Running Head: Confirmation bias in jurors.

Title: Faith in thy Threshold.

Lee J. Curley[1*], Jennifer Murray[2], Rory MacLean[1], Phyllis Laybourn[1], & David Brown[3]

1. School of Applied Sciences, Psychology Subject Group, Edinburgh Napier University, Edinburgh, Scotland, EH10 4BN.
2. School of Health and Social Care, Edinburgh Napier University, Edinburgh, Scotland, Mental Health Theme.
3. Currently not affiliated with any Universities.

Word count (exc. tables/references): 5755.

*Requests for reprints should be addressed to Lee Curley, School of Applied Sciences, Psychology Subject Group, Edinburgh Napier University, Edinburgh, Scotland, EH10 4BN. (email: Lee.Curley@napier.ac.uk)

Abstract

The current study focussed on the decision making processes of jurors. The study investigated how jurors make a decision, if they integrated information within their decision making process, and if cue utilisation thresholds promoted confirmation bias. To do this, 108 participants listened to one of nine cases. These participants were asked to give a likelihood of guilt rating after each piece of evidence, to state what was the last piece of information they needed to make a decision and give a final verdict at the end of a trial. The results highlighted that threshold decision making was being utilised, that information integration may allow thresholds to be reached and that thresholds may promote confirmation bias to reduce cognitive dissonance. In conclusion, this suggests that jurors integrate information until they reach a leading verdict, then the evaluation of information is distorted to support the leading threshold. Implications relate to von Dire and legal instructions.

Keywords: Decision Science; Confirmation Bias; Information Integration; Diffusion Threshold Model; Juror Decision Processes.

Acknowledgements: This research was carried out in partial fulfilment of the first author's PhD. The Mcglashan Charitable trust provided the lead researcher with funding to help pay tuition fees when this research was being conducted. The study was awarded ethical approval from Edinburgh Napier University's School of Applied Sciences' Research Integrity Committee in 2017. The authors have no competing interests to declare.

Guarantor: Lee Curley will act as the guarantor for this work, accepting responsibility for the conduct of the study, holds access to the data, and controlled the decision to publish, though all study authors were involved in the preparation of the manuscript, interpretation of the data, had access to the data, and informed the study design.

Statement of Contributions: The conceptualisation for the study initially arose from Lee Curley, with the remaining authors assisting in shaping the specific aim and design of the research. Lee Curley collected and analysed the data. Lee Curley, Rory MacLean and Jennifer Murray interpreted the data. Lee Curley prepared the first and final drafts for publication; Jennifer Murray contributed to re-drafting and responding to reviewer feedback; and Rory MacLean and Phyllis Laybourn assisted in preparing drafts of the manuscript. Mr David Brown helped design the researching materials, and was crucial in allowing the study to be conducted.

Faith in thy Threshold

The distortion of evidence by jurors could have a disastrous impact on both individuals and society. Evidence distortion can be brought about by a specific cognitive bias called confirmation bias [1]. Confirmation bias relates to individuals perceiving evidence that confirms a belief positively and disconfirming evidence negatively [1]. Confirmation bias has been cited as having a negative influence in the courtroom [2, 3]. David Camm was convicted twice for murdering his wife and children. After 13 years in prison, he was finally released after an appeal found that the evidence from the prosecution in the previous two trials should have been excluded [2, 3]. It has been suggested that confirmation bias caused the innocent David Camm to be falsely imprisoned [2, 3]. The reasoning behind confirmation bias being charged for David Camm's injustice is because Camm's charges were not dropped despite the prosecution's evidence being found to be unreliable/inaccurate [2]. His charges remained, and a trial commenced based on the beliefs of the prosecutors rather than the facts.

It is therefore evident that injustice may be an outcome of a distorted decision making process, and this may be caused by confirmation bias [4, 5]. Confirmation bias prevents jurors, police officers, and judges from using all of the available evidence within a case, thus increasing the chances of an error occurring [4]. Injustices have a number of significant and negative impacts on the legal system. First, an innocent individual has been failed in being given their right to a fair defence. Second, a victim has been failed in their right to justice and retribution. Third, a guilty perpetrator is still free to recidivate. Finally, the criminal justice system is failing in their obligation to society of deterring potential criminals, rehabilitating offenders, and incarcerating offenders [6]. Confirmation bias undermines justice and the aims of the legal system.

Confirmation bias occurs when a decision maker searches for information or interprets information in a way that supports their initial hypotheses, predictions, and attitudes [4, 1]. This

bias can skew information search in naturalistic decision making tasks where the decision maker can choose which pieces of information they want to use [1].

Confirmation bias may cause jurors to perceive information that supports their initial decision/threshold and/or belief more positively than pieces of evidence that contradicts their view [1]. A related construct to confirmation bias, pre-decisional distortion, may also occur in the courtroom. This is where information is skewed to favour prior biases [7], thereby strengthening the effects of confirmation bias. Confirmation bias may also be exaggerated in jurors as it has been found that the outcome that is favoured does not distort pre-decisional search, but does distort pre-decisional evidence evaluation [8].

One piece of research from Estrada-Reynolds, Gray, and Nuñez [7] found that prior beliefs (pro-prosecution, pro-defence and neutral) could forecast the final sentence recommendations given by jurors. This highlights that novel evidence is distorted so that it is more in line with the leading verdict. Carlson and Russo [9] showed that mock jurors distorted information from a civil trial, and that the prior beliefs of the juror (either pro-plaintiffs or pro-defendants) influenced the verdicts they finally gave. Interestingly, confidence increased with distortion, thus highlighting that evidence distortion and confirmation bias allowed jurors to believe that their leading verdicts were the most appropriate verdicts to give [9].

One potential explanation for confirmation bias relates to cognitive dissonance [10]. Cognitive dissonance occurs when the beliefs that a decision maker has and the information (i.e., the new competing belief system) that is being supplied to the decision maker contradict one another [10]. This variance between the original belief and the novel information may cause psychological and physiological discomfort [11]. To reduce this discomfort, the decision maker may try to reduce the cognitive dissonance [10]. The most 'cognitively easy' way to do this is to stop searching for disconfirming information, and to focus information search on confirming

evidence [10]. However, this would not be possible within a courtroom. Instead, jurors may seek to reduce their dissonance through viewing confirming evidence more positively than disconfirming evidence. Similarly, they may attend more to confirming evidence when making a decision and when within the deliberation room [12].

Some research has found, however, that disconfirming evidence is used when jurors are making decisions. Estrada-Reynolds et al.[7] aimed to test whether or not death sentence recommendations were distorted throughout a capital punishment case in accordance with pre-trial biases (pro-prosecution, pro-defence or neutral); they did this by asking participants to give sentence recommendations (one = more likely to sentence life; and seven = more likely to sentence death) at eight separate points in time. Estrada-Reynolds et al.[7] found that that sentence recommendations were changed in the appropriate direction, sentence recommendations increased with prosecution evidence and decreased with defence evidence, which does suggest that jurors may integrate information throughout a trial.

Information integration theory proposes that juror interpretations of the information presented in the courtroom can be described by a weighted average of a pre-trial belief alongside the probability of guilt inferred by every piece of evidence [13, 14]. In this theory, belief of guilt is adjusted after each piece of information [13, 14], jurors do not ignore disconfirming evidence [7], and through evaluating and integrating evidence a juror is able to make a decision [15]. Ostrom Werner, and Saks [16] found that jurors start with an innocent-until-proven-guilty belief and that said belief is integrated with trial evidence to produce the final verdict. Kaplan and Kemmerick [17] discovered that both legal and extra-legal (e.g., gender of the defendant) pieces of information allow a unitary evaluation of which verdict is most appropriate. Kaplan and Miller [13] found evidence for the information integration theory through showing that the influence that pre-trial biases have could be reduced by increasing the reliability of the information

presented in court, thus highlighting that a weighted averaging of prior beliefs and the perception of guilt gained from the evidence had occurred.

In contrast to confirmation bias, the information integration theory would suggest that jurors do not ignore or distort information and that information integration occurs, which then influences the decision [7]. Previous research has found support for both confirmation bias/evidence distortion and information integration occurring in jurors [7, 9]; therefore, a model of decision making that could explain both of these respective information evaluation methods would hold great utility.

One recent explanation for the decision making processes of jurors could be explained using a unified Diffusion Threshold Model of decision making [18]. This model allows both rational and intuitive processes to be encompassed by having a varying threshold; high thresholds are equivalent to rational processing, while low thresholds are equivalent to frugal processing [19]. In the Diffusion Threshold Model, the decision maker reaches their relative threshold through information integration [20]. Therefore, the information integration theory fits well with the Diffusion Threshold Model in that both suggest that information integration allow a decision to be made.

Previous research that supports the information integration theory within juror decision making, such as Kaplan and Miller [13], also supports the Diffusion Threshold Models use of information integration to explain how thresholds are reached. However, the Diffusion Threshold Model discussed differs from the information integration theory in one important way. In the information integration theory, a decision cannot be made until all of information has been integrated [7], whereas in the Diffusion Threshold Model, information integration may only occur until a threshold has been reached [20]; this difference is more apparent within a juror context than it is within everyday decision making. In everyday decision making environments,

the Diffusion Threshold Model would propose that individuals integrate available information until a threshold is reached, which would stop information search, and an outcome would then be chosen [20]. Juror decision making differs from everyday decision making though, as jurors do not control when information search/presentation is halted: lawyers and judges have control over the quantity of information shown in court.

Simon [21] suggested that when making decisions our cognition and environment interact. Consequently, jurors may still integrate information until a threshold has been reached; however, the environmental control over information search/presentation may have an impact on what occurs post-threshold. This difference in regard to what occurs once a threshold has been reached may explain why jurors have been found to be anxious about their decision making [22], as decision making that occurs every day may not be challenged post-threshold, whereas juror decision making may be. Decision makers in non-legal environments are not forced fed information (both confirming and disconfirming) once a threshold is reached, meaning that anxiety and regret does not have the chance to arise. Conversely, jurors may be provided with information post-threshold, which then allows disconfirming evidence to challenge the current threshold, meaning that anxiety or cognitive dissonance may occur [10]

The above paragraph raises an important question: what happens to the evidence presented once the last threshold has been reached? Previous research would propose that confirmation bias and evidence distortion might occur. There are a number of reasons for this. First, the reaching of a threshold may create a hypothesis in the juror (e.g., "the defendant is guilty"), which may promote belief persistence and cause beliefs to remain despite the presence of disconfirming evidence [4]. Second, the belief associated with the threshold may be challenged by disconfirming evidence, thus stimulating cognitive dissonance [10], which may lead to confirmation bias and evidence distortion to decrease said dissonance. Finally, Carlson and Russo[9] provide evidence of pre-decisional distortion in jurors, finding that juror interpretations

of new pieces of evidence were biased in order to support leading verdicts. Thresholds and leading verdicts are equivalent, as both represent tentative verdicts that are favoured by a juror during a trial [9], which suggests that thresholds may also promote evidence distortion. Furthermore, it seems likely that after a threshold is crossed, information may be distorted to support the verdict that corresponds to a particular threshold.

The Diffusion Threshold Model may therefore have two stages within a juror context: 1) an information integration stage that allows a threshold to be reached, and 2) an information distortion stage that allows the threshold to be supported. The first stage of this model is based upon traditional Diffusion Threshold Models, where thresholds are reached through information integration [20]. The second stage is based upon Carlson and Russo's [9] research, where leading verdicts were shown to cause evidence distortion. These two stages may make the Diffusion Threshold Model a more realistic and global model within a juror decision making context, as both information integration and evidence distortion has been found in previous juror decision making research [1, 7]. The current study has two main questions to test: '1) does information integration allow thresholds to be reached?' and '2) does the reaching of a threshold promote evidence distortion?'.

Method

Design.

This quasi-experimental design adopted a 2 x 3 x 2 mixed design, where threshold (pre- and post-threshold) and evidence type (prior (before any evidence was shown), prosecution and defence) were altered within-subjects factors, and verdict given (Guilty vs. Not Guilty) was a between-subjects factor. The response measure here was the likelihood of guilt ratings.

Participants.

One hundred and eight participants took part (12 people per vignette over nine vignettes; 73 female). The age range was 18–57 years, and the mean age was 26.1 (*SD* = 8.32) years old. Sixty-nine of the participants identified as students. Exclusion criteria included: people who were not on the electoral role; people who were not native to the UK; and those who did not have a visa. This was to ensure that the participants in this study were as close to a real life British jury population as possible. Opportunistic sampling was used, which included participants being recruited from websites (e.g., Facebook).

Materials.

Vignettes.

Nine vignettes were used in the current study and participants heard one of these vignettes each. This increased the generalisability of the current study as each vignette varied in relation to the evidence, the defendant and the victim. In addition, by reducing the number of vignettes each participant was presented with, it countered a criticism of the last quasi-experiment, as it was reported by the previous studies participants that too many vignettes were shown.

Vignette construction.

Nine separate vignettes were developed through consulting literature on vignette development (i.e., Ashill and Yavas [23]; Heverly, Fitt, and Newman [24]). The vignettes were developed to be short (eight pieces of information and two closing statements) and representative of court trials [25]. The vignettes in the current study were fictitious, but the structure and the evidence used in the vignettes were representative of previous court trials. This was ensured through the researcher consulting newspaper materials, court transcripts and visiting the High Court in Edinburgh. In addition, the vignettes were designed to be consistent, as all the victims were female, all the defendants were male, and all of the charges were homicide, which should

prevent confounding variables, such as crime type and gender bias, from having a negative impact on the study.

The type of information presented, and the order of the information was kept constant throughout the vignettes. This was to ensure that the vignettes were relatively similar. For example, participants first heard an opening statement from a judge, which provided contextual information, such as the name of the victim and the defendant. The first piece of evidence that participants were given was eyewitness testimony from the prosecution, which was then followed by a rebuttal from the defence. Participants then heard special witness testimony that supported the prosecution. This was followed by the defence's attempt to discredit the special witness's evidence in order to create reasonable doubt. The participants were then provided with a secondary confession (i.e., motive), which was discredited by the defence in the following piece of evidence. After this, the prosecution provided participants with information that highlighted that the defendant lacked an alibi. The defence rebutted this with a possible alibi, usually placing the whereabouts of the defendant in a location different from the victim. Finally, participants heard a closing statement from the prosecution, which was followed by a closing statement from the defence. The pieces of evidence, mentioned above, were selected as they had been used in previous forensic decision making experiments and are used in homicide trials [26, 27, 28].

Vignettes were designed to be of similar lengths in order for attentional biases to be reduced. The average word length of the vignettes was 704.11 (minimum = 606; maximum = 773; SD = 57.57).

The structure (i.e., prosecution evidence followed by defence evidence) of the vignettes were designed in a way to allow them to be ambiguous. This was conducted for two reasons. First, because court trials are inherently ambiguous, therefore juror vignettes should be equally ambiguous [29, 30]. Second, because ambiguity breeds bias, consequently ambiguous vignettes

are the best way to investigate decision making biases [31]. In addition, the structure was chosen for ecological validity purposes, as prosecution evidence is followed by defence evidence in real life court trials.

Audio recording device.

The vignettes were recorded in order to make the experiment more ecologically valid, as real jurors would listen to listen to the evidence presented in court. Therefore, in the current study, participants listened to the vignettes and did not read them. The vignettes were recorded with a voice actor in a quiet room at Edinburgh Napier University. The researcher used the audio recording app *Dictaphone – Audio Recorder* on an *iPhone 5* to record the vignettes. The audio recording was then edited on a program named *Audacity*. The average length of the audio vignettes was four minutes 16 seconds (minimum =3.48; maximum = 4.50; *SD* = .38).

Audio playback device.

The audio vignettes were saved as an mp3 file, and played out of *Windows Media Player*. The participants heard the vignettes from *Labtec Spin 95* speakers.

Verdict judgments.

Once participant had heard all the evidence available, they were asked to give a final verdict. Only two-verdict options were available here, participants could only give a Guilty or a Not Guilty verdict, thus creating the between-subjects factor of verdict given.

Threshold.

Participants were asked to state the last piece of information from the vignette that they needed to make a decision, this piece of information was marked as their threshold. For example, participants may after hear DNA evidence and this may allow them to favour a Guilty verdict

(i.e., reach a Guilty threshold). Therefore, it can be said that the DNA evidence allowed the Guilty threshold to be reached, and this threshold point can be measured by asking participants to state what the last piece of information they needed was to reach a decision, as it highlights when a Guilty verdict was initially preferred. Any responses (i.e., likelihood of guilt ratings) given both at and before this point were categorised in the pre-threshold condition, and any responses given after this point were marked within the post-threshold condition, thus allowing the within-subjects factor of threshold to be created.

Likelihood rating scale and cue utilisation.

Likelihood of guilty ratings varied between 0-100, with higher ratings suggesting a higher perception of guilt surrounding the defendant. Participants were asked to state a likelihood of guilt rating after the prior point, after each piece of evidence, and after each of the closing statements.

Demographics questionnaire.

The demographics questionnaire asked participants to state their: 1) age, 2) gender, and, 3) occupation.

Information sheet, consent form and debrief.

All participants received a standardised information sheet and consent form prior to participation in this study, and a debriefing sheet at the end. These materials made participants aware of exclusion/ inclusion criteria and ethical issues.

Procedure.

Participants read the standardised information sheet and completed the consent form within a laboratory room in the Psychology Laboratory at Edinburgh Napier University. Once the consent form was signed, participants filled out the demographics questionnaire. The researcher then played the recording of the opening statement to participants. The participants were then asked to state a prior likelihood of guilt rating (i.e., a likelihood of guilt rating given before any evidence has been provided); this was recorded by the researcher. Participants then heard the first piece of evidence, which was always a piece of eyewitness testimony supporting the prosecution, and were asked to state another likelihood of guilt rating, which was also recorded by the participant. This basic procedure was then repeated for the remaining seven pieces of evidence (three for the prosecution and four for the defence).

Once all of the evidence had been presented, participants were played closing statements from both the prosecution (first) and the defence. These closing statements were also rated in relation to likelihood of guilt ratings. After the closing statements, participants were asked to give a verdict: Guilty or Not Guilty. Then, participants were asked to identify the last piece of evidence that they needed to make their decision (i.e., they state their threshold). By asking participants to state this, it allowed the researchers to identify the piece of evidence that allowed a verdict preference to be initiated. Once the participants had finished, they were given a debrief sheet. The task took on average 20 minutes to complete.

Results

Generalised Estimating Equation (GEE) analysis for testing information integration and confirmation bias.

A Generalised Estimating Equation (GEE) was used to analyse the current data. A GEE is a type of analysis that has been used to analyse repeated measures designs, longitudinal data, and data that are correlated, and was developed to allow repeated measures data that is not normally distributed, or does not meet parametric assumptions, to be analysed [7].

The selection of a correlation matrix.

The corrected quasi-likelihood under independence model criterion (QICC) in the Goodness of fit box was lower for the independent correlation matrix (QICC = 603630.99) than it was for the unstructured (QICC = 648204.98), auto-regressive (QICC = 627448.46) and exchangeable (QICC = 603689.18) correlation matrices; indicating that the model was a better fit when an independent correlation matrix was used. Therefore, the independent correlation matrix was utilised.

GEE analysis of the likelihood of guilt values across threshold, evidence type and verdict given.

The Statistical Package for the Social Sciences (SPSS) version 23 was used to analyse the data. The GEE compared the factors of threshold (pre and post), evidence type (prior, prosecution and defence) and verdict given (Guilty and Not Guilty) across the response measure of likelihood of guilt: for more information on the factors (or explanatory variables) and response variables please see the methods section, specifically the design and materials section. For summary table of model output please see table.1.

Table 1

Summary Table of model output.

Variable	Wald Chi-Square	Degrees of Freedom	P value
Intercept	1256.73	1	<.001
Verdict Type	50.2	1	<.001
Threshold	4.38	1	0.04
Evidence type	276.38	2	<.001
Evidence Type* Verdict Given * Threshold	29.25	5	<.001

Verdict type was found to have a significant main effect on the likelihood of guilt ratings [Wald X^2 (1) = 50.20, p <.001]. Guilty verdicts (Estimated Marginal Means (EEM) = 65.43) had a higher likelihood of guilt rating in comparison to Not Guilty verdicts (EMM = 44.65). Threshold was found to be having a significant main effect on likelihood of guilt ratings [Wald X^2 (1) = 4.38, p =.04]. It was shown that the likelihood of guilt ratings were higher post-threshold (EMM = 62.28) in comparison to pre-threshold (EMM = 50.21). Evidence type was also found to have a significant main effect on likelihood of guilt ratings [Wald X^2 (2) = 276.38, p <.001]. The prior point (EMM = 34.85) had a significantly lower likelihood of guilt rating in comparison to the defence evidence condition (EMS = 54.64; p <.001). The likelihood of guilt rating for the defence evidence condition was significantly lower in comparison to the prosecution evidence condition (EMM = 65.53*;* p <.001). The prior point also had a significantly lower likelihood of guilt rating in comparison to the prosecution evidence condition (p <.001).

A significant interaction was found between the variables of evidence type, verdict given and threshold [Wald X^2 (5) = 29.25, p <.001]. Sidak post hoc tests were used in the current study as they avoid type 1 errors, they can take into account uneven conditions (i.e., less conservative

than Bonferroni tests), and they keep more statistical power in comparison to the Least Significant Difference Test (LSD). See figures 1, 2, 3, 4, and 5 for a visual illustration of the interactions found between the factors of evidence type, threshold and verdict given.

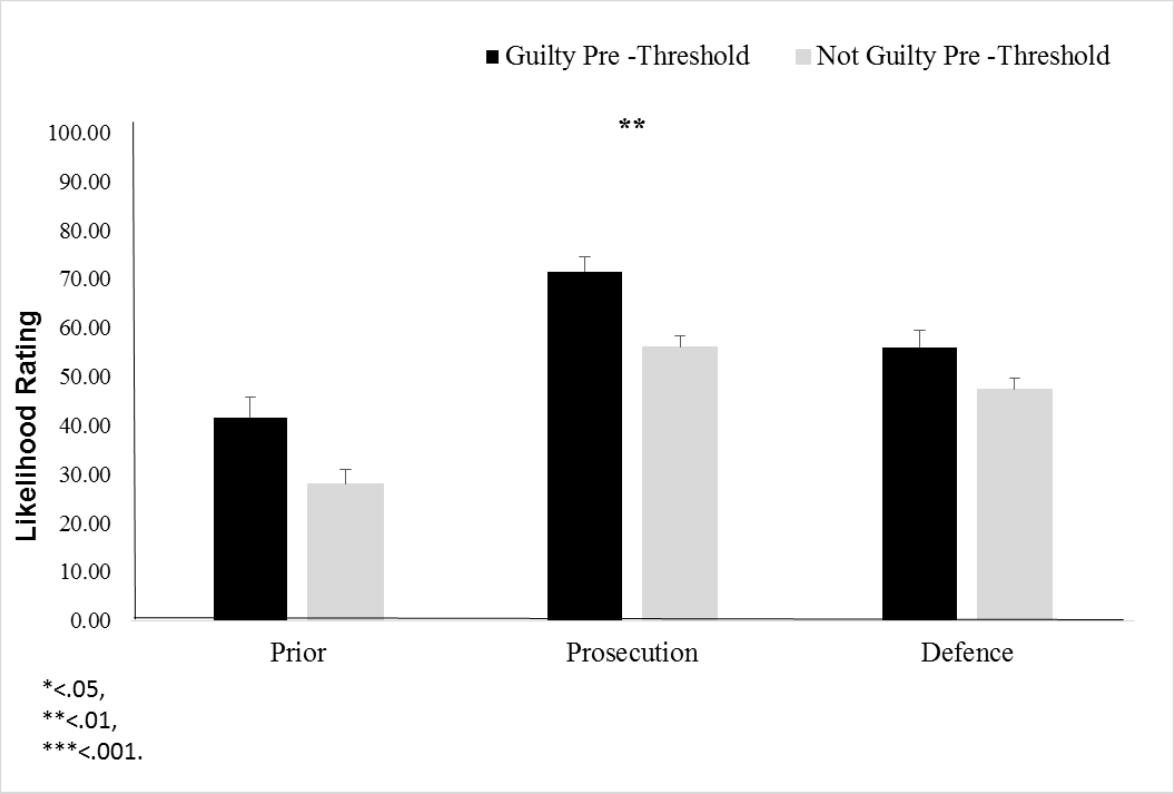


Figure 1. *Interaction between evidence type and verdict given within the pre-threshold condition.*

Figure 1 shows that in the pre-threshold condition, prosecution evidence generated significantly higher likelihood of guilt ratings in Guilty verdicts in comparison to Not Guilty verdicts.

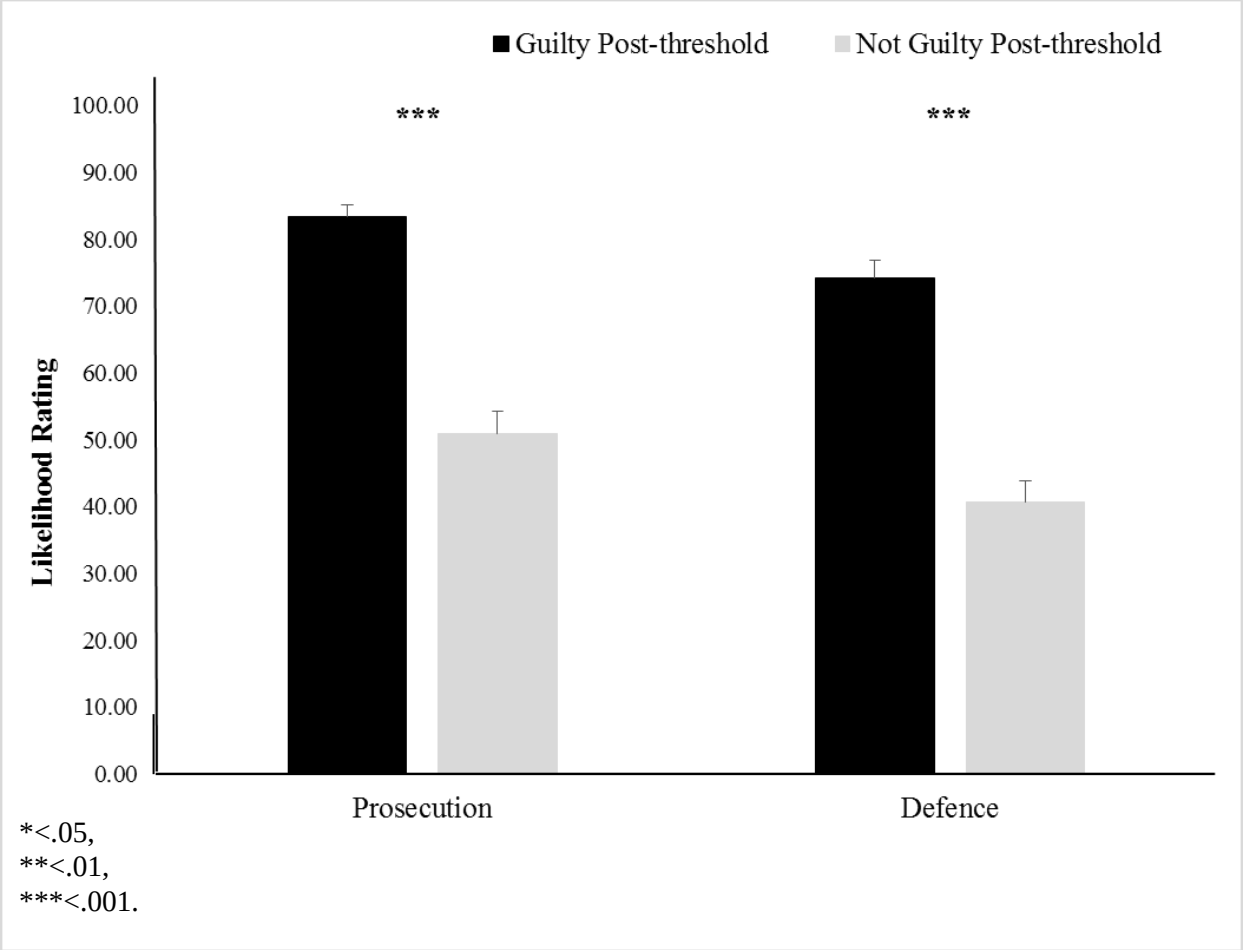


Figure 2. *Interaction between evidence type and verdict given in the post-threshold condition.*

Figure 2 highlights that in the post-threshold condition, both prosecution and defence evidence generated significantly higher likelihood of guilt ratings when a Guilty verdict was given in comparison to when a Not Guilty verdict was given, thus suggesting that thresholds may promote confirmation bias.

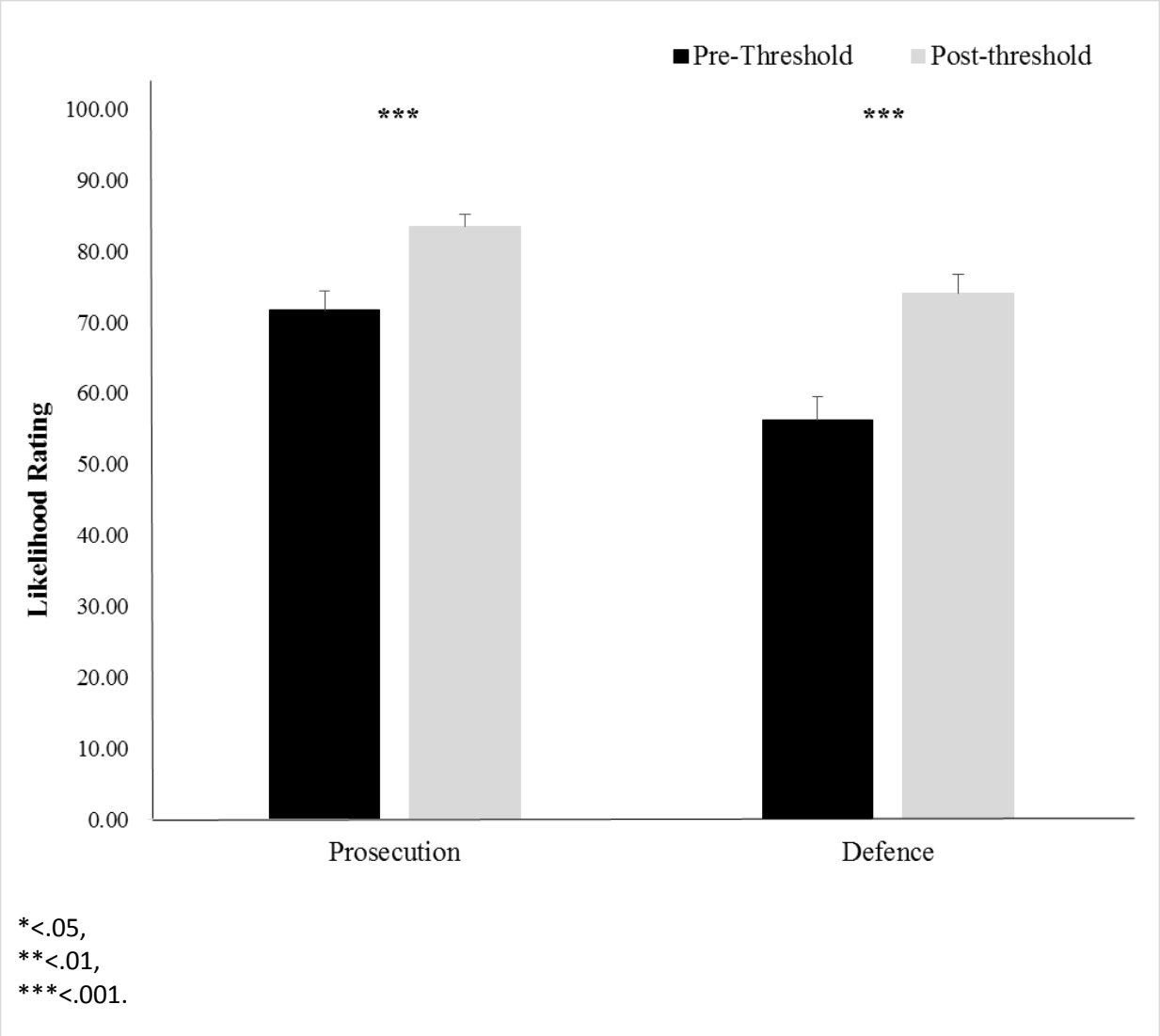


Figure 3. *Interaction between threshold condition and evidence type in Guilty verdicts.*

Figure 3 shows that in Guilty verdicts, both prosecution and defence evidence generated significantly higher likelihood of guilt ratings post-threshold in comparison to pre-threshold, thus suggesting that thresholds may promote confirmation bias.

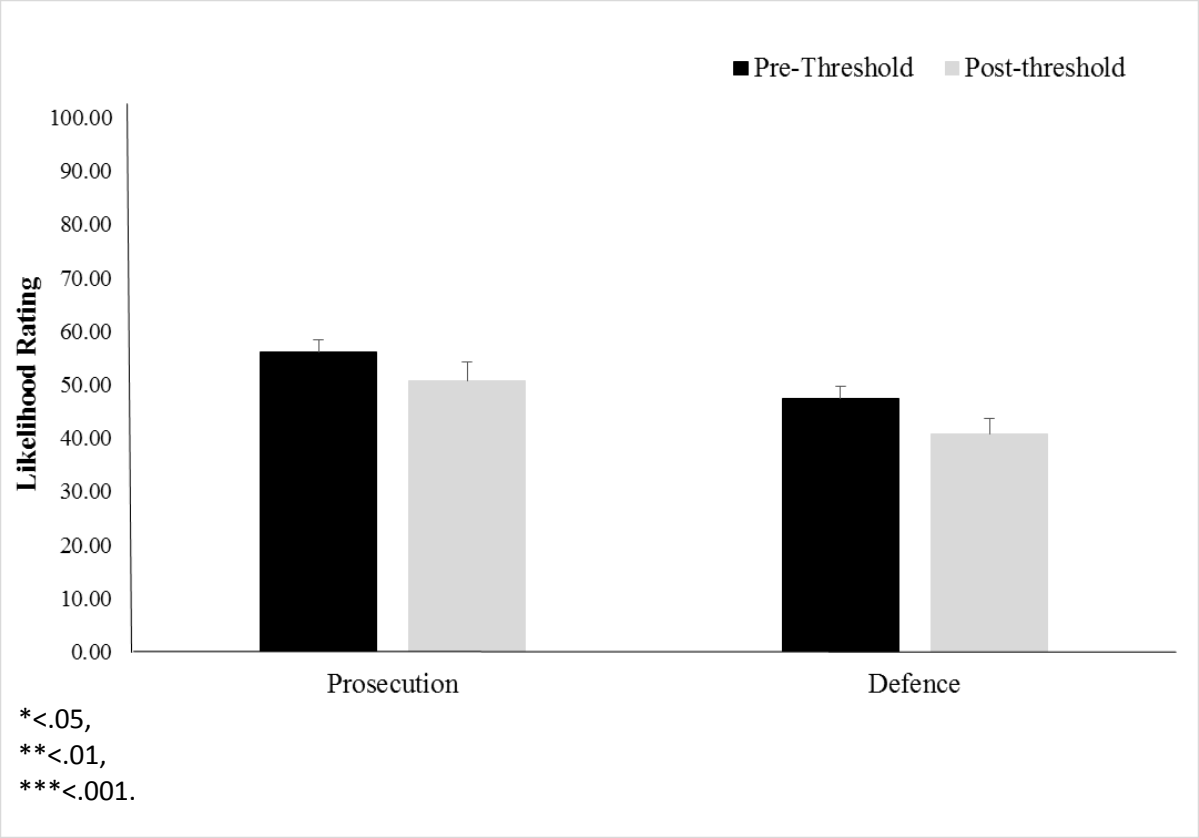
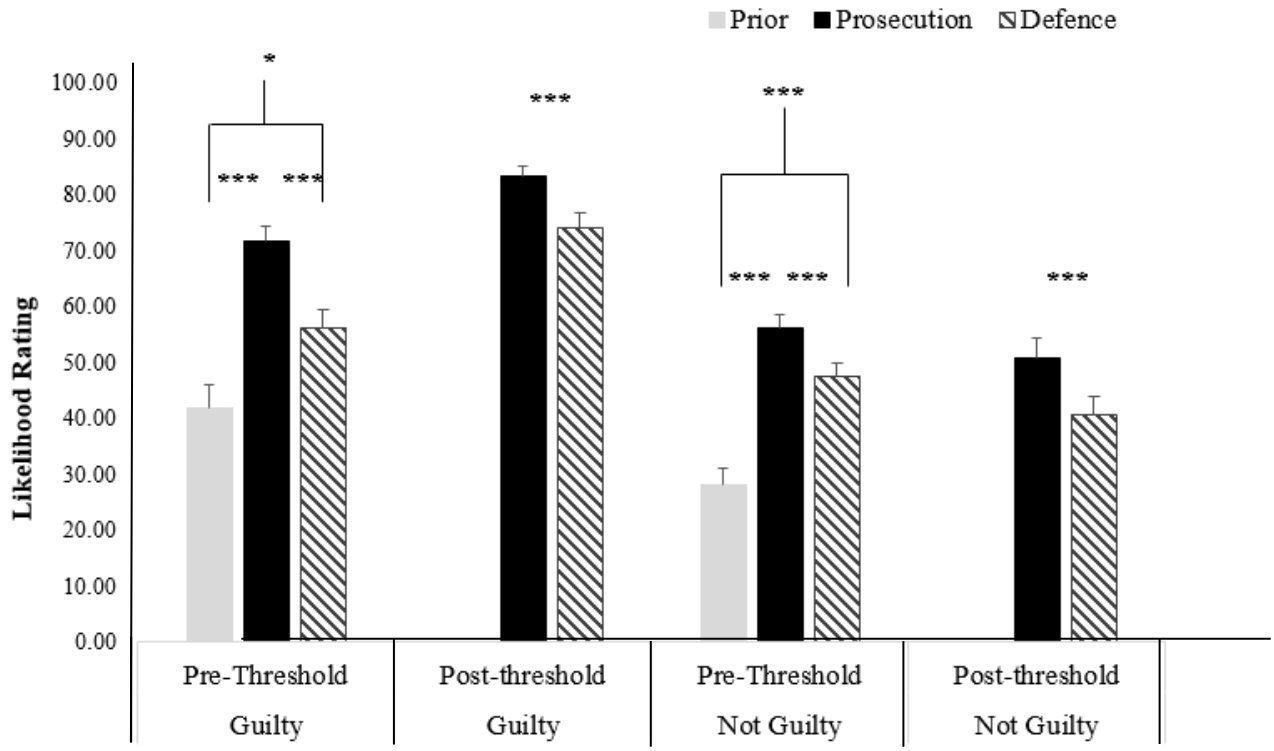


Figure 4. *Interaction between evidence type and threshold condition in Not Guilty verdicts.*

Figure 4 highlights that there was no significant differences between the pre-threshold and post-threshold likelihood of guilt ratings, for both prosecution and defence evidence, when a Not Guilty verdict was given.



*<.05,
**<.01,
***<.001.

Figure 5. *Interaction between evidence type, verdict given and threshold condition across the likelihood of guilt measure.*

Figure 5 highlights that regardless of threshold condition or verdict given, defence evidence generated significantly lower likelihood of guilt ratings in comparison to prosecution evidence, thus suggesting that information integration occurred throughout juror decision processes.

Discussion

Statement of principal findings

The current study had two main questions to test: '1) does information integration allow thresholds to be reached?' and '2) does the reaching of a threshold promote evidence distortion?'. The results highlighted that regardless of the threshold condition (pre vs. post) or the verdict given (Guilty vs. Not Guilty), prosecution evidence generated higher likelihood of

guilt ratings than defence evidence. These results support information integration and similar findings have been found in previous juror research. For instance, Estrada-Reynolds et al.[7] found that jurors in a death penalty case increased their sentence recommendations with prosecution evidence, and decreased their sentence recommendations with defence evidence. Kaplan and Miller [13] found that jurors integrate pre-trial biases and evidence from a trial when reaching a verdict. They also found that biases could be attenuated by increasing the strength of the evidence in court, which provides further evidence for information integration, as the integration of strong evidence counteracts the effects of pre-trial biases. The current results, therefore, are in line with previous research relating to information integration theory.

In relation to the first research aim, it is interesting that in the pre-threshold condition that prosecution evidence generated significantly higher likelihood of guilt ratings than defence evidence, regardless of the verdict given. These results suggest that jurors integrate information (increase with prosecution evidence and decrease with defence evidence), and that this information integration allows threshold to be reached. This is the first direct evidence of information integration allowing thresholds to be reached within a juror decision making setting. However, Ratcliff and Smith [20] did provide evidence of the same process occurring in perceptual decision making tasks.

The current study found that the variables of verdict given, evidence type, and threshold all significantly interacted with one another when producing likelihood of guilt ratings, which highlighted that juror's perceptions of guilt, based on the evidence presented, were altered depending on the threshold (i.e., Guilty threshold versus Not Guilty threshold) that was reached, thus suggesting that thresholds may promote confirmation bias. Each of the main findings will now be discussed.

First, it was found that regardless of the evidence type (prosecution and defence evidence) likelihood of guilt ratings were higher for Guilty verdicts post-threshold than they were for Not Guilty verdicts post-threshold, which suggested that confirmation bias was present. This suggests that when a Guilty threshold is reached that prosecution evidence is seen positively to confirm said threshold, and defence evidence is either disregarded (and does not affect the drift of the decision maker) or is distorted to support the current threshold [1]. When a Not Guilty threshold was reached, the opposite pattern may have occurred, with defence evidence being seen more favourably than prosecution evidence, thus causing the latter evidence type to either be disregarded or distorted to support the current threshold.

The current results are consistent with a number of pieces of previous research. Carlson and Russo [9] found that once jurors reached a leading verdict (i.e., threshold) then they distorted novel evidence to support said verdict. Ask, Rebelius, and Granhag [26] discovered that participants in a homicide trial rated disconfirming evidence as less reliable than confirming evidence. Smith and Bull [32] highlighted that individuals with a pro-prosecution bias towards forensic evidence perceived weak DNA evidence to be strong. When interpreting the current study's results in light of previous research, it is clear that once a juror has acquired a preference for a certain verdict through reaching a threshold, they will perceive supporting evidence positively [32], and disconfirming evidence will either be seen as less reliable [26] or will be distorted to support the current threshold/leading verdict[9].

Second, it was found that when a Guilty verdict was given that the likelihood of guilt ratings were significantly higher post-threshold in comparison to pre-threshold, regardless of evidence type. This suggests two things: 1) that jurors integrate more information pre-threshold than they do post-threshold; and 2) that in Guilty verdicts the evidence is perceived to support the prosecution more post-threshold in comparison to pre-threshold. In addition, the fact that this preference for the prosecution occurred regardless of the evidence type, suggests that defence

evidence was being distorted, or disregarded, to support the leading verdict [9]. The current results highlighted that thresholds may act as a catalyst for confirmation bias and evidence distortion.

Thresholds may promote confirmation bias/evidence distortion through cognitive dissonance [4]. Thresholds may symbolise the point where a juror generates a belief surrounding their preferred verdict (Guilty or Not Guilty), and once this belief is generated, it is open to attack from novel information that supplies competing belief systems. Cognitive dissonance may then occur in the juror as results of these conflicting beliefs of guilt [10]. For example, if a juror reaches a Guilty threshold and defence evidence is then presented, the juror may become victim to cognitive dissonance. The juror may then try to reduce this discomfort by distorting or disregarding disconfirming evidence, thus promoting confirmation bias. In addition, once a threshold is reached and a verdict belief is acquired, jurors may interpret novel information in line with their current threshold, which may then cause them to fall prey to confirmation bias [4]; see Figure 6 for visual illustration of how thresholds may promote confirmation bias.

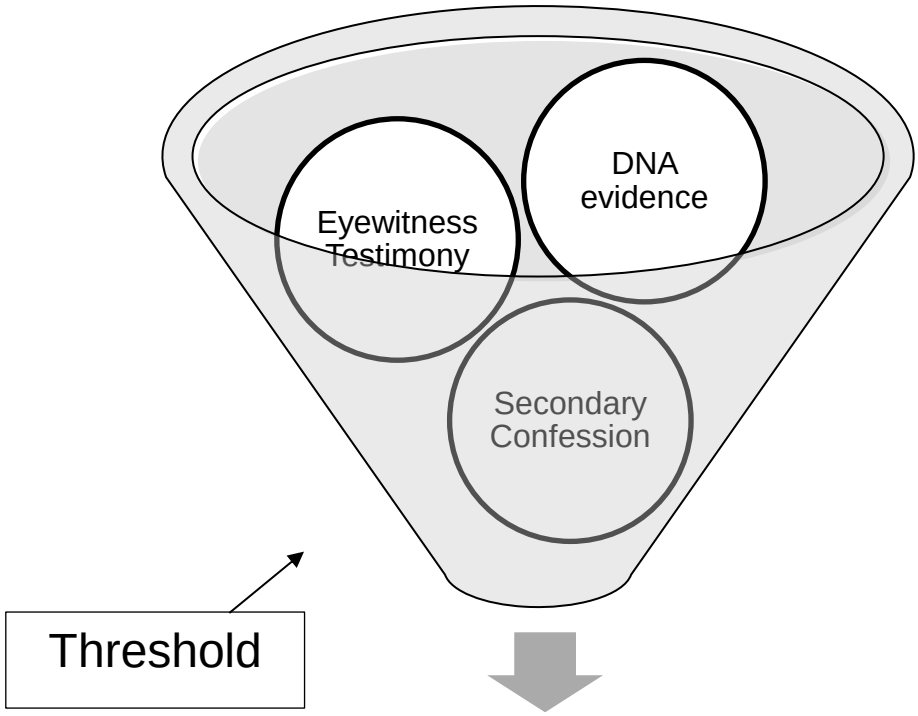


Figure 6. *A threshold effect on novel evidence.*

It was found in Not Guilty verdicts, however, that no significant differences existed in relation to likelihood of guilt ratings, for both evidence types, when comparing the pre-threshold condition with the post threshold condition. One explanation for this lack of significance is that the drift (perception of guilt) in Not Guilty verdicts was already low from an early stage in the decision making process, which may have caused a floor effect. This floor effect may have stopped the decline of likelihood of guilt ratings in the post-threshold condition, when compared to the likelihood of guilt ratings given in the pre-threshold condition, from reaching significance. Pre-trial biases may explain why some jurors (i.e., the ones that gave Not Guilty verdicts) perceived the defendant to be innocent from the beginning of the vignettes [33]. Previous research has shown that pre-trial biases influence juror perceptions of guilt at the beginning of a trial [34], and these pre-trial biases then affect how jurors interpret evidence throughout a trial [32]. In summary, pre-trial biases may have caused the jurors who gave Not Guilty verdicts in

the current study to be biased towards the defence from the beginning of the vignettes and to have low perceptions of guilt throughout said vignettes, which may have produced a floor effect when comparing the pre- and post-threshold conditions in relation to the likelihood of guilt ratings.

The results highlights that prosecution evidence generated significantly higher likelihood of guilt ratings in the post-threshold condition than defence evidence, regardless of verdict type. However, in the post-threshold condition both prosecution and defence evidence generated significantly higher likelihood of guilt ratings in Guilty verdicts than they did in Not Guilty verdicts (as mentioned above). These results taken together show two things: 1) that evidence is distorted to confirm the threshold that is reached; and, 2) that information integration occurs post-threshold.

These findings are similar to those found by Estrada-Reynolds et al.[7], as their research provided evidence for both information integration and evidence distortion. However, Estrada-Reynolds et al.[7] could not adequately explain the dissonance in their findings. The current study proposes that once a threshold is reached that a juror generates a belief surrounding which verdict they prefer, thus causing cognitive dissonance (as stated above), and that disconfirming evidence can be distorted to reduce cognitive dissonance. Nevertheless, in some scenarios the evidence (e.g., DNA evidence) may be too strong to ignore or distort [26]. The juror may then integrate said piece of evidence into their perception of guilt without changing their leading verdict/threshold, thus reducing cognitive dissonance through, limited, information integration. If the evidence is strong enough, it may be possible for verdict reversals to occur, although this enquiry is outwith the scope of the current study.

Overall, the current study highlighted that jurors integrated information until they reached a threshold [20]. This threshold then acted as a leading verdict, which competed with novel belief

systems provided by new evidence [10]. To reduce the cognitive dissonance formed by these competing beliefs, evidence was distorted to support the leading threshold. Some information integration may have occurred post-threshold as to integrate strong pieces of evidence within the decision process, without causing verdict reversals.

Strengths and weaknesses of the study

One limitation relates to how the threshold variable was measured. Participants were asked to state the last piece of information they needed to make a decision, and this response was used to create the threshold variable. However, participants may have fallen prey to hindsight bias, and this may have caused them to overestimate how early they reached their threshold [35]. The current study could not have investigated information integration or the influence of thresholds on evidence distortion without asking participants to state both likelihood of guilt ratings after each piece of evidence and the last piece of information they needed to make their decision.

Meaning of the study: possible mechanisms and implications for clinicians or policymakers

One implication from the current piece of research is that thresholds may allow confirmation /evidence distortion to occur, as thresholds may symbolise the first time a juror creates a belief regarding which verdict they prefer. This promotion of confirmation bias may impede on the legal system's ability to deliver justice to society, and may make the trial by jury process unfair for legal actors (i.e., the prosecution and the defence). To prevent said negative effects, thresholds with a high cue utilisation (i.e., using all the evidence in a trial) should be promoted. Judges do currently instruct jurors to only make a verdict once they have heard all of the evidence provided in court, but these instructions are normally given once all the evidence has already been presented [36]. Therefore, the current study would propose that jurors be given said instructions at the beginning of a trial[9], as this may help to stop thresholds being reached

prematurely and may attenuate confirmation bias in jurors. In addition, the current research, and similar research, should be used by judges to highlight to jurors the dangers of reaching a threshold (i.e., favouring a verdict) before all the evidence has been heard[9].

Unanswered questions and future research

Despite meeting the research aims/questions of the thesis so far, many questions have been left unanswered or have indeed been raised by the findings of the current study. One variable that was not tested in the current study relates to pre-trial biases. Estrada-Reynolds et al. [7] showed that pre-trial biases (pro-prosecution, neutral and pro-defence) did allow verdicts to be forecasted. Pre-trial biases may also predict the thresholds that participants are likely to reach and how much information they may need to reach these thresholds. Future research is needed to test this. In addition, the Scottish legal system with its Not Proven verdict is a unique legal system that has not been extensively research. Consequently, it would be interesting to investigate if Not proven verdicts are reached using this threshold mechanism and if the inclusion of the Not Proven verdict has any attenuating effects on confirmation bias.

The current piece of research has shown two things: 1) that information integration allowed thresholds to be reached; and 2) that confirmation bias may be promoted by a threshold being reached. The results highlight that in the pre-threshold condition, prosecution evidence generated significantly higher likelihoods of guilt ratings in comparison to defence evidence (regardless of verdict type), which suggested that information integration allowed thresholds to be reached. It was also found that regardless of evidence type (prosecution and defence), likelihood of guilt ratings were significantly higher for Guilty verdicts post-threshold than they were for Not Guilty verdicts post-threshold, thus highlighting that confirmation bias may be promoted through a threshold being reached. In addition, the results show that in Guilty verdicts both evidence types generated significantly higher likelihood of guilt ratings post-threshold in comparison to pre-threshold, which once again suggest that confirmation bias

might be promoted through a threshold being reached. The current study suggests that the trial by jury process should be amended to increase thresholds in relation to cue utilisation, thus allowing confirmation bias to be attenuated.

References

1. Nickerson RS. Confirmation bias: A ubiquitous phenomenon in many guises. Review of general psychology. 1998 Jun;2(2): doi:10.1037/1089-2680.2.2.175
2. Cargill, N. Confirmation Bias and the Law, http://nataliecargill.com/ (2016, accessed 21 December 2016).
3. LaRosa, P. Juror's insights on David Camm acquittal at 3rd murder trial. CBS news, 1 August, 2014.
4. Findley KA, Scott MS. Multiple Dimensions of Tunnel Vision in Criminal Cases, The. Wis. L. Rev.. 2006:1-107. Retrieved from: http://wisconsinlawreview.org/
5. McFarlane BA, Cordner SM. Wrongful convictions: the effect of tunnel vision and predisposing circumstances in the criminal justice system. Toronto: Government of Ontario; 2008.
6. Burke A. Neutralizing cognitive bias: An invitation to prosecutors. NYUJL & Liberty. 2006;2: 512-530. Retrieved from: http://scholarlycommons.law.hofstra.edu/faculty_scholarship/671
7. Estrada-Reynolds VC, Gray JM, Nuñez N. Information Integration Theory, Juror Bias, and Sentence Recommendations Captured Over Time in a Capital Trial. Applied Cognitive Psychology. 2015 Sep 1;29(5):713-22.
8. Chaxel, Anne-Sophie, J. Edward Russo, and Neda Kerimi. "Preference-driven biases in decision makers' information search and evaluation." *Judgment and decision making* 8, no. 5 (2013): 561-576. Retrieved from: http://search.proquest.com/?accountid=16607
9. Carlson KA, Russo JE. Biased interpretation of evidence by mock jurors. Journal of Experimental Psychology: Applied. 2001 Jun;7(2): 91-103. doi:10.1037//1076-898X.7.2.91
10. Jonas E, Schulz-Hardt S, Frey D, Thelen N. 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. 2001 Apr;80(4):557-571. doi:10.103700022-3514.80.4.557

11. Elliot AJ, Devine PG. On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort. Journal of personality and social psychology. 1994 Sep;67(3): 382-394. doi:10.1037/0022-3514.67.3.382
12. Pritchard ME, Keenan JM. Does jury deliberation really improve jurors' memories?. Applied Cognitive Psychology. 2002 Jul 1;16(5):589-601. doi:10.1002/acp.816
13. Kaplan, M. F., & Miller, L. E. (1978). Kaplan MF, Miller LE. Reducing the effects of juror bias. Journal of personality and social psychology. 1978 Dec;36(12):1443-1455. doi:10.1037/0022-3514.36.12.1443
14. MacCoun RJ. Experimental research on jury decision-making. Science. 1989 Jun 2;244(4908): 1046-1050. Retrieved from: http://www.jstor.org/
15. Shanteau, J., & Nagy, G. F. (1984). Information integration in person perception: Theory and application. In M. Cook (Ed), *Issues in person perception* London: Methuen. pp. 48-86.
16. Ostrom TM, Werner C, Saks MJ. An integration theory analysis of jurors' presumptions of guilt or innocence. Journal of Personality and Social Psychology. 1978 Apr;36(4): 436–450. doi:10.1037/0022-3514.36.4.436
17. Kaplan MF, Kemmerick GD. Juror judgment as information integration: Combining evidential and nonevidential information. Journal of Personality and Social psychology. 1974 Oct;30(4):493-499. doi:10.1037/h0037034
18. Curley, MacLean, and Murray Threshold Utilisation in juror decision making. in review.
19. Lee MD, Cummins TD. Evidence accumulation in decision making: Unifying the "take the best" and the "rational" models. Psychonomic Bulletin & Review. 2004 Apr 1;11(2):343-352. doi: 10.3758/BF03196581
20. Ratcliff R, Smith PL. A comparison of sequential sampling models for two-choice reaction time. Psychological review. 2004 Apr;111(2): 333-267. doi: 10.1037/0033-295X.111.2.333
21. Simon HA. Rational choice and the structure of the environment. Psychological review. 1956 Mar;63(2): 1-11. doi: 10.1037/h0042769.
22. Bornstein BH, Miller MK, Nemeth RJ, Page GL, Musil S. Juror reactions to jury duty: Perceptions of the system and potential stressors. Behavioral sciences & the law. 2005 May 1;23(3):), 321-346. doi: 10.1002/bsl.635
23. Ashill NJ, Yavas U. Vignette development: an exposition and illustration. Innovative Marketing. 2006;2(1):28-36. Retrieved from: https://businessperspectives.org/component/option,com_journals/task,journal/id,5/Itemid,74/
24. Heverly MA, Fitt DX, Newman FL. Constructing case vignettes for evaluating clinical judgment: An empirical model. Evaluation and Program Planning. 1984 Jan 1;7(1):45-55. doi:10.1016/0149-7189(84)90024-7
25. Miles, M.B. and Huberman, A.M. *Qualitative Data Analysis: An Expanded Sourcebook*. Thousand Oaks, CA: SAGE Publications, 1994.

26. Ask K, Rebelius A, Granhag PA. The 'elasticity'of criminal evidence: A moderator of investigator bias. Applied Cognitive Psychology. 2008 Dec 1;22(9): 1245–1259. doi: 10.1002/acp.1432
27. Lemieux JM. *The relationship between beliefs, strength of evidence, statistical presentation, and expert testimony on jury decision-making in DNA cases*. Doctoral dissertation, Carleton University, CA, 2007.
28. Wetmore SA, Neuschatz JS, Gronlund SD. On the power of secondary confession evidence. Psychology, Crime & Law. 2014 Apr 21;20(4): 339-357. doi:10.1080/1068316X.2013.777963
29. Arkes HR, Mellers BA. Do juries meet our expectations?. Law and Human Behavior. 2002 Dec;26(6):625-639. doi:10.1023/A:1020929517312
30. De La Fuente L, De La Fuente EI, García J. Effects of pretrial juror bias, strength of evidence and deliberation process on juror decisions: New validity evidence of the Juror Bias Scale scores. Psychology, Crime & Law. 2003 Jun 1;9(2):197-209. doi: 10.1080/1068316031000116283.
31. Dror IE. A hierarchy of expert performance. Journal of Applied Research in Memory and Cognition. 2016 Jun 1;5(2):121-127. doi: 10.1016/j.jarmac.2016.03.001
32. Smith LL, Bull R. Validation of the factor structure and predictive validity of the Forensic Evidence Evaluation Bias Scale for robbery and sexual assault trial scenarios. Psychology, Crime & Law. 2014 May 28;20(5):450-466. doi: 10.1080/1068316X.2013.793340
33. Lecci L, Myers B. Individual differences in attitudes relevant to juror decision making: Development and validation of the Pretrial Juror Attitude Questionnaire (PJAQ). Journal of Applied Social Psychology. 2008 Aug 1;38(8): 2010-2038. doi: 10.1111/j.1559-1816.2008.00378.x
34. Ruva C, McEvoy C, Bryant JB. Effects of pre-trial publicity and jury deliberation on juror bias and source memory errors. Applied Cognitive Psychology. 2007 Jan 1;21(1):45-67.doi:10.1002/acp.1254
35. Arkes HR, Wortmann RL, Saville PD, Harkness AR. Hindsight bias among physicians weighing the likelihood of diagnoses. Journal of Applied Psychology. 1981 Apr;66(2):252-254. doi: 10.1037/0021-9010.66.2.252