and post-trial conditions. The results are presented in Table 5.

Table 5

The effect of pre vs. post-trial instruction on coherence-shift of DNA evidence using Mann Whitney U tests

Comparison	Groups compared	Coherence-shift on DNA	
		Defence	Prosecution
9	1 (JI-PE1-DE1)	1(*Mdn*=31.38) and 5(*Mdn*=36.70)	1(*Mdn*=36.99) and 5(*Mdn*=30.92)
	5* (PE1-DE1-JI)	*Z*=-1.167, *p*=.124	*Z*=-1.345, *p*=.091
10	2 (JI-PE2-DE2)	2(*Mdn*=33.31) and 8(*Mdn*=35.69)	2(*Mdn*=39.76) and 8(*Mdn*=29.24)
	8* (PE2-DE2-JI)	*Z*=-.537, *p*=.299	*Z*=-2.333, *p*=.010
11	3(JI-PE1-DE2)	3(*Mdn*=29.95) and 6(*Mdn*=34.11)	3(*Mdn*=29.92) and 6(*Mdn*=34.15)
	6 * (PE1-DE2-JI)	*Z*=-.930, *p*=.178	*Z*=-.968, *p*=.169
12	4 (JI-PE2-DE1)	4(*Mdn*=30.82) and 7(*Mdn*=29.26)	4(*Mdn*=33.84) and 7(*Mdn*=26.53)
	7* (PE2-DE1-JI)	*Z*=-.367, *p*=.360	*Z*=-1.701, *p*=.045

Note. * indicates the group receiving post-trial instruction, predicted to have the greater coherence-shift.

Hypothesis 4. The hypothesis that DNA evidence presented earlier in the trial would be associated with an increase in confidence was unsupported. The data was negatively skewed, but homogeneity of variance was not violated (Appendix H). Therefore, we used non-parametric tests to carry out *a priori* comparisons.

Group 8 (*Mdn*=37.24), received the defence DNA last and had higher confidence ratings than Group 7 (*Mdn*=28.35) and *Z*=-1.922, *p*=.027. There was no significant difference in other comparisons. They are presented in Table 6.

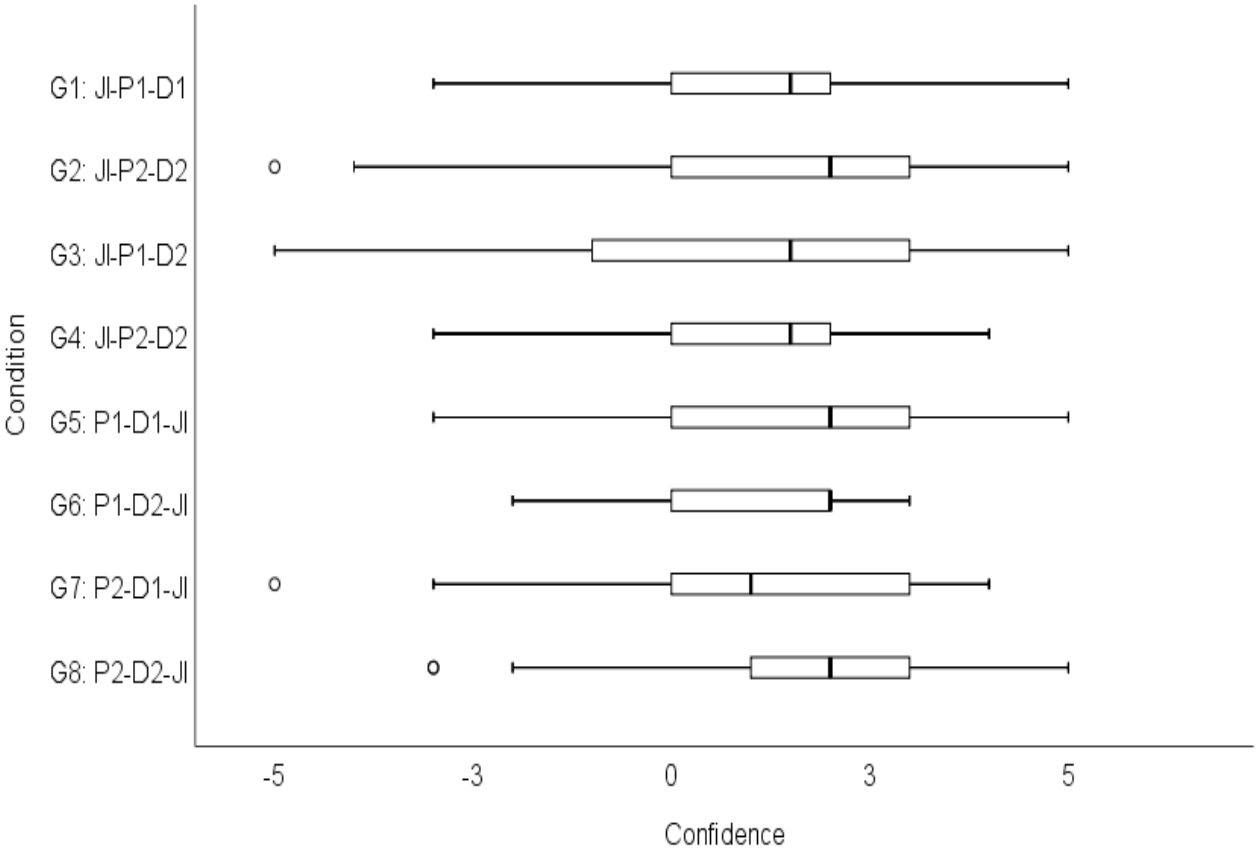


Figure 7. Boxplot displaying confidence ratings in verdict across conditions.

-5 = ‘not at all confident’, 0 = 'not sure' and 5 = 'very confident’

Table 6

A priori comparisons using Mann-Whitney U tests for confidence ratings across conditions

Comparison	Groups compared	Confidence rating
1	2 (JI-P2-**D2**) (*Mdn*=33.53)	*Z*=-.995, *p*=.162
	4 (JI-P2-**D1**)* (*Mdn*=29.04)	
2	2(JI-**P2**-D2) (*Mdn*=35.31)	*Z*=-.798, *p*=.215
	3(JI-**P1**-D2)* (*Mdn*=31.58)	
3	1(JI-**P1**-D1)* (*Mdn*=31.90)	*Z*=-.196, *p*=.424
	4(JI-**P2**-D1) (*Mdn*=31.02)	
4	1(JI-P1-**D1**)* (*Mdn*=33.85)	*Z*=-.156, *p*=.440
	3(JI-P1-**D2**) (*Mdn*=33.13)	
5	7(P2-**D1**-JI)* (*Mdn*=28.35)	*Z*=-1.922, *p*=.027
	8(P2-**D2**-JI) (*Mdn*=37.24)	
6	6(**P1**-D2-JI)* (*Mdn*=29.32)	*Z*=-1.527, *p*=.064
	8(**P2**-D2-JI) (*Mdn*=36.35)	
7	5(**P1**-D1-JI)* (*Mdn*=35.45)	*Z*=-1.328, *p*=.093
	7(**P2**-D1-JI) (*Mdn*=29.35)	
8	5(P1-**D1**-JI)* (*Mdn*=34.06)	*Z*=-.704, *p*=.243
	6(P1-**D2**-JI) (*Mdn*=30.84)	

Note. P-values are reported using the Exact statistic one-tailed. * indicates the group predicted to have the higher confidence rating.

Hypothesis 5. This hypothesis predicted that post-trial instruction and early presentation of DNA evidence by the prosecution would result in the highest conviction rate. Distribution of verdicts are presented in Table 7. A chi-squared test of independence indicated the relationship between condition and verdict was significant, χ^2 (7, N=257) = 14.46, p =.041. However, the hypothesis that the combination of early presentation of DNA evidence and post-trial instruction (Group 6) would result in the most convictions was unsupported. The two groups with the highest conviction rate were Group 2 receiving both pieces of evidence last and a pre-trial instruction and Group 8 receiving both pieces of evidence last and a post-trial instruction.

Table 7

Distribution of verdict across condition

Group	Guilty	Not guilty	Average probability of guilt
1(JI-P1-D1)	32%	68%	60
2(JI-P2-D2)	44%	56%	63
3(JI-P1-D2)	15%	85%	49
4(JI-P2-D1)	14%	86%	52
5(P1-D1-JI)	33%	67%	50
6(P1-D2-JI)	32%	68%	65
7(P2-D1-JI)	29%	71%	61
8(P2-D2-JI)	44%	56%	61

3.3. *Post-hoc* exploratory analysis

All exploratory analyses report the two-tailed statistic.

Open-ended responses

Open ended responses were deemed too unreliable for content analysis, given the variety and optional nature of responses.

Range of probabilities associated with each verdict category

Figure 8 illustrates the large range of overlap for probabilities of guilt associated with each dichotomous verdict.

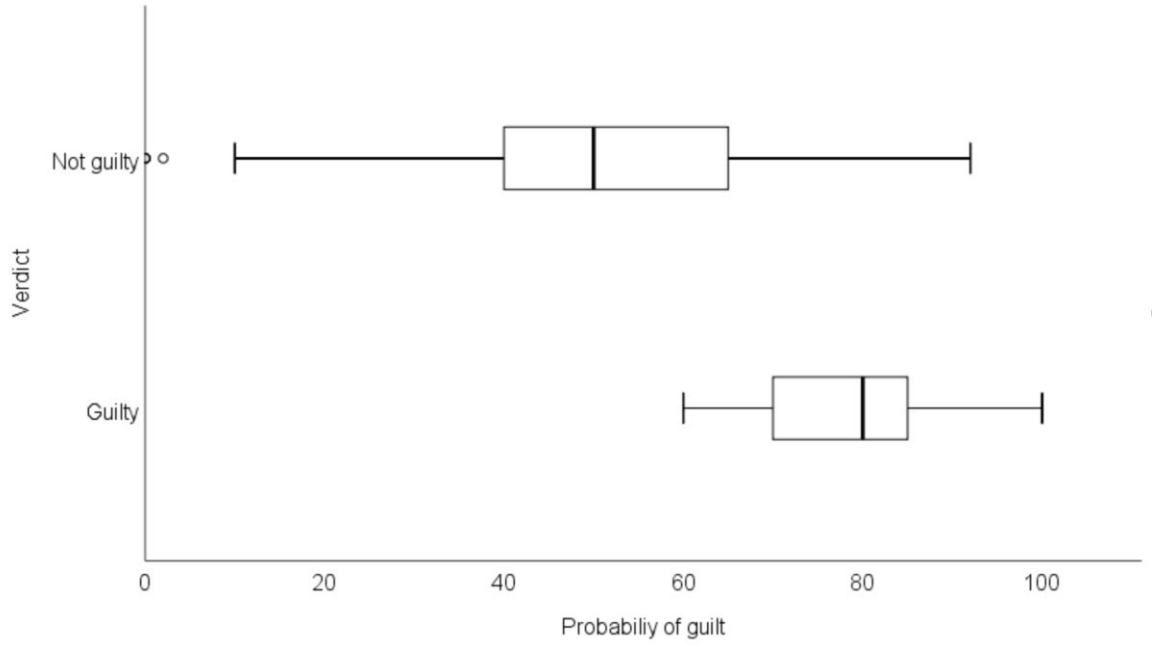


Figure 8. Boxplot of probability of guilt and verdict.

Confidence post hoc-analysis

Verdict was dummy coded to run a non-parametric test. Participants voting to convict (*Mdn*=166.01) were more confident that those voting to acquit (*Mdn*= 112.58), *Z*=-5.392, p=.000. Figure 9 below illustrates the distribution of confidence ratings by verdict.

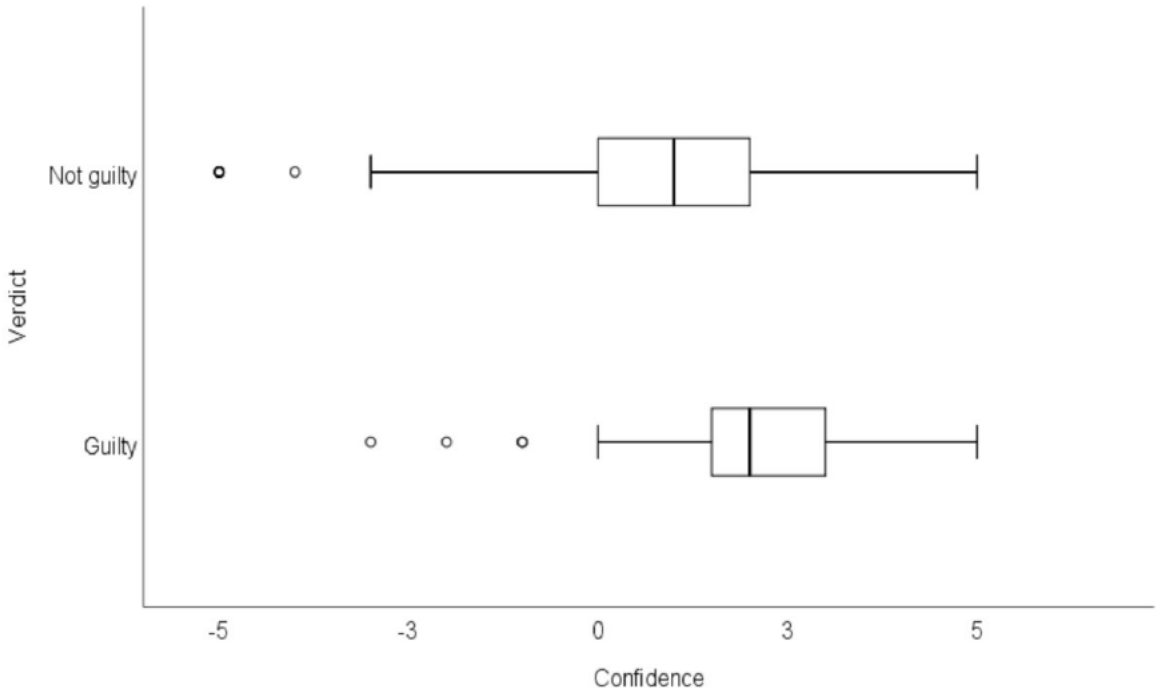


Figure 9. Boxplot showing distribution of confidence ratings by verdict.

-5= not at all confident, 0= not sure, +5=very confident

Verdict and understanding of DNA

There was a significant relationship between verdict and understanding of DNA evidence indicated by answers to question 42 (Appendix E), $\chi 2$ (3, N=257) = 11.744, p =.008. Correct understanding was associated with a lower conviction rate. Further, there was a significant association between response to the question and experimental condition, $\chi 2$ (21, N=257) = 39.74, p=.008. Table 8 below indicates participants receiving pre-trial instruction were more likely to correctly answer the question about DNA.

Table 8

Cross tabulation between condition and response

Group	Both statements are equally likely to be true (both types of fallacious reasoning)	Neither statements are likely to be true (**correct**)	The first statement is most likely to be true (prosecutors' fallacy)	The second statement is most likely to be true (defence attorneys' fallacy)
1(**JI**-P1-D1)	17 (50%)	6 (18%)	7 (20%)	4 (12%)
2(**JI**-P2-D2)	11 (32%)	8 (24%)	5 (15%)	10 (29%)
3(**JI**-P1-D2)	8 (25%)	13 (41%)	8 (25%)	3 (9%)
4(**JI**-P2-D1)	16 (57%)	6 (21%)	3 (11%)	3 (11%)
5(P1-D1-**JI**)	11 (33%)	5 (15%)	4 (12%)	13 (39%)
6(P1-D2-**JI**)	18 (58%)	2 (6%)	6 (19%)	5 (16%)
7(P2-D1-**JI**)	15 (49%)	5 (16%)	5 (16%)	6 (19%)
8(P2-D2-**JI**)	19 (56%)	5 (15%)	8 (24%)	2 (6%)

Analysis on gender effects

There were slight differences in the number of males and females across groups illustrated in Table 8. Given the materials included connotations to domestic violence gender differences were examined. However, there was no association between gender and verdict, χ^2 (1, N=257) = .856, p =.399.

Table 9

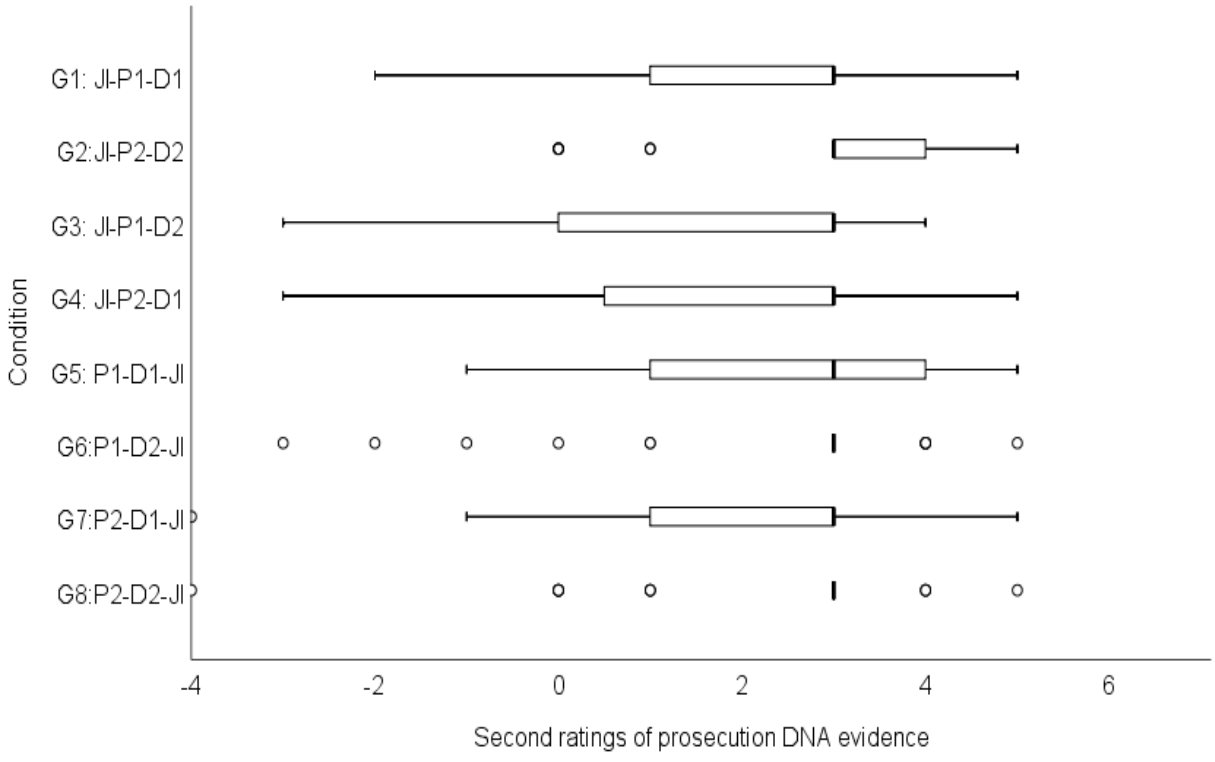
Demographic information including, gender and age across conditions

Group	Female	Male	Age	Number of participants
1(JI-P1-D1)	56%	68%	30	34
2(JI-P2-D2)	69%	56%	25	34
3(JI-P1-D2)	59%	85%	33	32
4(JI-P2-D1)	68%	86%	28	28
5(P1-D1-JI)	64%	34%	33	33
6(P1-D2-JI)	58%	42%	31	31
7(P2-D1-JI)	64%	36%	31	31
8(P2-D2-JI)	73%	27%	34	34
Total	64%	54%	31	57

Comparisons of secondary ratings pre-decision for DNA evidence

There was largely no significant difference amongst comparisons for the second rating of either prosecution or defence DNA evidence. Only one comparison between Group 2 (*Mdn*=39.12) and Group 3 (*Mdn*=27.52), indicated a significant difference, *Z*=-2.666, *p*=.008. The results of all comparisons are reported in Table 10 and 11 below.

Figure 10 below shows ratings of prosecution DNA evidence before deliberation. Further, there was no difference between verdict and total coherence-shift (H (1) =1.129, *p*=.288), with a mean rank of 125.72 for not guilty, and 136.39 for guilty. A boxplot displays total coherence-shift by verdict where total coherence-shift score is a summed difference between initial and secondary ratings.



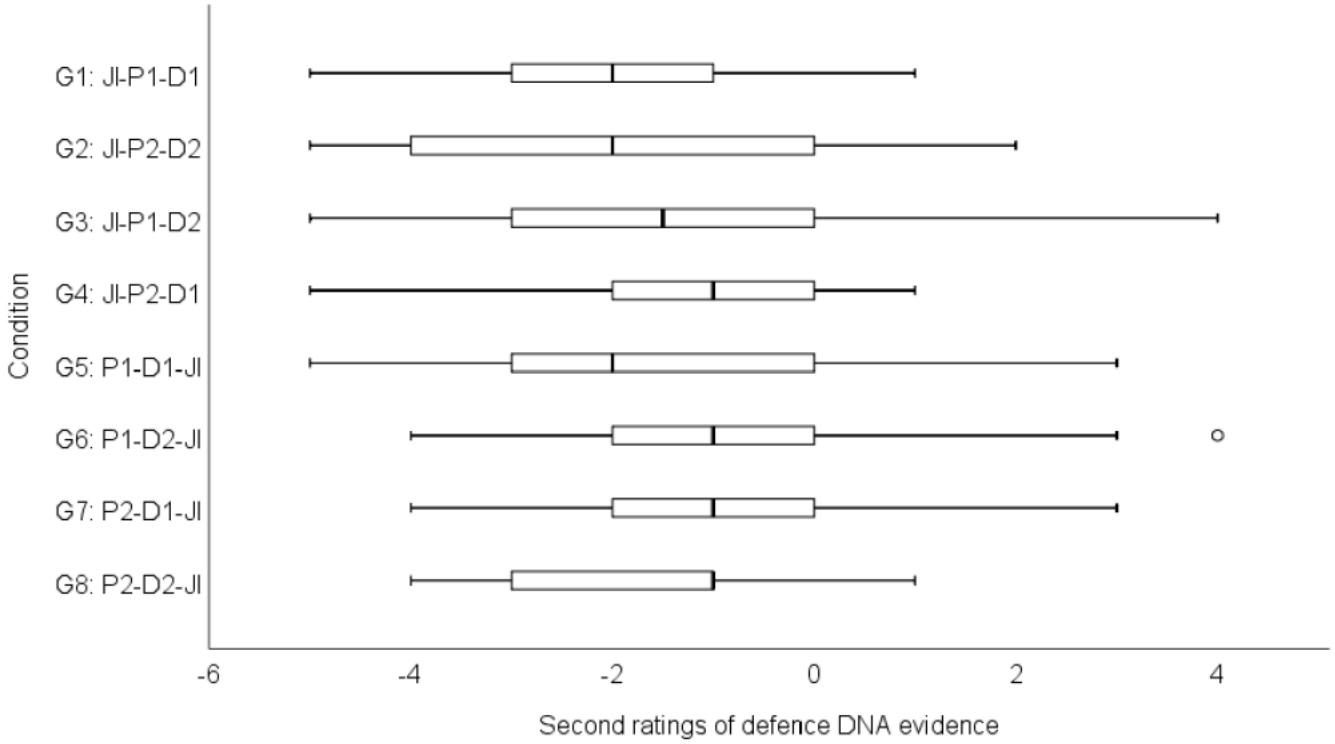


Figure 11. Boxplot showing secondary ratings of defence DNA evidence

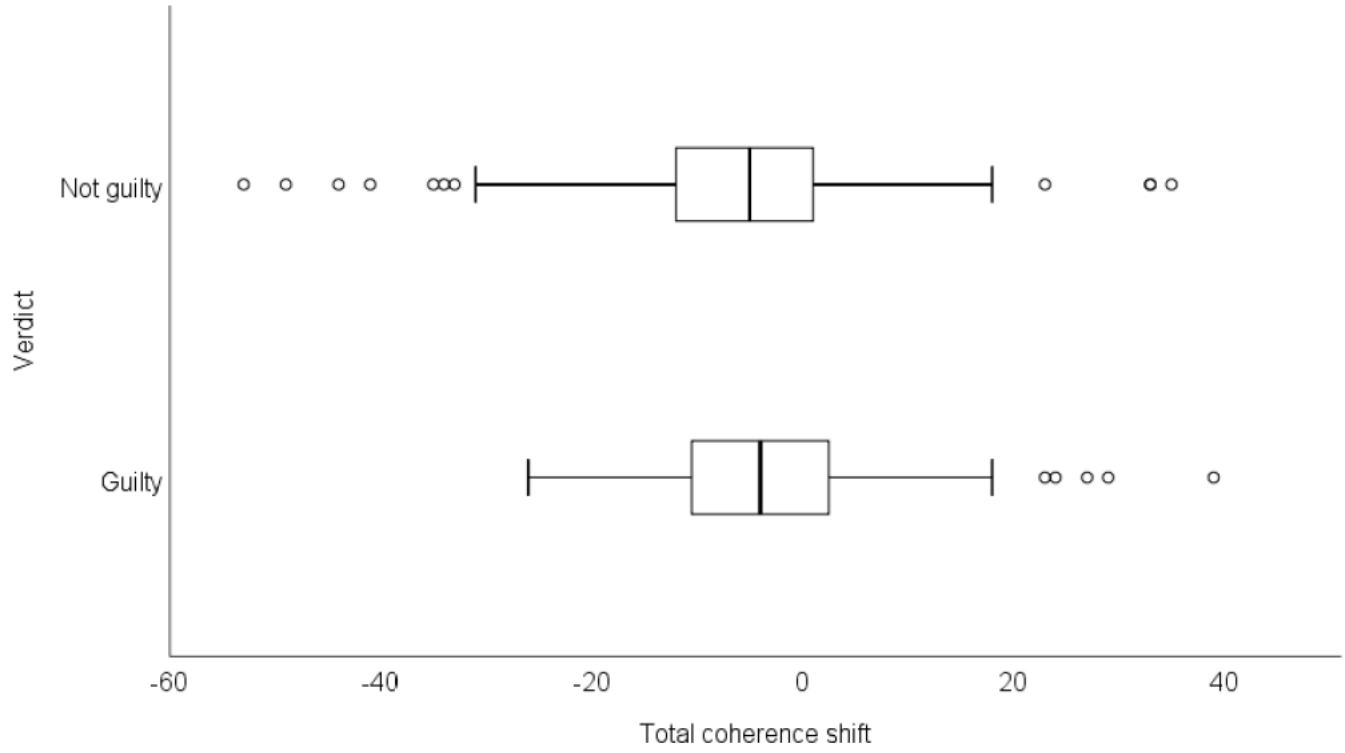


Figure 12. Boxplot showing total coherence-shift by verdict

Table 10

Comparisons for second ratings of prosecution DNA evidence

Comparison groups	Prosecution second ratings Mann Whitney U tests
2 (JI-**P2**-D2)	Group 2 (*Mdn*=39.12)
3 (JI-**P1**-D2)	Group 3 (*Mdn*=27.53)
	Z=-2.666, *p*=.008
6(**P1**-D2-JI)	Group 6 (*Mdn*=34.42)
8(**P2**-D2-JI)	Group 8 (*Mdn*=31.71)
	Z=-.669, *p*=.504
5(**P1**-D1-JI)	Group 5 (*Mdn*=31.19)
7(**P2**-D1-JI)	Group 7 (*Mdn*=33.73)
	Z=-.577, *p*=.564
1 (JI-**P1**-D11)	Group 1 (*Mdn*=34.53)
4 (JI-**P2**-D1)	Group 4 (*Mdn*=27.82)
	Z=-1.562, *p*=.118

Note. All two tailed statistics are reported.

Table 11

Comparisons for second ratings of defence DNA evidence

Comparison groups	Defence second ratings Mann Whitney U test
7(P2-**D1**-JI)	Group 7 (*Mdn*= 35.00)
8(P2-**D2**-JI)	Group 8 (*Mdn*=31.18)
	Z=-.834, *p*=.409
5(P1-**D1**-JI)	Group 5 (*Mdn*=30.67)
6(P1-**D2**-JI)	Group 6 (*Mdn*=34.45)
	Z=-.829, *p*=.412
1(JI-P1-**D1**)	Group 1 (*Mdn*=32.72)
3(JI-P1-**D2**)	Group 3 (*Mdn*=34.33)
	Z=-.347, *p*=.733
2(JI-P2-**D2**)	Group 2 (*Mdn*=28.49)
4(JI-P2-**D1**)	Group 4 (*Mdn*=35.16)
	Z=-1.470, *p*=.144

Note. All two-tailed statistics are reported.

Chapter 4: Discussion

4.1 Overview

The aim of this study was to investigate timing effects on DNA evidence with the goal to help jurors make objective decisions. The results must be considered in light of the limitations discussed below, however, they suggest that recency effects and coherence-shifts work in opposite directions. This process appears to result in similar ratings of DNA by the end of the trial. The different findings may indicate different cognitive processes for both evaluation and integration and for prosecution and the defence.

DNA evidence

DNA evidence presented at the end of the trial was rated more highly, yet DNA evidence presented at the beginning was associated with greater coherence-shifting *for the prosecution*. DNA evidence did not inflate subsequent evidence or cause an increase in confidence when presented early.

Judicial instruction

Pre-trial instruction was associated with lower initial ratings for prosecution DNA, yet greater coherence-shifts. Pre-trial instruction did not impact defence DNA and was associated with increased understanding of DNA evidence.

4.2 What effect does timing of DNA evidence in a criminal trial have on mock-juror evaluation of evidence?

Hypothesis 1 was partially supported as DNA presented first was associated with greater coherence-shifts. This confirms the literature indicating that strong inculpating evidence triggers coherence-shifts (Holyoak & Simon, 1999; Simon, 2004; Greenspan & Scurich, 2016). In opposition, both prosecution and defence initial ratings were higher when DNA was presented last. This supports the recency effects found by Constabile and Klein (2010) and Charman et al. (2016).

Although contradictory, the results are compatible with CBR. As belief builds during a trial, evidence given last, may enter a set of propositions that provide a supporting framework for acceptability (Thaggard, 1989; 2005). The results indicate this causes higher ratings of DNA compared to when that same evidence enters the representation early in the trial. The increased coherence-shift for early presentation of *prosecution* DNA supports the bidirectional nature of reasoning in previous studies (Holyoak & Simon, 1999; Simon, 2004; Lundberg, 2004). Overall, the contradictory effects may by symptomatic of evaluation and integration being separate cognitive processes as has been alluded to by Charman et al (2013).

The defence DNA was less impacted by recency effects, and not impacted by the coherence-shifts at all. These inconsistencies may signal increasing constraint on the representation explained by CBR and Explanatory Coherence (Thaggard, 1989; Glöckner & Betsch, 2008; Holyoak & Simon, 1999; Glöckner & Engel, 2013; Glöckner, Betsch & Schindler, 2010). As the trial progresses, the mental representations become increasingly *fixed* by the prosecution case. Therefore, in contrast the prosecution may allow for more coherence-shifting compared to the defence who must put forth their case in light of what the prosecution has presented.

Due to a lack of similar studies and unknown effect size, it cannot be ascertained if the frameworks explain these inconsistencies across the prosecution and defence, or if they indicate a lack of power.

4.3 Can a pre-trial instruction mitigate coherence-shifts better than a post-trial instruction?

Mixed results provide partial support for Hypothesis 4. Pre-trial instruction was associated with lower *initial* ratings of prosecution DNA and increased understanding of DNA evidence. The latter indicates pre-trial instruction is more effective at combating the prosecutor's fallacy (Thompson, 1989) and provides an explanation for participants lowered initial ratings of DNA. Two groups receiving pre-trial instruction had low conviction rates of

14-15%. This indicates support for the effectiveness of pre-trial instruction found in previous studies (Kassin & Wrightsman, 1979; ForsterLee, Horowitz & Bourgeois, 1993; Elwork, Sales, & Alfini, 1977). It is also consistent with the proposition that pre-trial instruction provides a legally appropriate index for jurors to evaluate and integrate evidence throughout the trial (Pennington & Hastie, 1993). The low rates of DNA comprehension (especially Group 6 with only 6% of participants correctly understanding DNA) indicate participants in post-trial conditions struggled to retrospectively correct their initial understandings.

However, inconsistencies, particularly Group 2 (44% conviction rate) who received pre-trial instruction did not conform to this trend. Two comparisons did not reach significance for pre-trial instruction reducing initial ratings of prosecution DNA evidence, which is in line with other studies which failed to find a significant effect (Wheatcroft & Keogan, 2017; Rowe, 1997). This further indicates that the current study is underpowered despite having 257 respondents.

Further, pre vs. post-trial instruction had no significant effect on initial ratings of defence DNA. As discussed under 4.1 above, this may be due to the institutional structure of a trial. Although Charman et al (2016) found rumination had no impact on verdict, proximity of the prosecution case to the pre-trial instruction may be a mediating factor for its effectiveness. This would explain the lack of difference on the side of the defence in the current study.

Pre-trial instruction was associated with an increased coherence-shift for two comparisons of *prosecution* DNA, yet the manipulation had no significant effects for the defence. The increasing constraint on the network may explain this. Further, although shifting more, examination of the distribution (see Figure 5) indicates that coherence-shift scores fell around 0 (0 = no coherence-shift), with post-trial instructed groups more likely to shift towards the defence. Therefore, a likely explanation is that the lowered initial ratings with pre-trial instruction accounted for the slight increase in coherence-shift.

Finally, there was largely no significant difference on second ratings of DNA. Given the mixed results and indication that timing effects of DNA mediate one another, the timing manipulations alone cannot be attributed as the only cause for the discrepant conviction rates. However, examination of Figure 10 indicates the distribution of pre-instructed groups (excluding Group 2) were skewed to favour the defence more so that post-instructed groups.

Ultimately, as concluded in 4.1 above, inconsistent results indicate a lack of power and results and interpretation are therefore not reliable.

4.4 Strengths

A focus on the manipulated DNA coherence-shift as a dependent measure enables more insight into the cognitive process, which previous studies focusing on verdict and probability of guilt do not afford (Charman et al., 2016; Glöckner & Engel, 2013).

In addition, although studies on coherence-shifts have tracked ratings of evidence, they have not looked at the impact of ordering effects on coherence-shits, instead changing facts (Holyoak & Simon, 1999; Greenspan & Scurich, 2016; Fawcett, 2016). Therefore, to our knowledge, this is the first study to simultaneously inspect the cognitive ordering effects on coherence-shifting for both the prosecution and the defence.

The current study has quite good internal validity. Minor changes to the temporal order of the same evidence excludes the possibility that the dependent measure is influenced by separate inferences caused by changing the strength or content of a piece of evidence as has been done in prior studies eliciting coherence-shifting (Holyoak & Simon, 1999; Greenspan & Scurich, 2016; Fawcett, 2016). The current study also avoids potential priming effects, elicited by the use of pre-test vignettes. Further, the cognitive process can potentially be generalized to other types of prejudicial evidence.

Finally, the current study enabled the recent assimilation hypothesis (Rassin, 2017) to be tested with more ecologically valid stimulus materials.

4.5 Limitations

Pre-trial instructed participants did not read an instruction at the end compared to post-trial instructed participants. This resulted in different time delays between second ratings of evidence for participants in pre vs. post-instructed conditions. Although Charman et al. (2016) found no effect of rumination, memory recall has been proposed as a mediating factor in the literature for recency effects (Constabile & Klein, 2010) and pre-trial instruction effectiveness (Elwork, Sales, & Alfini, 1977; Kassin & Wrightsman, 1979). The longer delay for post-trial instructed jurors may have reduced memory recall for the defence DNA and this may confound the coherence-shift measure.

Similarly, proximity of the instruction to the DNA evidence may constitute a mediating factor for weightings of evidence. This ecological confound arises by the nature of attempting to conduct an experiment within the constraints of what is externally applicable in the legal institution (i.e. defence presenting case last). Given it applies externally in the legal system it is a valid confound.

The lack of a significant difference on second ratings of evidence could also indicate that participants tired of rating the evidence by the end of the trial or alternatively in a 40-minute reading time were able to consciously remember their initial ratings. In a courtroom it is likely jurors are anchored by their initial impressions of evidence, however, the methodology in the current study may pronounce this effect. Therefore, future research should consider including a distractor task in between the end of trial materials and the re-rating of evidence, to account for this limitation.

The current study did not account for belief bias. This may account for the inconsistent results, and in particular, the high conviction rate in Group 2 (44%) who received a pre-trial instruction. Prior beliefs and experiences form a large portion of Pennington and Hastie's

experimental research on the Story Model, influencing representations constructed (1988, 1992; Klauer, Musch, & Naumer, 2000). Open-ended responses indicated that some participants were pro-prosecution. Further, a body of research on the 'CSI effect' indicates individuals can vary in their prior expectations on forensic evidence depending on exposure to forensic evidence in television (Schweitzer & Saks, 2007).

This study does not account for trial deliberation. Some research indicates groups may further polarize facts during the decision-making process (Sunstein, 2000) although several studies claim that individual preference for verdict predicts group verdict (Devine, Clayton, Dunford, & Seying, 2001). This limits the current study's external validity.

Analysis indicated no association between gender and verdict, supporting findings from a recent meta-analysis, (Devine & Caughlin, 2014). Despite this, the gender imbalance together with the lack of screening for jury eligibility makes the sample bias. Specifically, a large portion of participants were students. Although Bornstein et al., (2017) provided support for the use of student samples in mock-juror research, Carlson and Russo (2001) did not find the same level of predecisional distortion in their student sample compared to prospective jurors, which may explain the lack of a significant difference between verdict and total coherence-shift scores. Lastly, the online platform excluded individuals without computer access. The above all limit the external validity of the study sample.

Finally, the inconsistencies discussed above indicate the study is does not have the power to detect consistent effects between-groups with the current subtle manipulations.

4.6 Implications for the literature

The literature has demonstrated proleader and antitrailer distortion (Carlson & Russo, 2001). Yet the results in the current study do not replicate these findings, indicated by the lack of difference between verdict and total coherence-shift. Despite this, results support that those

voting to convict are more confident (Glöckner & Engel, 2013; Simon 2004). Interestingly, proleader and antitrailer distortion (Blanchard, Carlson & Meloy, 2014) has been attributed as the cause of increased confidence, yet the current study did not find a significant difference between total coherence-shift and verdict. This indicates confidence is more associated with reaching the standard of proof.

Further, the current study failed to replicate the recent assimilation hypothesis (Rassin, 2017) by finding no inflation of subsequent evidence. Participants in the current experiment were required to evaluate 12 pieces of ambiguous evidence compared to participants in the initial experiment (Rassin, 2012) where participants only evaluated four short vignettes which, all implicated the defendant. This indicates more power is required in an ecologically valid experiment to detect the effect.

The manipulations in the current experiment may lack the strength required to see an effect. For example, participants may have identified the "neighbours' testimony" (see Appendix E) as being 'obviously' weak.

The range of probabilities of guilt associated with each verdict raises questions about the use of probability of guilt as a dependent measure in the literature. It indicates probability of guilt in place of verdict is not an externally valid measure. The median probability for conviction was 80%, aligning with previous findings by Glöckner and Engel (2013).

Finally, open-ended responses provided support for the Story Model. Participants reported inferences not given in the transcript. Examples include; "*Mostly relied on the dna evidence being on her finger nails and thought she might have grabbed his hair in pain as she begged him to stop?"* and *"…. other ways he could have committed the crime – like borrowing his girlfriend's car or other things like that."* Both of these indicate story construction. One participant responded; *"When dealing with complex scientific arguments that are sometimes*

conflicting, such as DNA, I find myself relying more heavily on narrative and emotion," indicating a level of self-awareness on reliance on story creation.

4.7 Implications for the legal system

Poor comprehension of judicial instruction has been associated with reliance on heuristic processing, and racial stereotypes (Lynch & Haney, 2000). Therefore, any improvement on jurors' ability to comprehend and apply instruction is relevant for the legal system. Current findings that pre-trial instruction is associated with less fallacious reasoning (specifically, the prosecutor's fallacy) could be included in a report to the Courts Administration Authority reviewing current recommendations on appropriate timing of judicial instructions.

However, despite this the majority of participants still failed to correctly understand the DNA evidence, indicating that, contrary to the judgement of *Aytugrul v The Queen*, judicial instruction is not an adequate safeguard. Therefore, the results also contribute to the growing support for the prohibition of the exclusion percentage in court.

4.8 Directions for future research

Cognitive research on juror decision making is disjointed and uncertain. A systematic review would provide valuable direction for future research. Given the unknown effect size, future research aiming to replicate the current study should aim for a larger sample size until a meta-analysis is viable.

Extensive research on serial position effects illustrates that both first and last pieces of information in a sequence are recalled more easily than those in the middle (Murdock, 1962; Anderson, Bothell, Lebiere, & Matessa, 1998; de Bruin, 2005). The current study did not examine the effect of presenting DNA mid-trial. DNA presented at the beginning and end of trial could both be more prejudicial than mid-trial presentation. Such a study examining DNA mid-trial would also provide clarity and support for CBR. If mid-trial presentation were found to be more prejudicial than DNA at the beginning, it would further support that belief builds

throughout the trial. However, if both first and last presentation is associated with increased ratings this would provide clarity over whether memory is a mediator of context effects.

To account for forensic evidence belief bias, future research should consider using the FEEBS scale as a pre-assessment measure to analyse prior beliefs as a mediating factor for DNA weightings (Smith & Bull, 2012; 2014). Further the measure could assess if pre-trial instruction combats prior belief bias over forensic evidence more so than post-trial instruction. As the current study indicates belief builds, future research could consider breaking "the story" presented by a party by objections, and cross-examination at different time points. Further proximity of judicial instruction to DNA may be a mediating factor and future research could specifically examine this.

The literature has focused on predicting verdict with qualitative research, but given the open-ended responses discussed in 4.6, future research should consider designing a mixed-methods or qualitative study. This would add depth to participants' experience of story construction and enable exploration of how pre-trial bias impacts jurors' decisions.

4.9 Concluding remarks

Despite evidentiary thresholds, the courts demand for an impartial juror, capable of objectively evaluating incomplete, contradictory and complex evidence appears incompatible with an adversarial system where the reliability of forensic evidence is obscure. In light of the subtle manipulation in the current study, the large discrepancy in conviction rates across conditions supports this. It begs the question; can we expect jurors to remain impartial as forensic science continues to evolve beyond the scope of what most laypeople can critically evaluate?

This is particularly relevant in light of the prejudicial nature of DNA evidence, the fallibility of human cognitive processes and the grave consequence that a criminal conviction can have on individual liberty. Despite this, the experimental results indicate that timing of DNA

evidence alone does not have a huge impact on weight attribution by the end of the trial. The recency effect and the coherence-shift working in opposite directions seem to mitigate one another. The effect of pre-trial instruction is less clear but provides partial support for its effectiveness, and perhaps indicates more power is needed achieve consistent results. The limitations discussed above make for tentative conclusions, and future research needs to replicate, improve on the current studies limitations discussed above and examine these effects further.

References

Amaya, A. (2007). Formal models of coherence and legal epistemology. *Artificial Intelligence Law*, 15(4): 429-447.

Anderson, C.A., & Sechler, E., S. (1986). Effects of explanation and counter explanation on the development and use of social theories. *Journal of Personality and Social Psychology,* 50 (1): 24-34. doi: 10.1037/0022-3514.50.2.24.

Anderson, J.R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language,* 38(4):341-380. doi: 10.1006.jmla.1997.2553.

Appleby, S. C., & Kassin, S.M. (2016). When Self-Report Trumps Science: Effects of Confessions, DNA, and Prosecutorial Theories on Perceptions of Guilt. *Psychology, Public Policy, and Law*, 22(2): 127-140.

Aytugrul v The Queen [2012] HCA 15.

Baguley, C. M., McKimmie, B.M., & Masser, B.M. (2017). Deconstructing the simplification of jury instructions: How simplifying the features of complexity affects jurors' application of instructions. *Law and Human Behavior*, 41(3), 284-304. doi:10.1037/lhb0000234

Blanchard, S.J., Carlson, K.A., & Meloy, M.G. (2014). Biased predecisional processing of leading and nonleading alternatives. *Psychological Science*, 25(3): 812-816. doi: 10.1177/0956797613512663.

Bornstein, B. J., Golding, J.M., Neuschatz, J., Kimbrough, C., Reed, K., Magyarics, C., & Luecht, K. (2017). Mock juror sampling issues in jury simulation research: A meta-analysis. *Law and Human Behavior,* 41(1). doi:10.1037/lhb0000223

Byrne, M. D. (1994). The Convergence of Explanatory Coherence and the Story Model: A Case Study in Juror Decision. In J.D. Moore & J.F. Lehman (Eds.), *Proceedings of the*

Seventeenth Annual conference of the Cognitive Science Society. Mahwah, NJ: Lawrence Erlbaum Associates.

Carlson, K. A., & Guha, A. (2011). Leader-focused search: The impact of an emerging preference on information search. *Organizational Behaviour and Human Decision Processes*, 115, 133-141.

Carlson, K. A., & Pearo, L.K. (2004). Limiting predecisional distortion by prior valuation of attribute components Organizational Behavior and Human Decision Processes, 94, 48-59.

Carlson, K. A., & Russo, J.E. (2001). Biased interpretation of evidence by mock jurors. *Journal of Experimental Psychology: Applied*, 7, 91-103.

Carlson, K. A., Meloy, M.G., & Russo, J.E. (2006). Leader driven primacy: Using attribute order to affect consumer choice. *Journal of Consumer Research*, 32, 513-518.

Charman, S.D. (2013). The forensic confirmation bias: A problem of evidence integration, on just evidence evaluation. *Journal of Applied Research in Memory and Cognition,* 2:56-58. doi: 10.1016/j.jarmac.2013.01.101

Charman, S.D., Carbone J., Kekessie, S., & Villalba, D.K. (2016). Evidence Evaluation and Evidence Integration in Legal Decision-Making: Order of Evidence Presentation as a Moderator of Context Effects. *Applied Cognitive Psychology,* 30: 214-225. doi: 10.1002/acp.3181.

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(1):76-90. doi: 10.1037/a0014682.

Colombo, C. (2018). Hearing the Other Side? - Debiasing Political Opinions in the Case of the Scottish Independence Referendum. *Political Studies*, 66(1), 23-42.

Constiable, K.A., & Klein, S.B. (2010). Finishing Strong: Recency Effects in Juror Judgements. *Basic and Applied Social Psychology,* 27(1):47-58.

Cutler, B. L., & Penrod, S.D. (1995). *Mistaken identifications: the eyewitness, psychology, and the law*. Cambridge, UK: Cambridge University Press.

d'Avila Garez, A. S., Gabbay, D.M. & Luis, C.L. (2014). A neural cognitive model of argumentation with application to legal inference and decision-making. *Journal of Applied Logic*, 12, 109-127.

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 Behaviour*, 33(5). doi:10.1007/s10979-008-9151-y.

Dartnall, S., & Goodman-Delahunty., J. (2006). Enhancing Juror Understanding Of Probabilistic DNA Evidence. *Australian Journal of Forensics,* 38(2):86-96. doi: 10.1080/00450610609410635.

Davis, J.H., Tindale, R.S., Nagao, D., Hinsz, V.B., & Robertson, B. (1984). Order effects in multiple decisions by groups: A demonstration with mock juries and trial procedures. *Journal of Personality and Social Psychology,* 47(5), 1003-1012.

de Bruin W.B. (2005). Save the last dance for me: unwanted serial position effects in jury evaluations. *Acta Psychologica*, 118(3): 245-260. DOI:10.1016/j.actpsy.2004.08.005

DeKay, M. L., Miller, S.A., Schley, D.R., & Erford, B. M. (2014). Proleader and antitrailer information distortion and their effects on choice and post-choice memory. *Organizational Behavior and Human Decision Processes*, 125, 134-150.

DeKay, M. L., Stone E.R., & Miller, S.A. (2011). Leader-driven distortion of probability and payoff information affects choices between risky prospects. *Journal of Behavioural Decision Making*, 24, 394-411.

DeKay, M. L., Stone, E., R., & Sorenson, C.M. (2012). Sizing up information distortion: Quantifying its effect on the subjective values of choice options. *Psychonomic Bulletin & Review*, 19, 349-356.

Devenport, J. L. C., B.L. (2004). Impact of defense-only and opposing eyewitness experts on juror judgements. *Law and Human Behavior*, 28(5), 569-576.

Devine, D.J., Clayton, L.D., Dunford, L.D., & Seying, R.P.J. (2001). Jury decision making: 45 years of empirical research on deliberating groups. *Psychology, Public Policy, and Law,* 7(3), 622-727.

Dioso-Villa, R. (2015). A Repository of Wrongful Convictions in Australia: First Steps Toward Estimating Prevalence and Causal Contributing Factors. *Flinders Law Journal,* 17(2).

Dror, I.E., & Hampikian, G. (2011). Subjectivity and bias in forensic DNA mixture interpretation. *Science & justice,* 51(4). doi: 10/1016/j.scijus.2011.08.004

Eastwood, J. P., & Caldwell J. (2015). Educating Jurors about Forensic Evidence: Using an Expert Witness and Judicial Instructions to Mitigate the Impact of Invalid Forensic Science Testimony. *Journal of Forensic Sciences*, 60(6).

Ecker, U. K., Lewandowsky, S. & Tang, D.T. (2010). Explicit warnings reduce but do not eliminate the continued influence of misinformation. *Memory and Cognition*, 28(8): 1087-1100.

Elwork, A., Sales, B.D., & Alfini, J.J. (1977). Juridic Decisions in Ignorance of the Law or in Light of It? *Law and Human Behaviour*, 1:163-289.

Englich, B., Mussweiler, T. & Strack, F. (2005). The last Word in Court - A Hidden Disadvantage for the Defense. *Law and Human Behavior*, 29(6), 705-722.

Englich, B., Mussweiler, T. & Strack, F. (2006). Playing Dice with Criminal Sentences: The Influence of Irrelevant Anchors on Experts' Judicial Decision Making. *Personality and Social Psychology Bulletin,* 32(2), 188-200. doi: 10.1177/01461670582152

Estrada-Reynolds, V.C., Gray, J.M., & Nunez, N. (2015). Information Integration Theory, Juror Bias, and Sentence Recommendations Captured Over Time in a Capital Trial. *Applied Cognitive Psychology,* 29(5): 713-722. doi 10.1002/acp.3155.

Faigman, D.L., & Baglioni, A.J. (1988). Bayes' theorem in the trial process. *Law and Human Behaviour,* 12(1), 1-17.

Fawcett, H. (2016). The effect of evidence timing and witness motivation upon juror evaluations of alibi witnesses and defendants. *Psychiatry, Psychology and Law*, 23(575-587).

ForsterLee, L., Horowitz, I.A. & Bourgeois, M.J. (1993). Juror Competence in Civil Trials: Effects of Preinstruction and Evidence Technicality. *Journal of Applied Psychology,* 78(1), 14-21.

Furnham, A. The Robustness of the Recency Effect: Studies Using Legal Evidence. *The Journal of General Psychology,* 113(4), 351-57. Doi:10.1080/00221309.1986.9711045.

Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics,* 40(1), 35-42. doi: 10.1016/j.socec.2010.10.008.

Glöckner, A. (2016). The reversed description-experience gap: Disentangling sources of presentation format effects in risky choice. *Journal of Experimental Psychology: General,* 14, 486-508.

Glöckner, A., & Betsch, T. (2008). Modelling option and strategy choices with connectionist networks: Towards an integrative model of automatic and deliberate decision making. *Judgement and Decision Making,* 3(3), 215-228.

Glöckner, A., & Engel, E. (2013). Can We Trust Intuitive Jurors? Standards of Proof and the Probative Value of Evidence in Coherence-Based Reasoning. *Journal of Empirical Legal Studies,* 10(2):230-252.

Glöckner, A., Betsch, T., & Schindler, N (2010). Coherence Shifts in Probabilistic Inference Tasks. *Journal of Behavioural Decision Making,* 23: 439-462. doi: 10.1002/bdm.668.

Goodman-Delahunty, J. & Hewson., L. (2010). Improving jury understanding and use of expert DNA evidence. *AIC Reports: Technical and Background Paper*, No 37. Canberra: Australian Institute of Criminology.

Greenspan, R., & Scurich, N. (2016). The Interdependence of Perceived Confession Voluntariness and Case Evidence. *Law and Human Behavior,* 40(650-659).

Guthrie, C., Rachlinski, J.J. & Wistrich, A.J. (2001). Inside the Judicial Mind. *Cornell Law Review,* 86(4).

Hampikian, D. (2011). Forensic examiners beliefs about the suspect can influence interpretation of DNA evidence. *Science & Justice,* 51(4), 204-208.

Hastie, R., Penrod, S.D., & Pennington, N. (1983). *Inside the Jury*. Cambridge, Mass: Harvard University Press.

Heuer, L., & Penrod, S. D. (1988). Increasing jurors' participation in trials: A field experiment and question asking. *Law and Human Behaviour*, 12(3): 231-261. doi: 10.1007/BF01044383.

Heuer, L., & Penrod, S.D. (1989). Instructing jurors - A field experiment with written and preliminary instructions. *Law and Human Behavior*, 13(4), 409-430. doi:10.1007/BF0156412.

Holyoak, K. J., & Simon, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. *Journal of Experimental Psychology*: *General*, 128(1), 3-25.

Holyoak, K.J., & Simon, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. *Journal of Experimental Psychology: General,* 128(1). 3-31.

Hope, A, L. Memom, P., & McGeorge, P. (2004). Understanding Pretrial Publicity: Predecisional Distortion of Evidence by Mock Jurors. *Journal of Experimental Psychology: Applied*, 10(2), 111-119. doi: 1037/1076-898X.10.2.111.

Johnson, M. D. R. J. T. (1998). How Not to Enhance the Confidence-Accuracy Relation: The Detrimental Effects of Attention to the Identification Process. *Law and Human Behaviour*, 22(4), 409-428.

Jonas, E., Schulz-Hardt, S., Frey, D., Thelen, N. (2001). Confirmation bias in sequential information search after preliminary decisions: an expansion of dissonance theoretical research on selective exposure to information. *Journal of personality and social psychology*, 80(4), 557-571.

Jones, E. E., Palmer, P. G., Jr., and Bandy, A.D. (2015). The Effect of Inconsistency on Evaluations of a Second Eyewitness: It Depends on Who Testifies First. *Psychiatry, Psychology and Law*, 22, 814-826.

Kassin, S. M., & Wrightsman, L.S. (1979). On the Requirements of Proof: The Timing of Judicial Instruction and Mock Juror Verdicts. *Journal of Personality and Social Psychology,* 37(10).

Kelley, H.H. (1973). The process of causal attribution. *American Psychologist*, 28, 107-128.

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

Kovera, M.B., & McAuliff, B.D. (2000). The effects of peer review and evidence quality on judge evaluations of psychological science: are judges effective gatekeepers? *Journal of Applied Psychology,* 85(4), 57-86. Doi:10.1037/0021-9010.85.4.574.

Kukucka, J. & Kassin, S.M. (2014). Do confessions taint perceptions of handwriting evidence? An empirical test of the forensic confirmation bias. *Law and Human Behavior*, 38(3).

Kukucka, J., & Kassin, S.M. (2014). Do confessions taint peceptions of handwriting evidnece? An empirical test of the forensic confirmation bias. *Law and Human Behavior,* 38(3), 256-70. doi: 10.1037/lhb0000066.

Lieberman, J. D., Carrel, C.A., Miethe, T.D., & Krauss, D.A. (2008). GOLD VS PLATINUM: Do Jurors Recognize the Superiority and Limitations of DNA Evidence Compared to Other Types of Forensic Evidence? *Psychology, Public Policy, and Law*, 14(1), 27-62. doi:10.1037/1076-8971.14.1.27.

Lundberg, C. G. (2004). Modeling and predicting emerging inference-based decisions in complex and ambiguous legal settings. *European Journal of Operational Research,* 153, 417-432. doi:10.1016/S0377-2217(03)00162-0.

Lundberg, C. G. (2011). Coherence and bidirectional reasoning in complex and risky decision-making tasks. *New Zealand Economic Papers*, 45(1-2), 161-181. doi:10.1080/00779954.2011.556077.

Martire, K. A., Kemp, R.I., Sayle, M. & Newell, B.R. (2014). On the interpretation of likelihood ratios in forensic science evidence: Presentation formats and the weak evidence effect. *Forensic Science International*, 240: 61-68. doi:10.1016/j.forsciint.2014.04.005.

McAuliff, B.D., Kovera, M.B., & Nunez, G. (2009). Can jurors recognize missing control groups, confound, and experimenter bias in psychological science? *Law and Human Behaviour,* 33(3), 247-57. doi. 10.1007/s10979-008-9133-0.

McCullough, G. W. (2007). Function of Text Structure in Jurors' Comprehension and Decision Making. *Psychological Reports,* 101, 723-730.

McQuiston-Surret, D. S., M.J. (2009). The Testimony of Forensic Identification Science: What Expert Witnesses Say and What Factfinders Hear. *Law and Human Behavior*, 33, 436-453. doi:10.1007/s10979-008-9169-1.

Miller, S. A., DeKay, M.L., Stone, E.R., & Sorenson C.M. (2013). Assessing the sensitivity of information distortion to four potential influences in studies of risky choice. *Judgement and Decision Making*, 8, 662-677.

Murdock, B.B. (1961). The serial position effect of free recall. *Journal of Experimental Psychology,* 64(5), 482-288.

Nance, D.A., & Morris, S.B. (2003). Juror Understanding of DNA Evidence: An Empirical Assessment of Presentation Formats for Trace Evidence with a Relatively Small Random-Match Probability. *The Journal of Legal Studies,* 32(2), 395-344.

National Research Council. 2009. *Strengthening Forensic Science in the United States: A Path Forward.* Washington, DC: The National Academics Press. doi. 10.17226/12/589.

Nurek, M., Kostopoulou, O. & Hagmayer, Y. (2014). Predecisional information distortion in physicians' diagnostic judgements: Strengthening a leading hypothesis or weakening its competitor? *Judgement and Decision Making*, 9(6), 572-585.

Ogloff, J. R. P., & Rose, V. G. (2005). The Comprehension of Judicial Instructions. In N. Brewer & K. D. Williams (Eds.), *Psychology and law: An empirical perspective* (pp. 407-444). New York, NY, US: Guilford Press.

Ola Svenson, N. G., Amina Memon & Torun Lindholm. (2017). Information about expert decision and post-decision distortion of facts on own decision. *Scandinavian Journal of Psychology*, 59(2).

Patiño-Echeverri, D., & Fischbeck, P. S. (2009a). Better safe than sorry: Precautionary reasoning and implied dominance in risky decisions. *Organizational Behavior and Human Decision Processes*, 109, 79-92.

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. doi: 10.1037/0022-3514.62.2.189.

Pennington, N., & Hastie, R. (1993). Reasoning in explanation-based decision making. *Cognition,* 49(1-2): 123-163. doi: 10.1016/0010-0277(93)90038-W.

Pennington, N., Hastie, R. (1986). Evidence Evaluation in Complex Decision Making. *Journal of Personality and Social Psychology*, 51(2), 242-258. doi:10.1037/0022-3514.51.2.242.

Pennington, N., Hastie, R. (1988). Explanation-Based Decision Making: Effects of Memory Structure on Judgement. *Journal of Experimental Psychology, Learning and Memory and Cognition*, 14(3). doi:10.1037/0278-7393.14.3.521.

Plous, S. (1993). *McGraw-Hill series in social psychology*. *The psychology of judgement and decision making*. New York, NY, England: Mcgraw-Hill Book Company.

Polman, E., & Russo, J.E. (2012). Commitment to a developing preference and predecisional distortion of information. *Organizational Behavior and Human Decision Processes,* 119(1), 78-88. Doi:10.1016/j.obhdp.2012.05.004.

President's Council of Advisors on Science and Technology, Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-comparison Methods, Executive Office of The President's Council of Advisors on Science and Technology. (2016). Washington DC. https://www.whitehouse.gov/administration/eop/ostp/pcast/docsreports.

Preston, C. C. & Coleman, A.M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104:2-15. doi: 10.1016/S0001-6918(99)00050-5/.

Price, H. L., & Dahl, L.C. (2013). Order and Strength Matter for Evaluation of Alibi and Eyewitness Evidence. *Applied Cognitive Psychology*, 28 (2): 143-150. doi: 10.001/acp.2983.

Rassin, E. (2017). Initial evidence for the assimilation hypothesis. *Psychology, Crime & Law,* 23 (10): 1010-1020. doi: 10.1080/1068316X.2017.1371307.

Rose, V.G., Chopra, S.R., & Ogloff, J.R.P. (2001). Evaluating the comprehensibility of jury instructions: A method and an example. *Law and Human Behaviour,* 25: 409-431.

Rowe, B.I. (1997). A Possible Solution for the Problem of Juries Slighting Nonscientific Evidence: A Bayesian-Like Judicial Instruction. *American Journal of Criminal Law,* 24 (3): 541-580.

Russo, E. P. J. E. (2012). Commitment to a developing preference and predecisional distortion of information. *Organizational Behavior and Human Decision Processes,* 119(1), 78-88.

Russo, J. E., & Chaxel, A.S. (2010). How persuasive messages can influence behaviour without awareness *Journal of Consumer Research,* 20(3), 338-342.

Russo, J. E., Carlson, K.A., Meloy, M.G., & Young, K. (2008). The goal of consistency as a cause of information distortion. *Journal of Experimental Psychology: General,* 137, 456-470.

Russo, J.E., Medvec., V. J., & Meloy, M.G. (1996). The Distortion of Information during Decisions. *Organizational Behavior and Human Decision Processes*, 66(1), 102-110. doi: 10.1006/obhd.1996.0041.

Schmittat, S. M., & Birte, E. If you judge, investigate! Responsibility reduces confirmatory information processing in legal experts. *Psychology, Public Policy, and Law*, 22(4), 386-400.

Schweitzer, N., & Saks, M. (2007). The CSI effect: Popular fiction about forensic science affects public expectations about real forensic science. *Jurimetrics,* 47:357-364.

Shaklee, H., & Fischhoff, B. (1982). Strategies of information search in causal analysis. *Memory & Cognition*, 10(6), 520-230.

Simon, D. (2004). A Third View of the Black Box: Cognitive Coherence in Legal Decision Making. *University of Chicago Law Review,* 71(2), 511-586.

Simon, D. (2012). *In Doubt*. Cambridge, Massachusetts. 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, Learning, Memory and Cognition*, 27(5), 1250-1260.

Simon, D., Snow, C.J., & Read, S.J. (2004). The Redux of Cognitive Consistency Theories: Evidence Judgements by Constraint Satisfaction. *Journal of Personality and Social Psychology,* 86(6):814-837. DOI: 10.1037/0022-3514.86.6.814.

Sloman, S. A. (1994). When explanations compete: the role of explanatory coherence on judgements of likelihood. *Cognition,* 52: 1-21. doi: 0010-0277/94/$07.00.

Smith, L. L., Bull, R. & Holliday, R. (2011). Understanding Juror Perceptions of Forensic Evidence: Investigating the Impact of Case Context on Perceptions of Forensic Evidence Strength. *Journal of Forensic Sciences*, 56(2). doi:10.1111/j.1556-4029.2010.01671.x.

Smith, V.L. (1991). Impact of Pretrial Instruction on Jurors' Information Processing and Decision Making. *Journal of Applied Psychology*. 76. 220-228. 10.1037/0021-9010.76.2.220.

Thaggard, P. (1989). Explanatory coherence. *Behavioural and Brain Sciences,* 12(3): 435-502. doi: 10.1017.S0140525X00057046.

Thaggard, P. (2005). Testimony, Credibility and Explanatory Coherence. *Erkenntnis.* 63:295-316. doi: 10.1007/s10670-005-4004-2.

Thompson, W. C. (1989). Are juries competent to evaluate statistical evidence? *Law and Contemporary Problems*, 52(4), 9-41.

Thompson, W. C., & Schumann, E.L. (1987). Interpretation of statistical evidence in criminal trials: The prosecutor's fallacy and the defense attorney's fallacy. *Law and Human Behavior*, 11 (3): 167-187.

Walmsley, S., & Gilbey, A. (2017). Debiasing visual pilots' weather-related decision-making. *Applied Ergonomics*, 65: 200-208. doi: 10.1016/j.apergo.2017.06.019.

Wheatcroft, J. M. & Keogan, H. (2017). Impact of Evidence Type and Judicial Warning on Juror Perceptions of Global and Specific Witness Evidence. *The Journal of Psychology,* 151(3), 247-26. doi: 10.1080/00223980.2016.1261077.

Appendix

[A-G]

Appendix A – Ethics Approval

THE UNIVERSITY OF ADELAIDE
AUSTRALIA

School of Psychology
University of Adelaide
North Terrace, Adelaide SA 5005
Ph. 61 8 8313 5693
Fax 61 8 8313 3770

School of Psychology: Human Research Ethics Subcommittee
Approval Sheet

Dear CAROLYN

The members of the subcommittee have considered your application:

Code Number: 18/61

Title:

EXPLORING STRENGTH OF DNA EVIDENCE

With [Student name, if applicable] [redacted]

I am writing to confirm that approval has been granted for this project to proceed. Approval is granted to 12 months from the date specified below.

Yours sincerely,

[redacted]

Deputy Convenor, Human Research Ethics Subcommittee

Name: PAUL DELFABBRO

Date: 20/5/18

Phone Number: 8313 4936

Email: paul.delfabbro@adelaide.edu.au

Appendix B – Poster

Innocent or guilty???

The study: Participation is voluntary. Jurors are comprised of the general population. This study aims to investigate the ways in which jurors integrate various pieces of evidence in an ambiguous trial.

Time: The time will depend on your reading speed, but it is estimated this study will take no more than 40 minutes. You will be required to read a transcript of a trial and render a verdict.

Risks: There are no immediate risks to your health of safety in completing this study. It can be done on any computer in any place. However, the trial does contain facts about a murder. If this is likely to cause any distress to you, we advise you consider not participating in this study.

Privacy: The data collected in this study will be made available on the **Open Science Framework**, but all of your responses and data will be de-identified. Only the quantitative (numerical) aspects of the data will be made available to researchers for future study. You will be able to withdraw from the survey at any point in time, however, once you have completed the survey your information will be deidentified. Thus, there will be no way to retract your results once they have been submitted, as there will be no way to identify your results.

Human research ethics committee approval number: [redacted]

Principle investigator: Dr Carolyn Semmler

Student researcher:

Student's degree: Honours in Psychology

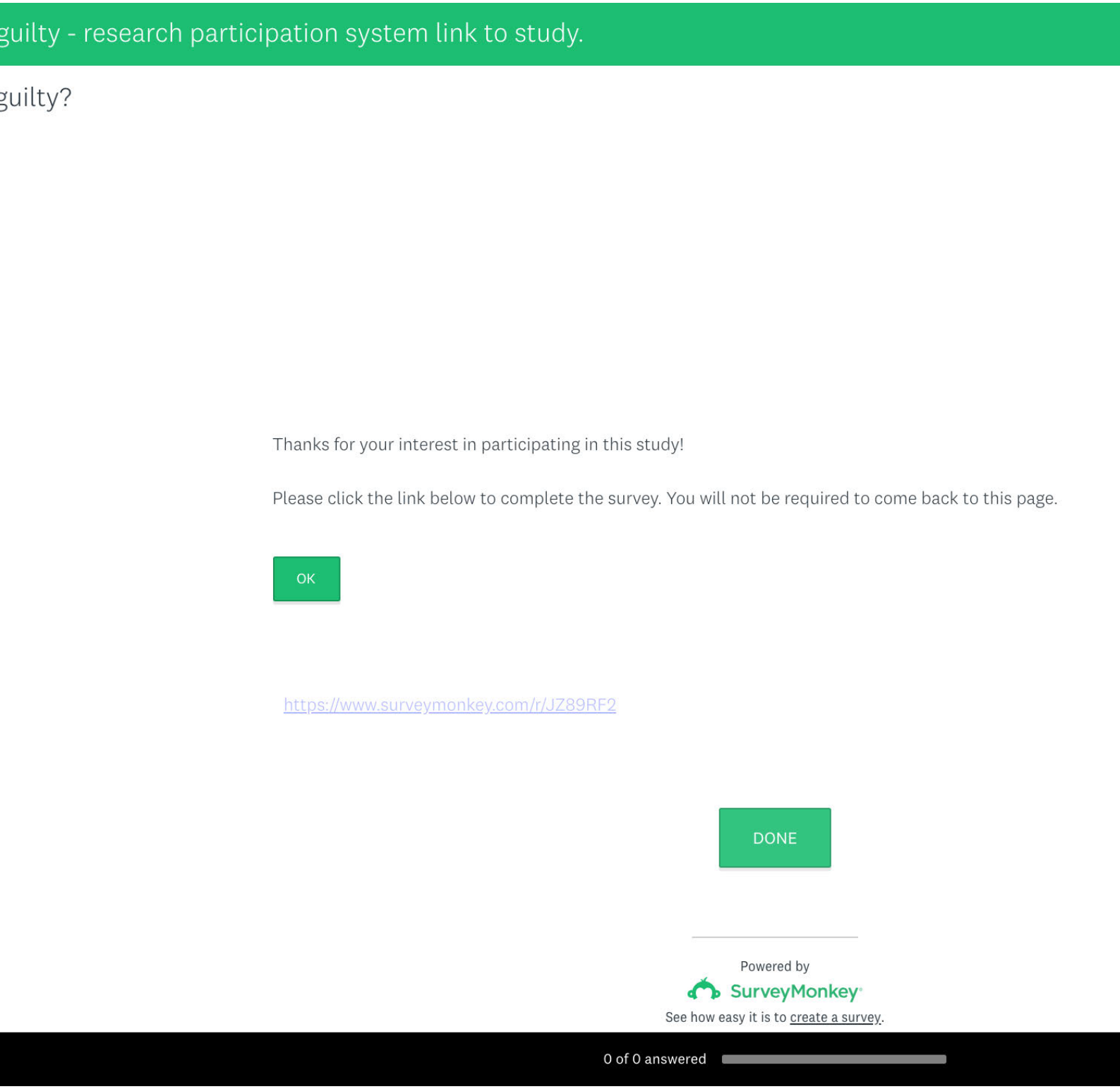
Participation in this study will mean you can enter the draw to win 1 of 3 $50 David Jones vouchers!

Contact the student researcher below!

[*Student researcher email tear strips was included at the bottom here *]*

Appendix C – Survey link to randomize

An initial Survey Monkey survey was used to randomize participants to different conditions by generating one of the 8 condition links randomly.



Appendix D – Participant preamble page and informed consent

Innocent or guilty? (#4)

Informed consent

You will be asked to answer some demographic questions and then be asked to read through a transcript of a criminal trial. The aim of this study is to learn more about the ways in which juror's reason. Since jurors are selected from the general population, you will be asked to imagine you are a juror, read a trial transcript, and then answer some questions, including rendering a verdict.

The time this study will take will depend on your reading speed, but should **not take longer than 30 minutes.**

You will not receive any benefit for participating in this study and are **free to withdraw** from the study **at any time** and there will be no consequence for not participating.

The data gathered during this study **will be made available on the Open Science Framework**. All data will be deidentified, and **only quantitative (numerical) data will be made available** on the Open Science Framework.

It may be used for future research and may be published but all information and results will be deidentified. No personal information or results will be known. If you quit the study before completion, your results will not be obtained. If you complete the study, you will not be able to retract your results, as all information will be deidentified. Therefore, there will be no way to tell which results yours are.

This study has been approved by the University of Adelaide Psychology ethics sub-committee **(Approval number:** [redacted]

If you have any questions or concerns then feel free to co[redacted] contact the supervisor of this study, Dr. Carolyn Semmler (email: Caro[redacted]

For concerns about ethical conduct contact Paul Delfabbro (email: paul.delfabbro@adelaide.edu.au)

Please be advised the transcript involves the murder of a woman who was stabbed to death. If this is likely to be distressing to you, you might want to consider not participating in this study.

1. Do you freely consent to participate in this study and understand participation is voluntary at all stages of the study?

- ◯ Yes
- ◯ No

Next

Appendix E – Survey

Demographic information

View site information

Please answer the following questions before continuing to the trial transcript.

2. What is your age?

3. What is your gender?

- ○ Male
- ○ Female
- ○ Prefer not to say
- ○ Other

4. Do you (or have you) studied any of the following:

- ☐ Law/legal studies
- ☐ Science
- ☐ Mathematics or statistics
- ☐ None of the above

5. What is your occupation?

6. Are you a current resident of Australia?

- ○ Yes
- ○ No

Prev Next

Conditions 5-8 received the instruction post-trial and did not see this at the beginning.

Bay [redacted]

You are sitting with the other jurors. Before the prosecution stand to deliver their opening address, the judge turns to you and says;

"You will hear forensic evidence concerning DNA. It will obviously form an important part of the trial, but as circumstantial evidence, you must view it in the context of all other evidence in the trial. You may hear that the chances of the DNA belonging to the defendant expressed in statistical terms from statistical studies, however, a DNA match establishes no more than the defendant could be the offender.

DNA cannot conclusively prove an individual to be a source of genetic material. It only establishes that the accused could be the source. Forensic evidence is subject to faults and is not to be attributed absolute weight of any kind of guilt or innocence. Alone, it cannot prove the guilt of the defendant beyond a reasonable doubt. No matter how persuasive, you should still carefully consider the possibility that any forensic evidence could be faulty. For example, none of the experts presenting today will assert 100% certainty over results, or conclusively tie the defendant to the crime or conclusively exculpate the defendant. **It is crucial that you understand** the following; **the probability of the DNA match is NOT the same probability of the defendant being guilty or innocent.**

By that I mean this; if the probability of the DNA occurring is 1 in 1,600, this does not mean that there is a 1 in 1,600 chance that the defendant is innocent. None of the experts will assert that the probability of a match, is equal to the probability of the defendant's guilt. A good way to do this is to take a moment after hearing each piece of evidence and consider three reasons why that evidence may not be indicative of guilt or innocence."

Prev Next

The judge continues;

"In appraising the forensic evidence, you must question if the evidence was collected and examined in a non-bias manner, or whether the evidence could have made its way to the crime scene, and still indicate innocence. For example, could the defendant's DNA have been found at the crime scene for some reason other than that they were there to commit the crime. You should continue to critically evaluate all the evidence throughout the trial and you should not begin to commit to a verdict before such time comes that you render the verdict.

As a juror, it is your responsibility to evaluate all of the evidence objectively. Further, it must be stressed, that the burden of proof lies with the prosecution. They must prove every element of the crime charged and this especially includes proving that the defendant was there to commit the crime. If you have any doubt, you must find the defendant not guilty.

This does not mean that the defendant is entitled to any far-fetched, or whimsical doubt which a quick mind might conjure up. It means that if you think that someone else other than the defendant might have reasonably committed the crime, you must find the defendant not guilty. If you are satisfied that no reasonable person would find doubt, you may find the defendant guilty."

Prev Next

You are instructed that there will now be a short break before the trial resumes. You are requested to **try** and list **three reasons** why DNA evidence may **not be indicative of guilt.**

7. List three reasons why **DNA** evidence may not indicate guilt:

Prev Next

For 5-8 this is where the survey began after the demographic questions were answered. This part remained in the same order for all conditions

Bayar v The Queen

Judge: "Mr. Bayar you have been charged with the murder of Ms. Evans on the 22nd of November 2005 under the Criminal Law Consolidation Act 1995. It is alleged you broke into her home with the intent to kill her. It is alleged you stabbed her a total of 7 times to the chest and abdomen, resulting in her death. Mr. Bayar, how do you plead?"

Mr. Bayar: "Not guilty."

The prosecution now stand to deliver their opening address.

Prosecution: "The Crown alleges that on the 22nd of November 2005 the defendant, Mr. Bayar, broke into Ms. Evans apartment and stabbed her 7 times. Mr. Bayar and Ms. Evans had a relationship which ended two years prior to her murder in 2005. Ms. Evans was a single mother. Mr. Bayar was not the father and did not enter their life until her daughter was four years of age. Ms. Evans ended the relationship and after this, Mr. Bayar stalked and harassed Ms. Evans up until the days before her death. He was angered by the break up. He sent threatening messages to her and we will show that DNA evidence, matching Mr. Bayar was found at the scene. The attack was fuelled by rage, so much so that one of the stab wounds penetrated Ms. Evans's lung. Further, the defendant told several lies to police over three consecutive interviews. There was a discrepancy in Mr. Bayar's alibi and he maintained he was 'friends' with Ms. Evans, when it is known that she went to great lengths to separate herself from him. We will prove he harassed Ms. Evans at her previous address and illegally called Centrelink to obtain her new address, all of which he denied at interview. We will leave no doubt that the defendant committed the crime of first degree murder and that he had the motive, and intent to kill."

Prev Next

You will read the following witness testimony in summary form as to read the transcript would take too long.

Neighbour: Ms. Evans neighbour testifies that she heard an argument coming from the deceased's apartment **a week before the murder**. She said it sounded like a man and a woman were fighting.

She then tells the court that the **night of the murder** she heard a high-pitched scream. She says that she remembers that she heard it around **8.20pm**, as the Harry Potter movie was on TV and was due to start at 8.30pm.

8. Please rate the strength of evidence regarding the neighbour's testimony:

Highly indicative of Mr. Bayar's innocence					Neutral					Highly indicative of Mr. Bayar's guilt
○	○	○	○	○	○	○	○	○	○	○

Prev Next

Elenor Jones: Elenor testifies that she had known Ms. Evans for two years. She says that during that time Ms. Evans cleaned her house every two weeks and they often chatted on those days. Ms. Evans had confided in Elenor and told her that the defendant Mr. Bayar had been stalking her. Elenor specifically tells the court that a few weeks prior to the murder, Ms. Evans told her that the defendant called her whilst she was shopping at IKEA asking to 'talk'. Ms. Evans told him she was busy, to which he replied 'no you are not. I know you are at IKEA.' Further, she submits that over the course of the year, Ms. Evans confided that she was being harassed. Elenor tells the court that Ms. Evans told her that over the year, she had her number plates stolen, black paint tipped on her car, her break lines cut, and a jelly like substance placed in her exhaust pipe. Elenor tells the court that Ms. Evans suspected it had been Mr. Bayar and that this was the reason that she had moved to another apartment, and not told Mr. Bayar where she was.

9. Please rate the strength of Elenor Jones testimony:

Highly indicative of Mr. Bayar's innocence					Neutral					Highly indicative of Mr. Bayar's guilt
○	○	○	○	○	○	○	○	○	○	○

Prev Next

Meg: Meg is a Police Officer who investigated some of the incidence of harassment described by Elenor Jones. She confirms that Ms. Evans had reported several incidences of harassment, and confirmed that black paint had been tipped on her bonnet, her break lines cut, and a jelly like substance in her exhaust, over the course of the year. She states that they were unable to find any evidence tying Mr. Bayar to any of the incidents.

10. Please rate the strength of Meg's testimony:

Highly indicative of Mr. Bayar's innocence					Neutral					Highly indicative of Mr. Bayar's guilt
○	○	○	○	○	○	○	○	○	○	○

Prev Next

John: John was one of the Police Officer's in charge of the investigation. He interviewed Mr. Bayar on three occasions prior to his arrest. He tells the court that Mr. Bayar told three lies. He tells you that there was a discrepancy in Mr. Bayar's alibi on the night of the murder. Mr. Bayar told police he was at home, eating dinner with his mother and his son on the couch. His mother told police that he was at home with his son eating dinner at the table. He tells you that Mr. Bayar maintained that he was friends with Ms. Evans and that he had never been to her previous address looking for her, when evidence has tied him to Ms. Evans's old apartment and there is evidence that Ms. Evans attempted to cut contact with Mr. Bayar.

11. Please rate the strength of John's testimony:

Highly indicative of Mr. Bayar's innocence					Neutral					Highly indicative of Mr. Bayar's guilt
○	○	○	○	○	○	○	○	○	○	○

Prev Next

Eruld Tunc: Eruld says he was with Ms. Evans that night, until about 8.06pm when his wife called to remind him about a party. He was having an extra marital affair with Ms. Evans. He says he left shortly thereafter and attended the party. He testifies that he said goodbye to Ms. Evans at **8.16pm** as he went to his car and from there called his wife to tell her he was on his way. The phone call and his presence at the party has been confirmed.

12. Rate the strength of Eruld's testimony:

Highly indicative of Mr. Bayar's innocence					Neutral					Highly indicative of Mr. Bayar's guilt
○	○	○	○	○	○	○	○	○	○	○

Prev Next

Asil Beckman: Asil testifies that prior to, and at the time of the murder, he worked at Centrelink. He was friends with the defendant. He testifies that Mr. Bayar called him a week prior to the murder and convinced him to use his position at Centrelink to obtain private details about Ms. Evans, including her new address which Ms. Evans had kept from Mr. Bayar. In the prosecutions' examination of Asil, they suggest that because Mr. Bayar had to illegally obtain her address through a friend and did not know her current address due to her deliberate attempt to get away from him, Mr. Bayar's DNA could not have been found at the crime scene at any other time than during the course of the murder.

13. Rate the strength of Asil's testimony:

Highly indicative of Mr. Bayar's innocence					Neutral					Highly indicative of Mr. Bayar's guilt
○	○	○	○	○	○	○	○	○	○	○

Prev Next

The prosecution tender the following documents.

Prosecution: "5 months before Ms. Evans was brutally murdered, Mr. Bayar published a poem in the *Turkish Weekly News*. The poem I am about to read has been translated to English. We submit that this poem is evidence that even over 2 and a half years after the break up, the defendant was still obsessed with the deceased, an obsession which as we will show shortly, only escalated, resulting in the brutal murder of Ms. Evans. The poem reads:

'My heart is burning for you like mad...

Even if you don't want to remember my name,

Don't want to hear my voice...

Even if you say you give up, I can not give you up

Even if you say forget, I can not forget the beautiful

Days we lived

Even if you cry all your hate, say give up,

I can not give up'

14. Rate the strength of evidence regarding the poem:

Highly indicative of Mr. Bayar's innocence					Neutral					Highly indicative of Mr. Bayar guilt
○	○	○	○	○	○	○	○	○	○	○

Prev Next

The prosecution continue;

"I will now read to you a series of communications from the 11th of November, ten days before the murder, between Ms. Evans and Mr. Bayar. At 1am on the 11th of November Mr. Bayar send a text message to the deceased saying;

'you are sad and crying and I feel it.'

It so happens that night that Ms. Evans had had an argument with her father. We submit to you, that Mr. Bayar knew she was crying as he was in the vicinity of her home that night. 20 minutes after this first message, Mr. Bayar send a message saying;

'I'm sure because I asked about your number plates you are blaming me for it.'

And shortly after;

'Alright, understood, there's a god above. What can I say. I pray for your health.'

We submit that the first message is evidence for his confession that he had harassed the deceased prior to the murder, by tampering with her number plates, and submit that the second text 'I pray for your health' was a threat. At nearly 2am this same morning, 6 phone calls wer made by Mr. Bayar to Ms. Evans. The content of the phone calls is unknown; however, we submit that these communications illustrate the defendant's obsessive and threatening behaviour towards the deceased."

15. Please rate the strength of evidence regarding the communication between Mr. Bayar and Ms Evans:

Highly indicative of Mr. Bayar's innocence					Neutral					Highly indicative of Mr. Bayar's guilt
○	○	○	○	○	○	○	○	○	○	○

Prev Next

The prosecution introduce their last witness.

Prosecutor: "Can you please state your name and occupation for the record?"

Pineda: "Dr. Seleena Pineda, I am Technical Leader and Associate Director at Cellmake, a private DNA company based in Texas, USA."

Prosecutor: "Thank you. Can you please describe to the jury, the piece of evidence which was given to you for analysis?"

Pineda: "Of course. It was a small body-hair, on which I conducted a mitochondrial DNA analysis."

Prosecutor: "And where had the small-body hair come from?"

Pineda: "It was stuck to the dried blood on the deceased's fingernail."

Prosecutor: "And, broadly, how did conduct your analysis on the DNA?"

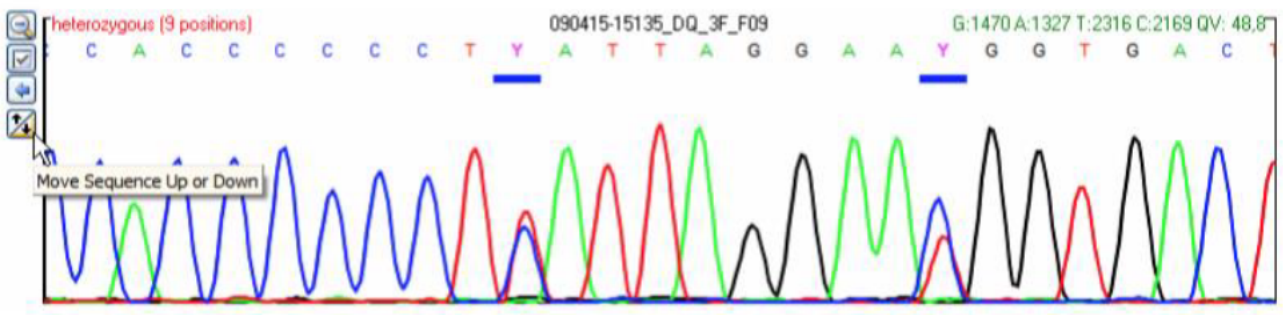
Pineda: "I compared the sample of the body-hair found at the crime scene with a buccal swap of the defendant – and then ran the DNA type through the SWGDAM database, which is a North America database."

Prosecutor: "Can you elaborate for the jury, what is meant by a buccal swab?"

Pineda: "A buccal swab is a sample of saliva, taken from a person's mouth."

Prosecution: "And can you give a bit more detail for the jury, so they can understand what DNA actually is, how it is analysed and interpreted?"

Pineda: "DNA is often referred to as the 'genetic blueprint' of a person. This blueprint is unique for each person. All cells contain DNA. As such, the body-hair found at the crime scene contained DNA, and forensic experts like myself use this to conduct a comparison between the DNA found at the scene and a sample of the suspects DNA. This is done by looking at an electropherogram. Here is a **just example of an electropherogram not the electropherogram used in Mr. Bayar's comparison in this trial**.... but hopefully it can make clearer the comparison I conducted:"



Pineda: "On the diagram we can see the different alleles; a DNA comparison involves a comparison of two electropherograms where I look to see how the loci on the two samples measure up to determine a match."

Prosecutor: "And in comparing the buccal swab taken from the defendant, Mr. Bayar, with the body-hair found at the crime scene what did you find?"

Pineda: "The DNA profile taken from the hair is consistent with the defendant's DNA profile on the buccal swab. In reaching this conclusion based on my interpretation of the electropherogram, I discounted a portion of the defendant's DNA profile not found on the hair. There is a chance of variability in relation to that component of the profile, but it is conventional to disregard this, and, in these circumstances, it would be considered a perfect match."

Prosecution: "Did you compare this body-hair with any other DNA samples?"

Pineda: "Yes, I compared the hair with Eruld Tunc, who I believe can be excluded as the source of the hair; it did not match the DNA profile of the hair."

Prosecution: "Was there any chance that this sample was contaminated at any point?"

Pineda: "It is highly unlikely. Given the nature of the forensic evidence, it was treated with care – there was an audit trail on who had been in contact with the DNA and so far, as I can say, this evidence could not have been contaminated."

Prosecution: "...so how likely is it that the hair found at the scene, matches the DNA of the defendant?"

Pineda: "Well, firstly, the DNA profile in the hair and in the sample taken from Mr. Bayar occurs 1 in 4, 839 people. However, by implementing a standard confidence interval (95%) I can conclude that 1 in 1,600 people in the world would be expected to share the DNA profile found in the hair. In other words, 99.9% of people in the general population would not be expected to have a DNA profile matching that of the hair."

Prosecution: "...and this figure of 1 in 1600, is it the upper limit or the lower limit?"

Pineda: "The frequency ratio of 1 in 1, 600 people is the upper limit. That is to say, this is the figure which gives the most benefit to the defendant."

Prosecution: "Thank you Ms. Pineda. No further questions."

16. Please rate the strength of Ms Pineda's forensic evidence:

Highly indicative of Mr. Bayar's innocence					Neutral					Highly indicative of Mr. Bayar's guilt
○	○	○	○	○	○	○	○	○	○	○

Prev Next

Bayar v The Queen
Prosecution closing

"We submit the following occurred the night of the murder;

Mr Bayar, becoming increasingly obsessed with Ms Evans, witnessed her lover, Eruld Tunc entering and leaving her home that night. It is no coincidence that it was only a matter of minutes after her lover Eruld left the apartment she was murdered. Mr. Bayar witnessed her affair that night, and the building obsession reached a boiling point. He was unable to contain his rage. He threatened her, and then, he followed through on his threat. I urge you to think carefully about the evidence we have presented; the increasing pattern of obsession two years after the break up. The lengths Ms. Evans went to avoid the accused. Mr. Bayar's DNA at the crime scene, stuck directly to Ms. Evans, by the blood on her finger, which does not match Eruld Tunc. There is no reason for which Mr. Bayar's DNA could have found its way to the scene, but in the course of the murder. These pieces of evidence are not coincidental, and I urge you to find justice here today."

Prev Next

For groups receiving the defence DNA first:

Bayar v The Queen
Defence forensic evidence

For groups receiving defence DNA last it read "defence call their last witness"

The defence call their first witness.

Defence: "Can you please state your name and occupation for the record?"

Ben Buckleton: "My name is Dr. Ben Buckleton and I am a scientist, employed by the New Zealand Government."

Defence: "Thank you. Can you start by telling us about the hair found at the crime scene?"

Ben Buckleton: "It was a small body-hair stuck to the dried blood on the fingernail of the deceased."

Defence: "Is there any way for an expert to establish how old a fragment of DNA is... to rephrase were you able to establish whether the body hair made its way to the scene at the time of the murder, or could it have been in the vicinity for some weeks or months prior to its examination?"

Ben Buckleton: "There is no way to tell how or when precisely the DNA came to be at the scene of the crime."

Defence: "And now can you tell us about the analysis you conducted on the body hair found at the crime scene and the sample taken from my client?"

Ben Buckleton: "Certainly. The analysis I conducted followed the same procedure as that used by Ms. Pineda from the prosecution. My interpretation of the results is that the DNA profile found in the hair might be found in 1 in 1,000 people in the non-Turkish population. However, in the Turkish population, this DNA type could be found between 1 in 50 to 1 in 100 people."

Defence: "So, the probability of a DNA match is influenced quite heavily by the database from which is compared with – is that correct?"

Ben Buckleton: "Yes, in my opinion, that is certainly correct. Further – I disagree with the evidence presented by the prosecution... in my view the discrepancy between the DNA profile of the accused, and the DNA profile found in the blood stuck to the victim's thumb is notable and indicates a genuine difference in profile.... Therefore, I would not treat these as a full match."

Prev Next

Defence: "So, to confirm you mean to say that 1 in 50 people could have this DNA type – so if we have a football stadium of two Turkish teams playing each other in football and ten thousand people attend.... Within that football stadium of ten thousand people there, in your view, would be on average, two hundred people who have the same DNA type as my client? But this would be a stretch as in your opinion they are not a full match?"

Ben Buckleton: "Yes, that is correct. As I stated that is conservative."

Defence: "So, then it is entirely possible that another man other than my client may be the source of the DNA?"

Ben Buckleton: "Yes, it is possible. I suppose there if around a 1 in 50 chance in the Turkish population. In the general population that chance would increase to closer to 1 in 1000. But in the Turkish population, that chance is much smaller. So, I suppose it depends on what proportion of the suspects are Turkish."

Defence: "How many source samples did you receive?"

Ben Buckleton: "I received two; one was the small body hair stuck to the blood on the deceased's fingernail, the other source was a sample of saliva taken from the deceased's chin."

Defence: "Did you run an analysis on the second source – the saliva found on her chin?"

Ben Buckleton: "Yes, I did. That DNA matched that of Eruld Tunc."

Defence: "Is there a chance there could have been a contamination or mix up of the DNA?"

Ben Buckleton: "There is always a chance for contamination. So yes, it is possible."

Defence: "No further questions, your Honour."

17. Please rate the strength of Buckelton's DNA evidence:

Highly indicative of Mr. Bayar's innocence					Neutral					Highly indicative of Mr. Bayar's guilt
○	○	○	○	○	○	○	○	○	○	○

Prev Next

Groups not receiving DNA first began:

The defence calls the following witnesses, which you will read in summary form, as to read in transcript would take too long.

Mark Marlow: Mark tells you he works at the local mechanics and has been a fully qualified mechanic for ten years. He tells the court that Mr. Bayar had brought his car in four days before the murder. Mark says he inspected the car and found significant damage to the exhaust and engine. Mark tells you that in his opinion the car was drivable but would have been making a considerable amount of noise. He tells the court that in his opinion, if Mr. Bayar had been in the vicinity of Ms. Evans apartment that night, someone would have heard his car.

18. Please rate the strength of evidence regarding Mark's testimony:

Highly indicative of Mr. Bayar's innocence					Neutral					Highly indicative of Mr. Bayar's guilt
○	○	○	○	○	○	○	○	○	○	○

Prev Next

Police Officer John is recalled by the defence.

John: John says that Mr. Bayar cooperated at every point during the investigation. He states there was no blood found in Mr. Bayar's car or home which he allowed police to search without a warrant. The defence question John about any reports of a noisy car in the area that night. John says that none of the neighbours he interviewed heard a car that night. The prosecution stress that there were no reports of noise in the area that night. They argue this indicates the defendant was not in the area that night and could not have been there to commit the crime.

19. Please rate the strength of evidence regarding John's testimony:

Highly indicative of Mr. Bayar's innocence					Neutral					Highly indicative of Mr. Bayar's guilt
○	○	○	○	○	○	○	○	○	○	○

Prev Next

Eruld Tunc: Eruld admits to having an affair with Ms. Evans, which had remained secret until a week prior to the murder. The defence question Eruld Tunc about a statement made by Ms. Evans's daughter, disclosed during the course of the investigation, that Eruld borrowed $8,000 from Ms. Evans for a gambling debt. Eruld replies that this was a lie created by himself and Ms. Evans to explain his presence at her apartment and to cover for their affair. He states he never borrowed any money from Ms. Evans.

Eruld testifies that he was at Ms. Evans apartment at around 8pm the night of the murder. At 8.06pm, he received a call from his wife. Phone records confirm this. At 8.16pm he states he left the house and attended a party with his wife. His presence at the party shortly thereafter has been confirmed. Eruld admits to returning to the apartment later that evening at 12.30am. He states that he returned as he had not been able to reach Ms. Evans, and had become worried. The defence question Eruld Tunc about the argument which may have occurred a week prior to the murder. Eruld maintains that he does not remember having an argument.

20. Please rate the strength of evidence regarding Eruld's testimony:

Highly indicative of Mr. Bayar's innocence					Neutral					Highly indicative of Mr. Bayar's guilt
○	○	○	○	○	○	○	○	○	○	○

Prev Next

Mr. Evans (the deceased's father): It comes to light that on the 11th of November, the same night that Ms. Evans received text messages from the defendant, Ms. Evans had been crying as her father had hit her upon discovering that she was having an affair with Eruld.

21. Please rate the strength of evidence regarding Mr. Evans testimony:

Highly indicative of Mr. Bayar's innocence										Highly indicative of Mr. Bayar's guilt
					Neutral					
○	○	○	○	○	○	○	○	○	○	○

Prev Next

Bayar v The Queen

Ayla Jones: Testifies that she has been and at the time of the murder in a relationship with Mr. Bayar. She states at the time of the murder they were 'close' and he spent 3-4 nights a week at her house.

22. Please rate the strength of Ayla's testimony:

Highly indicative of Mr. Bayar's innocence										Highly indicative of Mr. Bayar's guilt
					Neutral					
○	○	○	○	○	○	○	○	○	○	○

Prev Next

Defence: "My client has loved poetry for most of his life. Further, in Turkish culture, it is normal and acceptable to express high emotion. We submit that the poem published in the *Turkish Weekly News*, was not a threat, but simply a signal of my clients lingering affection. Mr. Bayar was in a new relationship and spent 3 to 4 nights a week with his new partner. He was at home the night of the crime with his mother. The small discrepancy is a result of the fallibility of memory. My client was interviewed at three times over the months after Ms. Evans's death and spoke freely with police without the assistance of a lawyer. Over time small details of every day monotony can become forgotten. At every step during the investigation my client cooperated. He allowed police to search his car and house without a warrant, and this is because my client has nothing to hide. There was no blood found at either premise. His car was not heard in the area that night. This is because it was not my client who killed Ms. Evans. I ask you to think about the fact that Ms. Evans was having a secret affair with a married man. A married man whose DNA was also found at the crime scene. Her own father was angered by this discovery. Further Eruld, her lover conveniently left the scene only minutes before her murder. Her screams were heard, but my client was not heard in the area; this is no coincidence. That is because my client was at home that night, with his mother and his son. I urge you to consider the evidence and find my client not-guilty."

Please answer the following questions about the trial transcript you have just read.

23. Ms Pineda testified for the **prosecution**. Did she say that:

- ○ That 99.9% of the general population would not be expected to match this DNA
- ○ 1 in 50 people in the Turkish population would have this DNA type

24. Ms. Evans (the deceased) was:

- ○ Stabbed to death
- ○ Hit over the head with a hammer
- ○ Drowned in a swimming pool

25. A mechanic testified for the defence and stated that:

- ○ Mr. Bayar's car was perfectly fine
- ○ There was a problem with Mr. Bayar's car which would have caused lots of noise

Prev Next

Innocent or guilty? (#4)

Experimental responses

You will now be asked to re-rate the evidence in the trial. Please rate the following pieces of evidence.

26. Please rate Ms Pineda's forensic evidence (DNA expressed as 99.9% population excluded from having this DNA OR 1 in 1,600 expected to have the DNA type):

Highly indicative of Mr. Bayar's innocence					Neutral					Highly indicative of Mr. Bayar's guilt
○	○	○	○	○	○	○	○	○	○	○

27. Rate the strength of Asil Beckman's testimony (illegally obtained Ms. Evans address from Centrelink on behalf of Mr. Bayar):

Highly indicative of Mr. Bayar innocence					Neutral					Highly indicative of Mr. Bayar guilt
○	○	○	○	○	○	○	○	○	○	○

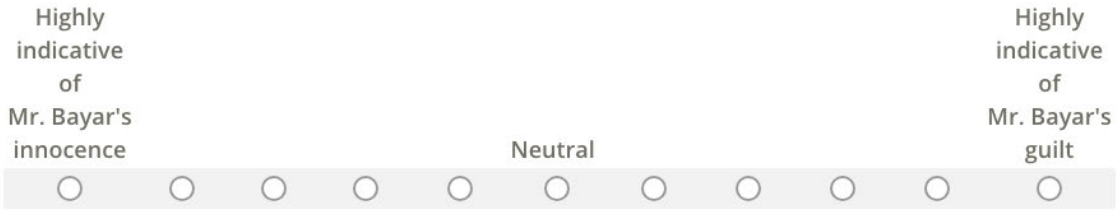
28. Rate the strength of Eruld Tunc's testimony (Ms. Evans lover, at the apartment the night of the murder but left at 8.16pm, denied borrowing money or having an argument):

Highly indicative of Mr. Bayar's innocence					Neutral					Highly indicative of Mr. Bayar's guilt
○	○	○	○	○	○	○	○	○	○	○

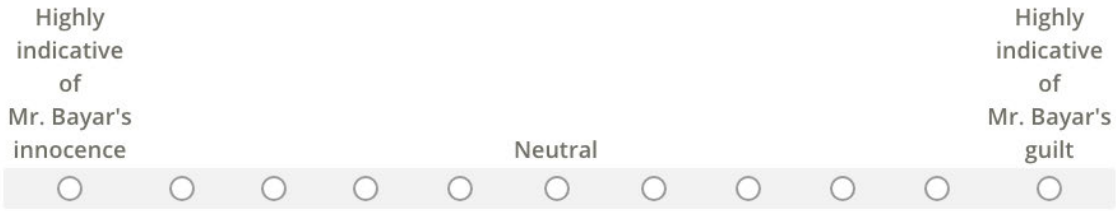
29. Please rate the strength of Police Officer John's testimony (stated Mr. Bayar lied at interview, but otherwise cooperated at all points in investigation):

Highly indicative of Mr. Bayar's innocence					Neutral					Highly indicative of Mr. Bayar's guilt
○	○	○	○	○	○	○	○	○	○	○

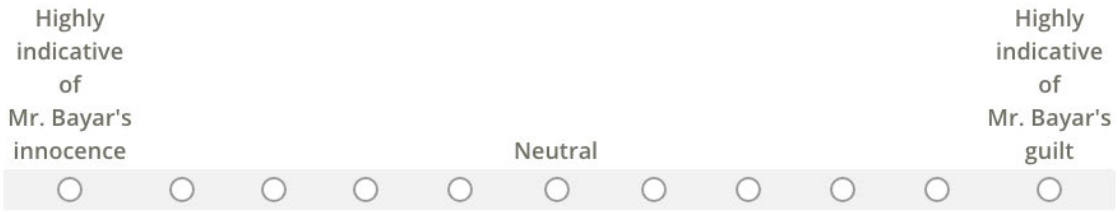
30. Please rate the strength of Police Officer Meg's testimony (confirmed incidents of harassment, but unable to link Mr. Bayar to any of the incidents):



31. Please rate the strength of Elenor Jones testimony (that Ms. Evans told her Mr. Bayar was stalking and harassing her):

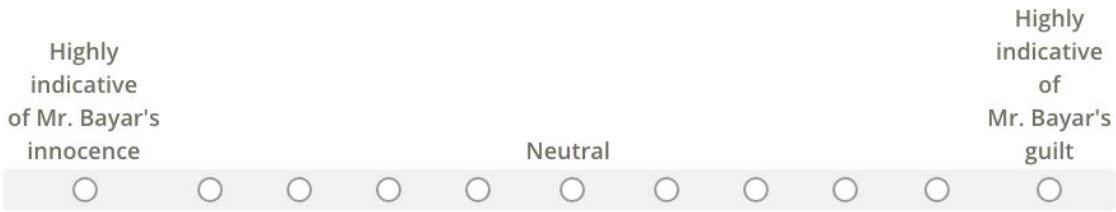


32. Please rate the strength of the neighbour's testimony (heard a scream at 8.20pm):



33. Rate the strength of evidence regarding the poem:

34. Please rate the strength of evidence regarding the communication between Mr. Bayar and Ms Evans (text messages and phone calls):



35. Please rate the strength of Ben Buckelton's DNA evidence (1 in 50 people expected to have this DNA type):

Highly indicative of Mr. Bayar's innocence					Neutral					Highly indicative of Mr. Bayar's guilt
○	○	○	○	○	○	○	○	○	○	○

36. Please rate the strength of Mark's testimony (the mechanic who stated the car would have made lots of noise):

Highly indicative of Mr. Bayar's innocence					Neutral					Highly indicative of Mr. Bayar's guilt
○	○	○	○	○	○	○	○	○	○	○

37. Please rate the strength of Mr. Evans's testimony (that he hit his daughter a week prior to the murder):

Highly indicative of Mr. Bayar's innocence					Neutral					Highly indicative of Mr. Bayar's guilt
○	○	○	○	○	○	○	○	○	○	○

38. Please rate the strength of Ayla's testimony (that at the time of the murder she was in a relationship with Mr. Bayar):

Highly indicative of Mr. Bayar's innocence					Neutral					Highly indicative of Mr. Bayar's guilt
○	○	○	○	○	○	○	○	○	○	○

Prev Next

Verdict

39. How probable do you think it is that Mr. Bayar is guilty?

0 100

Prev Next

Verdict

40. How do you find Mr. Bayar?

- ☐ Guilty
- ☐ Not guilty

Prev Next

Verdict

41. How confident are you in your decision?

Not at all confident					Not sure					Very confident
○	○	○	○	○	○	○	○	○	○	○

42. Please consider the two ways that Ms Pineda stated the results of the DNA test:

There was a 99.9% chance that the DNA belonged to Mr. Bayar.

There was a 1 in 1,600 chance that someone other than Mr. Bayar committed the crime.

Which of the following statements best reflects your belief that these statements are true?

- ○ The first statement is most likely to be true
- ○ The second statement is most likely to be true
- ○ Both statements are equally likely to be true
- ○ Neither statements are likely to be true

Prev Next

Optional

If you would like, please write a sentence describing the reasoning for your verdict i.e. the key pieces of evidence which contributed to your decision/missing information you would have liked.

43. Reasoning:

End of study

Please enter your email address below if you would like to be contacted with the results of this study or if you wish to go in the draw to win 1 of 3 $50 David Jones vouchers.

Your email address with be separated from the rest of your responses, and **will only be used to to inform you if you have won the voucher, or with the results.**

44. If you wish to be contacted with the results of this study:

- ○ No, I do not wish to be contacted
- ○ Yes, I wish to be contacted with the results of this study (please enter email in the box below)

45. If you wish to go in the draw to win 1 of 3 $50 David Jones vouchers:

- ☐ No, I do not wish to go in the draw
- ☐ Yes, I wish to go in the draw (please enter email in the box below)

46. For Psychology students requiring course credit, please enter your unique RPS code below:

Thank you for completing my study!

The results of this study will be used to inform research on juror comprehension of DNA evidence. The facts of this case were taken from an Australian case in 2012 'Aytugrul v The Queen' where Mr. Aytugrul the defendant was found guilty. However, the use of DNA in this trial has been highly contentious in the legal community.

Prev Done

Appendix F – Tests for normality hypothesis 1

Shapiro-Wilk test for normality and Levene's test for non-parametric data (Nordstokke & Zumbo, 2010)

Condition	Shapiro-Wilk			
	PDNAr1	CSPDNA	DDNAr1	CSDDNA
Group 1	.869, *p*=.001	.865, *p*=.001	.916, *p*=.011	.921, p=.015
Group 2	.789, *p*=.000	.802, *p*=.000	.919, *p*=.018	.786, p=.000
Group 3	.842, *p*=.000	.917, *p*=.018	.896, *p*=.005	.827, p=.000
Group 4	.824, *p*=.000	.786, *p*=.000	.902, *p*=.013	.905, p=.015
Group 5	.816, *p*=.000	.872, *p*=.001	.957, *p*=.215	.902, p=.006
Group 6	.834, *p*=.000	.847, *p*=.000	.935, p=.061	.964, p=.371
Group 7	.835, *p*=.000	.886, *p*=.003	.899, p=.007	.887, p=.003
Group 8	.837, *p*=.000	.776, *p*=.000	.920, p=.017	.872, p=.001
Non-parametric Levene's test	F(7,249)=2.309, p=.027	F(7,249)=.763, p=.619	F(7,249)=2.318, p=.026	F(7,249)=1.916, p=.068

Note. PDNAr1=initial rating of prosecution DNA evidence. DDNAr1=initial rating of defence DNA evidence. CSDDNA=coherence-shift defence DNA. CSPDNA=coherence-shift prosecution DNA.

Appendix G – Tests for normality Hypothesis 4

Shapiro-Wilk test for normality and Levene's test for homogeneity of variance for confidence ratings.

Condition	Shapiro-Wilk test
Group 1	.945, *p*=.**080**
Group 2	.915, *p*=014
Group 3	.894, *p*=.004
Group 4	.915, *p*=026
Group 5	.965, *p*=.**353**
Group 6	.905, *p*=.010
Group 7	.912, *p*=015
Group 8	.914, *p*=.011
Non-parametric Levene's test	*F* (7,249)=1.443, *p*=.189

Appendix H – Tests for normality for Hypothesis 2

Shapiro-Wilk test for normality and non-parametric Levene's test for testimonies following DNA evidence.

Condition	"Neighbours testimony" ratings	"Mechanics testimony" ratings
Group 1	.543, *p*=.000	.697, *p*=.000
Group 2	.641, *p*=.000	.903, *p*=.006
Group 3	.474, *p*=.000	.837, *p*=.000
Group 4	.599, *p*=.000	.863, *p*=.002
Group 5	.742, *p*=.000	.819, *p*=.000
Group 6	.595, *p*=.000	.742, *p*=.000
Group 7	.578, *p*=.000	.766, *p*=.000
Group 8	.713, *p*=.000	.672, *p*=.000
Non-parametric Levene's test	*F* (7,249) =.609, *p*=.748	*F* (7,249) =1.217, *p*=.294

Appendix I – Preliminary results email sent to participants

Preliminary results for the study 'Innocent or guilty?' Inbox ×

16:14 (3 minutes ago)

to bcc: s

Hi there,

Firstly, I'd like to thank you for your time in participating in the study 'Innocent or guilty?'

You are receiving this email as you expressed interest in being contacted with the results of this study. This email address will no longer be used for any future purpose and will be deleted.

The purpose of the study

As DNA evidence increasingly forms part of criminal trials, this study was specifically interested in it's impact. Although it has the potential to exonerate individuals of crimes, it is often used in circumstantial cases and is highly prejudicial. Perceived to be objective and reliable, DNA testing still involves a level of subjective judgement (if interested in how DNA testing is subjective see a great TED talk at - https://www.youtube.com/watch?v=tpPkmDeS3Dg)

The purpose of this study was therefore to investigate the effect of pre-trial vs. post-trial judicial instruction instruction alongside the impact of DNA evidence on the side of the prosecution and the defence being delivered at different times on assignments of probative weight. A judicial instruction is a direction from the judge to the jury about how to evaluate a case. Previous literature indicates that post-trial instruction is ineffective, and pre-trial instruction can decrease conviction rates by up to 20%.

Further, research on decision-making has shown that a robust phenomena known as coherence-shifting occurs when people are making decisions, whereby people polarise and change their weightings of different facts to arrive at a decision. It is a form of confirmatory reasoning. It was thought that DNA evidence may trigger a spreading coherence whereby following evidence would be inflated to support a decision; making it more prejudicial than DNA evidence given at the end of a trial.

The results

1. DNA evidence did not inflate subsequent evidence.
2. There was a recency effect; jurors receiving the DNA evidence last rated it significantly more highly than jurors receiving it at the beginning. This was for both the prosecution and the defence.
3. DNA given early resulted in greater coherence-shifts for the prosecution but not for the defence.
4. Pre-trial instruction DID decrease initial ratings of prosecution DNA evidence compared to post-trial instruction.
5. The difference in secondary ratings across groups for the DNA was not significant.

The results indicate that timing of DNA results in a different process of assessment however, jurors arrive at approximately the same weightings at the end of the trial. The pre-trial instruction is believed to have had a bigger impact on jurors verdicts outlined below.

Your participation contributes to an honours thesis and there is potential for these findings to be used in journal papers and other experiments in the future.

About the experiment:

You would have been allocated to 1 of 8 experimental conditions. Depending on this you would have received either a post-trial instruction from the judge, or the same instruction pre-trial; and you would have received prosecution DNA evidence and opposing defence DNA evidence.

The conviction rates and associated probability of guilt for each condition are outlined below. A total of 257 participants completed the study and all received the same information but at different time points.

Distribution of verdict across condition

Group	Guilty	Not guilty	Av. probability of guilt
1(JI-P1-D1)	32%	68%	60
2(JI-P2-D2)	44%	56%	63
3(JI-P1-D2)	15%	85%	49
4(JI-P2-D1)	14%	86%	52
5(P1-D1-JI)	33%	67%	50
6(P1-D2-JI)	32%	68%	65
7(P2-D1-JI)	29%	71%	61
8(P2-D2-JI)	44%	56%	61

I hope this is of interest to you. Once again, thank you for participating.

Kind Regards,

, ,Student Researcher; The University of Adelaide), Dr Carolyn Semmler (Researcher; The University of Adelaide).

[i] Also described as inference to the best explanation (Amaya, 2007); with subtractive (incoherent items removed), additive (bolstering facts) and reinterprative processes (revising evidence) to reach coherence.

[ii] Only groups with one change in acronym may be compared (i.e. Group 4 (*JI*-*P2*-D1) and Group 5 (*P1*-D1-*JI*) cannot be compared two things change – the instruction and the prosecution evidence).

[iii] Comparison = comparison uploaded on the Open Science Framework.