

Open Research Online

The Open University's repository of research publications and other research outputs

Proven and not proven: A potential alternative to the current Scottish verdict system

Journal Item

How to cite:

Curley, Lee John; Munro, James; Turner, Jim; Frumkin, Lara A.; Jackson, Elaine and Lages, Martin (2022). Proven and not proven: A potential alternative to the current Scottish verdict system. Behavioral Sciences & the Law (Early access).

For guidance on citations see [FAQs](#).

© 2022 The Authors



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Version of Record

Link(s) to article on publisher's website:
<http://dx.doi.org/doi:10.1002/bsl.2568>

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

Proven and not proven: A potential alternative to the current Scottish verdict system

Lee John Curley¹ | James Munro¹ | Jim Turner¹ |
Lara A. Frumkin¹ | Elaine Jackson² | Martin Lages²

¹Faculty of Arts and Social Sciences, School of Psychology and Counselling, The Open University, Milton Keynes, UK

²School of Psychology and Neuroscience, College of Science and Engineering, University of Glasgow, Glasgow, UK

Correspondence

Lee John Curley, Faculty of Arts and Social Sciences, School of Psychology and Counselling, The Open University, Milton Keynes, UK.

Email: Lee.Curley@open.ac.uk

Abstract

The current Scottish verdict system includes three verdicts: 'guilty', 'not guilty' and 'not proven'. The Scottish Government are currently reviewing the utility of the not proven verdict. Proponents of the not proven verdict suggest that it directs jurors to their true role of determining whether the prosecution's case has, or has not, been 'proven'. Reformists suggest a move to a system similar to England and Wales, with only guilty and not guilty verdicts. However, legal professionals have indicated a preference for an alternative system of proven and not proven. The aim of the current study was to test the effects of a proven and not proven system on verdicts given, when compared to alternative verdict systems (specifically, the current Scottish and Anglo-American verdict systems). 227 mock jurors watched a staged murder trial, filmed in a real-life courtroom, with legal professionals questioning witnesses and a judge giving legal direction. Jurors were significantly more likely to convict in a guilty and not guilty verdict system than either a proven and not proven or a guilty, not guilty and not proven verdict system. Future research should replicate this study with a focus on the impact of the not proven verdict in sexual offences.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. Behavioral Sciences & The Law published by John Wiley & Sons Ltd.

KEYWORDS

juror decision making, jury reform, proven and not proven, Scottish legal system, three-verdict system

Scotland, unlike other nations that use juries in criminal courts, has three verdicts available to juries (Curley, Munro, et al., 2021). These verdicts are 'guilty', 'not guilty' and 'not proven' (Curley, Munro, et al., 2021). The not proven verdict is an acquittal verdict with the exact same legal implications as a not guilty verdict (Curley, Murray, et al., 2021). The current three-verdict system is currently being publicly debated by stakeholders, including citizen activists and legal professionals, and the major political parties in Scotland have voiced support for either abolition or review of the not proven verdict (Richardson & Gardiner, 2021). The debate includes both points of legal principle and data on convictions and acquittals. For instance, legal professionals often view the not proven verdict as representing an important legal principle: it directs jurors to focus on the 'proof' provided by the case evidence and not impressions of guilt (or otherwise) of the accused (Curley, Munro, et al., 2021). However, politicians and activists commonly note that the not proven verdict is used disproportionately in murder and rape trials, and that Scotland has an even lower conviction rate for those crimes than England and Wales, implying that the not proven verdict may be a factor in this difference (Hope et al., 2008; Richardson & Gardiner, 2021; Scottish Government, 2019; Topping & Barr, 2020). The aim of this paper is to compare the current Scottish verdict system with a commonly proposed alternative (i.e. the Anglo-American system) and a reform that would be consistent with survey-based research into the views of Scottish legal professionals (i.e. a proven and not proven system; Curley, Munro, et al., 2021).

The Scottish verdict system has been reformed several times before it assumed its current form (Barbato, 2004; Curley, Murray, et al., 2021; Duff, 2021; Walker, 1988). The not proven verdict became part of the Scottish legal system through a convoluted, and unplanned, route (Barbato, 2004; Curley, Murray, et al., 2021; Walker, 1988). Before the late 17th century, Scottish juries could indicate their belief in the accused's guilt or innocence, although there was a lack of consistency in the wording (Barbato, 2004; Walker, 1988). However, in the late 17th century, juries were stripped of their ability to give general verdicts of guilty and not guilty, due to cases of refusal to convict those charged under laws considered to be oppressive (Barbato, 2004; Curley, Murray, et al., 2021). At this time, Scottish juries could only give verdicts of proven and not proven in relation to specific factual allegations (equivalent to modern-day criminal charges) as part of an indictment (Barbato, 2004; Walker, 1988). Judges then used these special verdicts to rule on the overall innocence or guilt of an accused individual (Smith & MacDonald, 1958). Juries in Scotland were given back the ability to return general verdicts (i.e. guilty and not guilty) in the first half of the 18th century. However, the not proven verdict was retained, thus creating the current three-verdict Scottish system (Barbato, 2004; Curley, Munro, et al., 2021; Curley, Murray, et al., 2021; Curley, MacLean, Murray, Laybourn, et al., 2019; Smith & MacDonald, 1958).

The debate on the not proven verdict has been somewhat divisive (Curley, Murray, et al., 2021; Richardson & Gardiner, 2021). Some supporters of the not proven verdict argue that it allows jurors to express doubt in the prosecution's evidence without taking a position on the accused's factual guilt (Curley, Munro, et al., 2021). Other proponents argue that it helps to distinguish a not guilty verdict as indicating belief in innocence, not just doubt in guilt (Jackson, 1998). In the Anglo-American system, a not guilty verdict can be given for two different reasons: 1) the jury concludes that the prosecution has failed to prove the case against the defendant beyond reasonable doubt; or 2) the jury believes that the defendant is innocent of the offence. Jackson (1998) explains that a not proven verdict may disambiguate these two outcomes, as it can be used when the jury believes that the crown has not met the burden of proof. This means that a not guilty verdict can be reserved for cases in which the jury believes the accused to be factually innocent. Opponents of the not proven verdict often note that it is used disproportionately in serious crimes such as homicide and rape (Hope et al., 2008; Richardson & Gardiner, 2021; Scottish Executive, 2006; Scottish Government, 2019). An argument is also made that the not proven verdict is incongruent with Article six of the European Convention on Human Rights (Hope et al., 2008), which states that everyone should be assumed innocent

until proven guilty. A verdict of 'not proven' rather than 'not guilty' may undermine the presumption of innocence, even after acquittal.

A number of studies have empirically investigated the influence of a two-verdict system (guilty and not guilty) compared to a three-verdict system (guilty, not guilty and not proven) on trial outcomes (Curley, Murray, et al., 2021; Hope et al., 2008; Smithson et al., 2007). For example, Curley, Murray, et al. (2021) found that the availability of the not proven verdict decreased the number of guilty verdicts by 38.1% and not guilty verdicts by 74.4%. This suggests that conviction rates could change significantly by moving from a three-verdict system to a two-verdict system. One explanation for this finding is that the availability of the not proven verdict, with its explicit reference to 'proof', might promote analytical thinking (Curley et al., 2018; Thomas & Millar, 2012). This could implicitly focus jurors' attention on weaknesses in the Crown's case (Curley, Munro, et al., 2021; Hope et al., 2008; McKenzie, 1985), which may lead to a shift in their decision criterion (higher reasonable doubt threshold) and lead to a lower conviction rate (Curley, Murray, et al., 2021).

Similar results have also been found in larger, more ecologically valid studies, such as Ormston et al. (2019). Ormston and colleagues used videos of a realistic staged trial (involving either a physical assault or a sexual assault), and asked participants to deliberate in mock juries for 90 min and then deliver a verdict. Half of these juries had two verdicts available to them (guilty and not guilty) and the other half had three verdicts available (guilty, not guilty and not proven). In physical assault trials, jurors were significantly more likely to give a guilty verdict in the two-verdict system than the three-verdict system, although the difference was not significant in a sexual assault trial. The findings of this experiment have been used by Scottish policymakers as evidence that the not proven verdict should be abolished (Duff, 2021), leaving a guilty and not guilty system.

Curley, Munro, et al. (2021) argued, however, that the voice of legal professionals was largely absent from the academic research, and conducted a survey to address this lacuna. They found that the majority of legal professionals were against abolition of the not proven verdict. The legal professionals perceived the Anglo-American verdict system to be a worse option than the Scottish three-verdict system and, if given the opportunity to redesign the Scottish verdict system *ad libitum*, they would favour a two-verdict system of proven and not proven. One rationale given for favouring a proven and not proven verdict system was that it may help to direct jurors to their true role of evaluating whether a case has been 'proven beyond reasonable doubt' (Curley, Munro, et al., 2021).

The history of the (not) proven verdict demonstrates that it was once used to remove from the jury the ability to judge the morality or guilt of the accused (Barbato, 2004). The word 'guilty' comes with implicit connotations that may drive emotional decision-making processes. However, it is not the role of a juror to decide on the *moral* guilt of the accused (which is a subjective reflection of the juror's own morality), only whether or not the Crown has proven their case beyond reasonable doubt (Curley, Munro, et al., 2021). Likewise, jurors are not tasked to establish whether the accused is factually guilty or not, rather their role is to evaluate the evidence and establish if it sufficiently supports the prosecution's case to warrant a conviction. The jury's task is one of proof, not truth. A move to a proven and not proven system may direct jurors to their intended role, of considering proof beyond reasonable doubt, more effectively than verdicts framed in terms of guilt (Curley, Munro, et al., 2021). This focus on proof over guilt (or truth) may also influence juror perceptions of the beyond reasonable doubt threshold; something this paper hopes to investigate.

In addition, there is evidence that jurors are more likely to deliver guilty verdicts if they are guided by intuitive decision-making strategies, often associated with emotion (Curley et al., 2018; Curley et al., 2019). This type of reasoning, referred to by Kahneman (2003) as 'system 1' reasoning has been associated with cognitive biases in the perception of information. One such bias is the 'framing effect', in which participants may perceive information differently depending on which features are emphasised. This bias has been found to be reduced in participants who are asked to "think like a scientist" (Thomas & Millar, 2012, p. 143). Further, terms such as 'proven' may promote similar analytical processes and lead to a lowering of the conviction rate when compared to a guilty and not guilty system. However, despite this, jurors may prefer terms such as 'guilty' and 'not guilty' over terms such as 'proven' and 'not proven'. The former may be perceived by jurors to reflect a likelihood of (or desire for) a more punitive outcome for the

accused (Curley, Munro, et al., 2021), and enabling punishment of offenders may be a key part of how jurors see their role. Further, the proven verdict might be seen by some as being 'too soft' on convicted individuals (Curley, Munro, et al., 2021) by not clearly labelling them as 'guilty'.

A final, and more pragmatic than theoretical, explanation is that in three-verdict systems, the 'middle' verdict (i.e. not proven) is seen as a compromise between guilt and innocence. For example, the not proven verdict may be selected if a juror is unsure or as a compromise between jury members divided between 'guilty' and 'not guilty'. In binary systems, this middling verdict option is not available, thus eliminating a 'compromise effect' (Hope et al., 2008) and increasing convictions relative to a three-verdict system. Further, if the removal of a 'compromise' verdict was aligned with a focus on proof, by having a proven and not proven verdict system rather than the guilty and not guilty verdict system, there should not be an increase in the instances of injustice (Curley, Munro, et al., 2021). Despite gaining valuable insights from a unique sample (i.e. legal professionals), Curley, Munro, et al. (2021) did not directly investigate the influence that a proven and not proven verdict system would have on the decision making of jurors. A proven and not proven system may increase convictions when compared to a three-verdict system, much as Ormston et al. (2019) found the guilty and not guilty verdict system to do. This is a key empirical question, which we address in this paper and which also allows for disambiguation between the mere existence of a third 'middle' verdict and the semantic and conceptual framing of the verdicts (i.e. 'guilt' vs. 'proof').

Our aim in this paper is to compare the current Scottish legal system with the most likely reform (i.e. the Anglo-American system) and the alternative reform that Scottish legal professionals preferred (i.e. proven and not proven).

The hypotheses are:

1. Jurors will be more likely to convict and less likely to acquit in the proven and not proven verdict system when compared with the guilty, not guilty and not proven Scottish verdict system.
2. Jurors will be more likely to convict and less likely to acquit in the guilty and not guilty Anglo-American verdict system when compared with the guilty, not guilty and not proven Scottish verdict system.
3. In convictions, jurors will show a preference between guilty and proven verdicts.
4. In acquittals, jurors will show a preference between not guilty and not proven verdicts.
5. Jurors will report differing levels of the perceived beyond reasonable doubt when delivering a guilty (conviction) verdict when compared to a proven (conviction) verdict.

1 | PILOT STUDY

Since this is the first time a proven and not proven system has been directly compared with the Scottish three-verdict and Anglo-American two-verdict system, we began by conducting a pilot study with 30 mock jurors, using a within-subjects design and including pre-trial bias as a covariate.

Although the legal system aims to operate as fairly as possible, research shows that jurors may have implicit biases that can affect their decision making (Curley, Murray, et al., 2021; Lecci & Myers, 2008). However, the interaction between these biases and verdict type has received little academic attention. For example, the individual biases of jurors may interact with the unclear nature of the *not proven* verdict to produce a system which inadvertently allows those biases to dominate the decision-making process (Curley et al., 2022). We therefore investigated the effects of pre-trial biases on jurors' decision making across different verdict systems in our pilot study. For this we used the pre-trial juror attitudes questionnaire (PJAQ), which has been used successfully in other research (Estrada-Reynolds et al., 2015) and has good internal validity (Lecci & Myers, 2008).

After reading a transcript (approximately 13,000 words long) about a fictitious common assault trial, the verdicts of 26 mock jurors (4 had to be excluded from the analysis) were compared in terms of (i) number of convictions and acquittals and (ii) confidence ratings. Although the number of convictions and acquittals did not vary significantly

between verdict systems, a linear regression on the confidence ratings for convictions indicated a significant effect of verdict system and a significant effect of pre-trial bias. Confidence in convictions was significantly decreased for the proven/not proven verdict system compared to the Anglo-American system (guilty/not guilty), but not compared to the Scottish (guilty/not guilty/not proven) system. There was no significant interaction between verdict system and PJAQ scores on verdicts or confidence. This suggests that having not proven as a verdict option significantly affected jurors' decisions, in and of itself. Further details of this study can be found on: osf.io/vpjuik.

1.1 | Main study

After piloting was completed, we designed our main study using more ecologically-valid material, based on a staged trial of a murder case. The rationale for moving to a murder case was that the not proven verdict is used disproportionately in trials for serious offences, including homicide and sexual assaults (Hope et al., 2008; Scottish Government, 2019) and homicide trials have received less attention in the literature (see Hope et al., 2008). The PJAQ (Lecci & Myers, 2008; Lecci et al., 2013) was not used in the main study, as our pilot study found the effect of verdict system to be independent of pre-trial bias.

2 | METHOD

2.1 | Design

Our hypotheses were tested using an experimental design with verdict system as a between-subjects factor. This factor had three levels: 1) guilty, not guilty and not proven verdict system; 2) guilty and not guilty verdict system; 3) proven and not proven verdict system. We will refer to these as the 'Scottish', 'Anglo-American', and 'Experimental' verdict systems respectively when discussing our methods and results. Our dependent measures, which were taken using an online questionnaire tool, were: 1) verdict; 2) confidence in verdict; 3) perception of accused's guilt; and, 4) decision difficulty. Conviction and acquittal rates were calculated from individual jurors' verdicts. These measures, additional questions and the stimulus video are detailed in the Materials section below. The study received ethical approval from the Human Research Ethics Committee at the Open University.

2.2 | Participants

The following exclusion and inclusion criteria were used to recruit participants: 1) participants had to be over 18 years old; 2) participants had to be registered as a parliamentary or local government elector in the United Kingdom (i.e. on the electoral roll); and 3) participants must have lived in the United Kingdom, Channel Islands or Isle of Man for any period of at least 5 years since they were 13 years old. These criteria correspond to the requirements for jury eligibility in Scotland.

The data collection phase of the study occurred in two stages. In the first stage, we recruited and collected data from 33 participants, to ensure that our online experiment was operating correctly (Andrade, 2020). These participants were recruited through sharing the mock trial and associated materials on social media sites such as Twitter and Facebook. We then opened up our second stage of recruitment, in which 194 eligible participants were recruited using the online platform Prolific. Fisher's exact test established that participants recruited from Prolific did not give significantly different verdicts from the participants that were recruited from social media [$\chi^2(3) = 1.38, p = 0.72, \phi_c = 0.08$], so the data from both stages were combined into one data set. The mean age of the participants was 37.41 years (SD = 12.76; min = 18; max = 71), and the sample consisted of 133 females, 87 males, six non-binary

individuals and one person who identified as transgender. One-hundred and thirty one individuals stated that they were employed, 29 reported they were self-employed, 20 individuals were in the process of looking for work, 12 were not able to work due to a disability, 10 were retired, 17 were not currently working (of whom 15 were students), five were not working due to other reasons and three preferred not to say.

All 227 participants were born in and/or lived in Scotland: 198 reported that they were born in Scotland, 207 lived in Scotland at the time of participation. The sample was predominantly Caucasian: 186 participants reported their ethnicity as white Scottish, 12 as white British, four as white English, four as white other, and two identified as a combination of different white ethnicities. Three participants reported their ethnicity as Pakistani Scottish/British, two as Chinese Scottish/British, six as black or African Scottish/British, and one as Latin. Six participants identified as mixed race. One participant preferred not to disclose their ethnicity. The sample was fairly representative of the Scottish population, as Scotland's ethnic make-up was 96% white and 2.6% Asian in 2018 (Statista, 2018), the last year for which complete demographic data were available before this study was conducted.

Based on our intended analyses, our sample of 227 participants meets the required N of 219 for a moderate effect size of $w = 0.3$ (with alpha of 0.05 and power of 0.95).

3 | MATERIALS

3.1 | Video trial

The mock trial used in the current study was developed by the Modern Studies Association of Scotland in association with members of the Faculty of Advocates and Bloody Scotland. It was inspired by the real case of Peter Queen, who was accused and ultimately convicted of murdering his lover in 1932. In the video, modern forensic evidence was included and current trial practices were used. The video allowed participants to watch a realistic mock trial, filmed in a real courtroom, with the use of legal professionals, a professional judge, and actors playing the roles of witnesses. The video lasted approximately 53 min. There were three variations of the video in relation to verdict directions to the jury, which reflect our experimental conditions: 1) Scottish: *guilty*, *not guilty* and *not proven*; 2) Anglo-American: *guilty* and *not guilty*; 3) Experimental: *proven* and *not proven*. In the video, the jury is instructed by the judge, several witnesses present evidence (an eyewitness, a police officer, a forensic pathologist, and a forensic scientist), and the jury hear closing statements from each of the respective counsels.

3.1.1 | Questionnaire

There were three main parts to the questionnaire: demographics; legal understanding inventory; and the post-trial questions. In the demographics part, participants were asked questions relating to their ethnicity, gender and occupation. In the legal understanding inventory, participants were asked three questions relating to the current Scottish jury system. For example, they were asked the following open-ended question: 'How many members does a Scottish jury have?'

The post-trial section was where questions relating to the main objectives of the study were asked. Here participants were asked a series of questions, using a combination of forced-choice, open-ended (free text) and 11-point Likert-type scale responses. An example of a forced-choice question was: 'Which verdict of the following is the most appropriate given the evidence?'. The options available here varied depending on the verdict condition (e.g. proven or not proven). An example of a Likert-type scale question was: 'How confident are you in the verdict you gave?', with 11 indicating certainty and zero showing that the participant was not confident at all. Finally, an example of an open-ended question was: 'What aspects of the trial led you to give this verdict? (Please specify)'.

Participants were also asked questions in relation to verdict preferences. For instance, for conviction verdicts, participants were asked the following: "Regardless of the verdict you gave, what would be your preferred type of conviction verdict? In a Scottish court, a **guilty** verdict means that the evidence has been enough to prove 'beyond reasonable doubt' that the accused person committed the crime. In a Scottish court, a **proven** verdict does not exist. If it did exist, it might mean that the evidence has been enough to prove 'beyond reasonable doubt' that the accused person committed the crime." They were then asked to select their preferred mechanism of conviction (guilty or proven). Participants were asked a similar question in relation acquittal verdicts (not guilty and not proven). These two questions were asked to target hypothesis three and four.

In addition, participants were asked questions in relation to how they perceived the beyond reasonable doubt threshold. Participants were asked to rate from 0 to 100 (0 = not at all certain; 100 = completely certain) how certain they would need to be in order to convict an accused individual. They were asked this in relation to both guilty and proven verdicts (the same definitions as used above were utilised). This question was asked in order to meet hypothesis five. Further details on specific questions are included in the Results section, below, where the associated response data are described.

3.2 | Procedure

Each participant first read an online information sheet before they gave their consent to participate. After consenting, they gave their demographic information (age, gender, ethnicity). Their understanding of legal matters was then tested by asking questions such as: 'Which of the following verdicts can a Scottish jury give? (Please select all that apply)', the answer options included fictitious and correct answers. After this, participants were presented with the mock trial. Watching the video footage took 52 min and 33 s and provided mock jurors with a realistic (albeit shortened) experience of proceedings in court. Jurors were presented with the indictment from the clerk of court, instructions from the judge, evidence from an eyewitness, a police officer, and two expert witnesses. Throughout the trial, participants were also presented with questions and statements from the prosecution and defence advocates. Once the trial had finished, participants were asked to give a verdict. The verdict options available to participants differed and were dependent on the condition they were randomly allocated to: Scottish, Anglo-American or Experimental verdict system. After giving their verdict, participants were asked six questions to ensure that they had watched and understood the video footage. If they answered more than three of the questions incorrectly, their data were excluded from the analyses.

Participants were then asked further questions about the case, such as their confidence in their verdict, and about the definitions and legal and social consequences of certain verdicts. Once this section had been completed, participants were provided with the debrief, including information on how to withdraw their consent and have their data removed if they wished to (none did), and the contact details of the principal investigator on the study.

4 | RESULTS

The first section of the results explores the influence that varying verdict systems (Scottish, Anglo-American, and Experimental) had on the likelihood of jurors convicting an accused individual (hypotheses one and two). The second section identifies which verdicts the jurors preferred and whether perceptions relating to the beyond reasonable doubt threshold differed depending on the conviction verdict that was available (hypotheses three, four and five). The final section of the results presents descriptive data about our participants' understanding of the current jury system. We had no a priori expectations or hypotheses about the latter, but the data are included here as they provide a useful insight into juror knowledge and expectations.

4.1 | Verdict system influence on convictions and acquittals

Table 1, below, presents the frequency of each verdict type that was given in each of the varying verdict systems; *note: some verdicts were only available in certain verdict systems (see design)*. In the Anglo-American verdict system, guilty and not guilty verdicts were transformed into conviction and acquittal frequencies, respectively. In the Scottish verdict system, guilty verdicts were totalled to give a conviction frequency and not guilty and not proven verdicts were integrated to give an acquittal frequency. In the Experimental verdict system, proven and not proven verdicts were transformed into conviction and acquittal frequencies, respectively.

A Pearson chi-square test of independence established a significant association between verdict systems and verdict [$\chi^2(2) = 15.83, p < 0.001, w = 0.26$]; a post-hoc power analysis also found that the result had a power of 0.96. The Scottish and Experimental verdict systems had higher acquittal rates than the Anglo-American verdict system, which correspondingly had a higher conviction rate. Statistical comparisons between pairs of verdict systems showed that jurors were significantly more likely to convict in the Anglo-American system than either the Scottish system [$\chi^2(1) = 8.72, p = 0.003, w = 0.24$] or the Experimental system [$\chi^2(1) = 13.80, p < 0.001, w = 0.31$]; these findings remain significant when adjusting the alpha level to 0.017 using a Bonferroni correction. However, jurors in the Scottish system did not differ significantly in their likelihood of giving a conviction verdict when compared with the Experimental system [$\chi^2(1) = 0.80, p = 0.37, w = 0.07$].

An additional analysis also showed that the Scottish verdict system significantly reduced the amount of *not guilty* verdicts given by jurors when compared to the Anglo-American system [$\chi^2(1) = 28.72, p < 0.001, \phi_c = 0.43$].

4.2 | Preferred verdicts and perception of beyond reasonable doubt

Binomial tests were used to test which conviction and acquittal verdicts the participants preferred. First, when focusing on conviction verdicts, participants preferred the guilty verdict ($N = 163$) significantly ($p < 0.001$) more than the proven verdict ($N = 64$). This preference was not influenced by the verdict system factor (i.e. what verdicts the participants had available to them) [$\chi^2(2) = 1.79, p = 0.41, w = 0.09$]. Second, participants did not differ significantly ($p = 0.69$) when comparing which acquittal verdict they preferred (not guilty, $N = 110$, versus not proven, $N = 117$). This preference was not influenced by the verdict system factor [$\chi^2(2) = 1.07, p = 0.59, w = 0.07$].

Due to deviation from normality, a related-samples Wilcoxon Signed Rank test was used to compare the perceived beyond reasonable doubt threshold for the proven verdict with the guilty verdict. This showed a significant difference, with the proven verdict ($Mdn = 95$) having a significantly lower perceived beyond reasonable doubt threshold than the guilty verdict ($Mdn = 99$) [$Z = -6.06, p < 0.001$]; this measure was from 0 to 100, with 100 indicating that people would need to be 100% certain before they could convict. For additional analyses, please see Appendix 1

4.3 | Juror knowledge of the current Scottish verdict system

Jurors were asked three questions relating to the current juror system. First, jurors were asked how many members were on a Scottish jury, 34 individuals correctly said 15. However, 193 individuals gave incorrect responses, with the

TABLE 1 Observed frequencies of verdicts, convictions and acquittals, by verdict system

Verdict System/Verdicts	Guilty	Not Guilty	Not Proven	Proven	Conviction frequency	Acquittal frequency
Scottish	28	4	47	N/A	28	51
Anglo-American	46	32	N/A	N/A	46	32
Experimental	N/A	N/A	50	20	20	50

minimum estimate being five and the maximum estimate being 100 (we suspect this may have been a typo for '10'). The modal response was 12, which was given by 122 participants and is the number of jurors in the Anglo-American system.

The participants were then asked to select, from a list of real and fictitious verdicts (guilty, not guilty, not proven, proven, undecided and hung), which verdicts could be given by Scottish juries. One hundred and twenty three individuals ticked combinations which were incorrect, and 104 individuals selected the correct options (and no incorrect options). Finally, participants were asked to answer a question relating to how many jurors needed to agree for a jury verdict to be reached: 126 participants gave the correct response (8 jurors), with 101 giving incorrect responses.

5 | DISCUSSION

Two main hypotheses were tested in the current study. The first hypothesis was: 1) Jurors will be more likely to convict and less likely to acquit in the proven and not proven Experimental verdict system when compared with the guilty, not guilty and not proven Scottish verdict system. This hypothesis was rejected. Instead, the findings demonstrated that jurors were no more likely to convict or less likely to acquit in either system. The second hypothesis was: 2) Jurors will be more likely to convict and less likely to acquit in the guilty and not guilty Anglo-American verdict system when compared with the guilty, not guilty and not proven Scottish verdict system. This hypothesis was accepted, replicating the finding of Ormston et al. (2019).

In addition, the Experimental system of proven and not proven was found to lead to significantly fewer convictions than the Anglo-American system, and to a similar number of convictions to the Scottish system. This indicates that the change in conviction rates identified in previous research (Curley, Murray, et al., 2021; Ormston et al., 2019) is not solely down to changing from a three to a two-verdict system; the labelling of the verdicts is also important. One explanation is that the label of '(not) proven' may emphasise the need for jurors to focus on the evidence and standards of proof, rather than more implicit, non-analytical impressions of 'guilt'. Previous research has indicated that jurors are more likely to use heuristic-based reasoning to reach guilty verdicts than to reach acquittal verdicts (Curley et al., 2018, 2019). Phrases such as 'think like a scientist' have been found to promote analytical thinking (Thomas & Millar, 2012, p. 143). Framing verdicts in terms of 'proof' may promote similar analytical decision making, as they implicitly direct jurors to be guided by the evidence and focus on the weaknesses of the prosecution's case (Hope et al., 2008). This may explain the lower conviction rate when compared to verdicts framed in terms of 'guilt', which may be perceived as being associated with emotion or morality (Curley, Munro, et al., 2021). This does not, of course, imply that all convictions are reached via emotive evaluation or that acquittals are reached through rational reasoning. Rather, it suggests that eliciting emotional rather than rational thought processes may increase the chances of a juror giving a conviction verdict. This could interact with crime type or severity, which could be an avenue for future research.

The debate regarding reform of the Scottish verdict system concerns two factors: the 'not proven' verdict specifically, and the existence of three verdicts rather than two as in the Anglo-American system (Curley, Munro, et al., 2021; Richardson & Gardiner, 2021). Our study found that the three-verdict Scottish system and the two-verdict Experimental system did not produce significantly different conviction or acquittal rates from each other. The only differences were between the Anglo-American verdict system, which does not include not proven, and the other two systems, which do. This suggests that it is the availability of the not proven verdict, regardless of number of verdicts available, that impacts conviction and acquittal rates.

Our study only investigated the proven and not proven system in physical violence cases (common assault in our pilot study and homicide in our main study). However, the low conviction rates in rape and sexual assault trials are also a prominent part of the legal reform debate (Richardson & Gardiner, 2021; Scottish Government, 2019). If the findings from our studies generalise to sexual assault and rape trials, it is unlikely that a proven and not proven system would be useful in increasing conviction rates. It is worth noting that conviction rates in sexual offence trials are also

low in other jurisdictions that already have two-verdict systems (Topping & Barr, 2020) and that the availability of the not proven verdict did not significantly influence the conviction rate in sexual assault trials in the study by Ormston et al. (2019). It may be that conviction rates in sexual offence trials are so low that experimentally determining any specific adverse effect of verdict system will be hampered by floor effects. Nevertheless, this is an empirical question and further research should investigate the utility (or otherwise) of a proven and not proven system in various trial types, including sexual offences, before policy changes are implemented or options discarded.

We also believe that it would be useful to investigate *why* the not proven verdict is used so frequently in sexual offence cases. Future research should assess this, in the context of other extra-evidential factors, such as corroboration, rape myths, and/or a lack of legal education about verdict options (see *Juror knowledge of the current Scottish verdict system section*). Currently the not proven verdict's role in the low conviction rates in rape and sexual assault trials is, ironically, not proven.

In addition to the two main hypotheses, three additional hypotheses were also tested. The third hypothesis was: In convictions, jurors will show a preference between guilty and proven verdicts. The hypothesis was accepted, with jurors showing a preference for the guilty over the proven verdict. The fourth hypothesis was: In acquittals, jurors will show a preference between not guilty and not proven verdicts. This hypothesis was rejected, and jurors had no preference regarding which acquittal verdict they favoured. The fifth hypothesis was: Jurors will report differing levels of perceived reasonable doubt when delivering a guilty (conviction) verdict when compared to a proven (conviction) verdict. This hypothesis was accepted, as the proven beyond reasonable doubt threshold was perceived to be lower than the guilty beyond reasonable doubt threshold.

Further, the data from the main study showed that jurors strongly preferred the guilty verdict to the proven verdict for convicting the accused, but were split on their preferred acquittal verdict. Jurors may find it more comfortable to indicate their belief in the 'guilt' of the accused rather than on the 'proof' provided by the prosecution's evidence. This is counter to both their intended role and the views of legal professionals (Curley, Munro, et al., 2021). The connotations of 'guilty' may be seen by jurors to reflect a more punitive outcome for the accused than 'proven' (Curley, Munro, et al., 2021), which may reflect how they see their role (Son et al., 2019). A proven verdict might be seen by some as being 'too soft' on an accused individual who they perceive as morally guilty. Interestingly, though, jurors had no preference in regards to which acquittal verdict they preferred, which may suggest that they care more about how the convicted are labelled than the acquitted.

Jurors may also see a guilty verdict as giving definitive closure (i.e. the accused definitely did it), whereas a proven verdict may suggest less certainty. If this were the case, then it might be expected that jurors would need to be more certain to give a guilty verdict than they would in order to give a proven verdict. Our jurors rated their beyond reasonable doubt threshold significantly lower for a proven verdict when compared to the guilty verdict, supporting this suggestion. However, this difference in reasonable doubt threshold did not correspond to the conviction rate or perception of guilt (see Appendix 1) data, as jurors were more likely to convict in the Anglo-American verdict system than the Experimental verdict system. Previous research has found that jurors do not have a clear concept of what the (somewhat nebulous) legal standard of 'beyond reasonable doubt' really means (Kramer & Koenig, 1990). In our study, jurors may have not only had a lack of clarity about 'beyond reasonable doubt' but also how 'proven' relates to that legal standard. This contrasts with the views of Scottish legal professionals, who were more in agreement on the meaning of a proven verdict than they were for any existing verdict (Curley, Munro, et al., 2021). Further research directly exploring juror perceptions of a 'proven' verdict, for example, relating it explicitly to concepts of 'reasonable doubt', would be informative.

In addition, a proven verdict may also remove the juror's ability to decide on the morality of the accused (Barbato, 2004; Curley, Munro, et al., 2021; Jackson, 1998). Jurors may use their role as 'masters of fact' to not only decide on the legal guilt of the accused, but to also judge the actions of the accused in relation to their own morality (Curley, Munro, et al., 2021). In other words, the crown may have failed their case but the jury still believed the accused to be morally guilty, which may then still lead them to reaching a guilty verdict. The mock jurors in this

study may have felt that reform from a guilty verdict system to a proven verdict system attenuated their ability to use subjective moral beliefs in their decision making and also attenuated the power of the jury altogether.

Finally, our study also found that jurors did not find the current Scottish verdict system any more confusing than the Anglo-American verdict system (see Appendix 1), which runs counter to the finding of Ormston et al. (2019). It may be that the coverage of the reform debate in the media made our participants more familiar with the Scottish verdict system than were Ormston et al.'s participants. However, this seems unlikely as the majority of participants were unable to identify the correct verdict options that would be available to them in a Scottish courtroom. A pressing issue for Scottish justice, then, may be the inadequacy of education on Scots law for the general public. This is likely exacerbated by over-representation of the Anglo-American legal system in media. Just as exposure to forensic dramas has led to concerns about a 'CSI Effect' on jurors (Kim et al., 2009), exposure to Anglo-American legal dramas may have a Scottish specific 'Law & Order Effect' that influences jurors' expectations of trials. In Scotland specifically, this could cause jurors to have the wrong expectations of jury size, required majority, and, of course, the available verdicts.

5.1 | Limitations, recommendations, and future research

A notable limitation of the current study is the presence of misunderstandings about aspects of the Scottish legal system demonstrated by our sample of participants. Despite being Scottish, the vast majority of our participants did not know basic details of how the Scottish legal system works (e.g. number of jurors, verdicts available). This general lack of understanding of the legal system perhaps outweighs any lack of understanding of any individual components of it, specifically verdict options. Can a legal system work effectively, when the people who have to make decisions do not know or understand the mechanisms and principles by which those decisions should be made? We therefore recommend investment in more, and better, legal education, both as part of the core curriculum of modern studies in schools and as a targeted intervention for people when they are called for jury duty.

A second limitation was that the study was conducted online rather than in a live setting. The research team was unable to run the study face-to-face due to Covid-19 pandemic restrictions at the time of data collection. This eliminated the opportunity to ask participants ad hoc follow-up questions during de-briefing, for example. It also meant that we were unable to control the participants' environment during the study, such as eliminating distractions. However, our attention-check questions should have identified any participants who did not pay full attention to the trial video, so that they could be excluded from the analysis. Online studies may also have sampling problems, if the online population is not representative of the target population in a relevant characteristic. However, we used an online recruitment platform (Prolific) which takes measures to provide a representative sample of a population by stratifying the intended sample size across age, sex and ethnicity to the same proportions of the national population (in the case of the current paper, Scotland via reference to the UK Office of National Statistics; Prolific 2022;). We can therefore be reasonably confident that our sample is broadly representative of our target population (i.e. potential Scottish jurors).

It is worth noting that, when we were designing our main study, Scottish Courts and Tribunals were already considering conducting virtual trials using video technology (Scottish Courts and Tribunals, 2020a). They subsequently began running such trials in order to keep the legal system operational during the Covid-19 pandemic (Scottish Courts and Tribunals, 2020b). Except for the length of a typical trial, which would typically be several days to weeks, the experience of mock jurors in our study would not have differed dramatically from real jurors in virtual trials. Furthermore, the trial video was developed based on real case information, in conjunction with legal advisors, and key roles were played by legal professionals. The material therefore has high ecological validity and reflects the type of video-mediated trial that real jurors may now experience.

A third limitation is that, despite our ecologically valid materials, there are some factors that our research does not address: 1) lack of consequentiality (i.e. our jurors knew that their verdicts were not going to directly affect real

people's lives); 2) length of trial (as noted above); 3) the absence of deliberations among jurors. Each of these issues has been critiqued in the jury literature before (Curley et al., 2022; Diamond, 1997; Krauss & Lieberman, 2017) and they do decrease the generalisability of jury studies.

Consequentiality and length of a trial are currently irresolvable issues, due to legal, ethical and practical constraints. For example, it would not be ethical to recruit research participants to mock trials that could last for weeks or months, nor would it be practical to pay their expenses to do so (Diamond, 1997; Krauss & Lieberman, 2017). It is currently not legal to research real jurors during a trial, so consequentiality will always be lacking from jury research until the law changes. We share the view of Krauss and Lieberman (2017) that jury research will always be imperfect and that the best solution is to gather data using multiple methods. If similar results are found by triangulating the data, then the conclusions and recommendations are likely to be valid, reliable and useful (Bekhet & Zauszniewski, 2012; Diamond, 1997; Krauss & Lieberman, 2017). Indeed, this is why we have conducted surveys of legal professionals in addition to our experimental work (Curley, Munro, et al., 2021).

Deliberations, on the other hand, are a viable area for research, although they have only rarely been included in study designs (Curley, 2021; Curley et al., 2022; Diamond, 1997; Krauss & Lieberman, 2017). This is largely due to time, costs and complexity (e.g. arranging groups of 12, or 15 in Scotland, jurors to participate simultaneously in a study). Due to this added complexity, some authors have argued that jury researchers should start with relatively simplified designs to identify variables of interest (Diamond, 1997; Krauss & Lieberman, 2017). Follow-up studies can then further investigate those variables using more complex designs, which may trade-off some experimental control (or even sample size) against greater ecological validity (e.g. including open-ended deliberations). Our present study is located towards the beginning of this process, and the influence (if any) of deliberations under different verdict systems will form part of a future research programme.

6 | CONCLUSIONS

In conclusion, the proven and not proven (Experimental) verdict system led to significantly fewer convictions than the guilty and not guilty (Anglo-American) verdict system, but a similar amount of convictions to the guilty, not guilty and not proven (Scottish) verdict system. This shows that verdict labels affect conviction rates separately from the number of verdicts available to jurors. One explanation for these findings may be that proven and not proven verdicts promote analytical thinking in jurors and/or a shift of the decision criteria for a conviction, thus decreasing jurors' probability of giving a guilty verdict. Future research should include rape and sexual assault trials to explore the role of the not proven verdict in Scotland's low conviction rate in such cases. An important, and unanticipated, finding was that our participants lacked basic knowledge and understanding of the Scottish jury system, despite all being jury-eligible Scots. We therefore strongly recommend improvements in education and communication about this important aspect of citizenship, both in modern studies classes in Scottish schools and specifically for those called for jury duty.

ACKNOWLEDGEMENT

The researchers would like to acknowledge the Modern Studies Association of Scotland, Bloody Scotland and the Faculty of Advocates. These organisations developed the materials and without their help, we would have not been able to conduct this research. Thank you! No funding was attached to this project.

CONFLICT OF INTEREST

All Authors declare no conflicts of interest.

REFERENCES

- Andrade, C. (2020). The limitations of online surveys. *Indian Journal of Psychological Medicine*, 42(6), 575–576.
- Barbato, J. M. (2004). Scotland's Bastard Verdict: Intermediacy and the unique three-verdict system. *Indiana International & Comparative Law Review*, 15, 543–581. <https://heinonline.org/HOL/LandingPage?handle=hein.journals/iicl15&div=25&id=&page=>
- Bekhet, A. K., & Zauszniewski, J. A. (2012). Methodological triangulation: An approach to understanding data. *Nurse Researcher*, 20(2), 40–43. <https://doi.org/10.7748/nr2012.11.20.2.40.c9442>
- Curley, L. J. (2021). Juror and jury decision making. In A. J. Turner, C. Hewson, K. Mahendran, & A. Strathie (Eds.), *Living psychology: From the everyday to the extraordinary* (pp. 361–394). Oxford University Press/The Open University.
- Curley, L. J., MacLean, R., Murray, J., Laybourn, P., & Brown, D. (2019a). The bastard verdict and its influence on jurors. *Medicine, Science & the Law*, 59(1), 26–35.
- Curley, L. J., MacLean, R., Murray, J., Pollock, A. C., & Laybourn, P. (2019b). Threshold point utilisation in juror decision-making. *Psychiatry, Psychology and Law*, 26(1), 110–128. <https://doi.org/10.1080/13218719.2018.1485520>
- Curley, L. J., Munro, J., & Dror, I. (2022). Cognitive and human factors in legal layperson decision making: Sources of bias in juror decision making. *Medicine, Science, and the Law*.
- Curley, L. J., Munro, J., Frumkin, L. A., & Turner, J. (2021). Informing reform: The views of legal professionals on the unique aspects of Scottish Law. *Medicine, Science & the Law*, 61(4), 256–265. <https://doi.org/10.1177/00258024211992913>
- Curley, L. J., Murray, J., MacLean, R., Laybourn, P., & Brown, D. (2018). Faith in thy threshold. *Medicine, Science & the Law*, 58(4), 239–250. <https://doi.org/10.1177/0025802418791062>
- Curley, L. J., Murray, J., MacLean, R., Munro, J., Lages, M., Frumkin, L. A., & Brown, D. (2021). Verdict spotting: Investigating the effects of juror bias, evidence anchors and verdict system in jurors. *Psychiatry, Psychology and Law*, 1–22. <https://doi.org/10.1080/13218719.2021.1904450>
- Diamond, S. S. (1997). Illuminations and shadows from jury simulations. *Law and Human Behavior*, 21(5), 561–571.
- Duff, J. (2021, April 4). Nicola Sturgeon says Scotland's not proven verdict to face review [online]. *The National: news*.
- Duff, J., & Findlay, M. (1997). Jury reform: Of myths and moral panics. *International Journal of the Sociology of Law*, 25(363), 367–370.
- Estrada-Reynolds, V. C., Gray, J. M., & Nuñez, N. (2015). Information integration theory, juror bias, and sentence recommendations captured over time in a capital trial. *Applied Cognitive Psychology*, 29(5), 713–722. <https://doi.org/10.1002/acp.3155>
- Hope, L., Greene, E., Memon, A., Gavisk, M., & Houston, K. (2008). A third verdict option: Exploring the impact of the not proven verdict on mock juror decision making. *Law and Human Behavior*, 32(3), 241–252. <https://doi.org/10.1007/s10979-007-9106-8>
- Jackson, B. S. (1998). Truth or proof?: The criminal verdict. *Revue Internationale de Semiotique Juridique*, 11(3), 227–273. <https://doi.org/10.1007/BF01110409>
- Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review*, 93(5), 1449–1475. <https://doi.org/10.1257/000282803322655392>
- Kim, Y. S., Barak, G., & Shelton, D. E. (2009). Examining the 'CSI-effect' in the cases of circumstantial evidence and eyewitness testimony: Multivariate and path analyses. *Journal of Criminal Justice*, 37(5), 452–460. <https://doi.org/10.1016/j.jcrimjus.2009.07.005>
- Kramer, G. P., & Koening, D. M. (1990). Do jurors understand criminal jury instructions – Analyzing the results of the Michigan Juror Comprehension Project. *U. Mich. JL Reform*, 23, 401–440. <https://repository.law.umich.edu/cgi/viewcontent.cgi?article=1779&context=mjlr>
- Krauss, D. A., & Lieberman, J. D. (2017). Managing different aspects of validity in trial simulation research. In M. B. Kovera (Ed.), *The psychology of juries* (pp. 185–205). American Psychological Association. <https://doi.org/10.1037/0000026-009>
- Lecci, L., Beck, C., & Myers, B. (2013). Assessing pretrial juror attitudes while controlling for order effects: An examination of effect sizes for the RLAQ, JBS, and PJAQ. *American Journal of Forensic Psychology*, 31(3), 41–66. <https://psycnet.apa.org/record/2013-27759-003>
- Lecci, L., & Myers, B. (2008). Individual differences in attitudes relevant to juror decision making: Development and validation of the pretrial juror attitude questionnaire (PJAQ) 1. *Journal of Applied Social Psychology*, 38(8), 2010–2038. <https://doi.org/10.1111/j.1559-1816.2008.00378.x>
- McKenzie, J. P. (1985). The editorial notebook; between guilt and innocence. *New York Times*, A28.
- Ormston, R., Chalmers, J., Leverick, F., Munro, V., & Murray, L. (2019). Scottish jury research: Findings from a large-scale mock jury study. *Crime and Justice*. <https://www.gov.scot/binaries/content/documents/govscot/publications/research-and-analysis/2019/10/scottish-jury-research-findings-large-mock-jury-study/documents/scottish-jury-research-findings-large-scale-mock-jury-study/scottish-jury-research-findings-large-scale-mock-jury-study/govscot%3Adocument/scottish-jury-research-findings-large-scale-mock-jury-study.pdf>
- Richardson, A., & Gardiner, R. (2021, April 6). Not proven – is evidence weighted against Scotland's 3-verdict justice system? [online]. *The Courier: News*. [thecourier.co.uk](https://www.thecourier.co.uk) [Accessed 29 April 2021].

- Scottish Courts and Tribunals. (2020a). *Extending virtual summary trials in Scotland*. Retrieved July 3, 2020, from <https://scotcourts.gov.uk/about-the-scottish-court-service/scs-news/2020/07/03/extending-virtual-summary-trials-in-scotland>
- Scottish Courts and Tribunals. (2020b, September 28). *First cinema remote jury centre trial set to start*. Retrieved July 28, 2020, from <https://www.scotcourts.gov.uk/about-the-scottish-court-service/scs-news/2020/09/28/first-cinema-remote-jury-centre-trial-set-to-start>
- Scottish Executive. (2006). *Criminal proceedings in Scottish courts: Statistical Bulletin: CrJ/ 2006/3*. <http://www.scotland.gov.uk/Publications/2006/04/25104019/11>
- Scottish Government. (2019). *Criminal proceedings in Scotland 2017-2018*. In *Criminal proceedings in Scotland 2017-2018 - gov.scot* (www.gov.scot) Accessed 29 April 2021.
- Smith, J. I., & MacDonald, I. (1958). *An introduction to Scottish legal history* (Vol. 20). Stair Society.
- Smithson, M., Deady, S., & Gracik, L. (2007). Guilty, not guilty, or...? Multiple options in jury verdict choices. *Journal of Behavioral Decision Making*, 20(5), 481–498. <https://doi.org/10.1002/bdm.572>
- Son, J. Y., Bhandari, A., & FeldmanHall, O. (2019). Crowdsourcing punishment: Individuals reference group preferences to inform their own punitive decisions. *Scientific Reports*, 9(1), 1–15. <https://www.nature.com/articles/s41598-019-48050-2>
- Statista. (2018). Distribution of non-white ethnic backgrounds in Scotland in 2018. *Statista* [Internet]. <https://www.statista.com/statistics/367842/scotland-ethnicity-of-population/>
- Thomas, A. K., & Millar, P. R. (2012). Reducing the framing effect in older and younger adults by encouraging analytic processing. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 67(2), 139–149. <https://doi.org/10.1093/geronb/gbr076>
- Topping, A., & Barr, C. (2020, July 30). *Rape convictions fall to record low in England and Wales [online]*. *The Guardian*: UK Politics. Available from: *Rape convictions fall to record low in England and Wales | Rape and sexual assault | the Guardian* [Accessed 29 April 2021].
- Walker, D. M. (1988). *A legal history of Scotland. 1. The beginnings to AD 1286*. Green.

How to cite this article: Curley, L. J., Munro, J., Turner, J., Frumkin, L. A., Jackson, E., & Lages, M. (2022). Proven and not proven: A potential alternative to the current Scottish verdict system. *Behavioral Sciences & the Law*, 1–15. <https://doi.org/10.1002/bsl.2568>

APPENDIX 1

Additional analysis was conducted on a number of measures. Shapiro-Wilk's tests suggest that none of the below measures are normally distributed across the verdict systems. Therefore, seven Kruskal-Wallis tests were conducted. None of the differences were significant; for exact *p* values, see Table A1.

TABLE A1 Verdict system influence on decision confidence, perception of guilt and other additional measures

Measure	Scottish Verdict System Mean (SD)	Anglo-American Verdict system Mean (SD)	Experimental Verdict system Mean (SD)	<i>p</i> value
Decision point (when a decision could be made)	6.15 (1.85)	6.03 (1.59)	6.59 (1.58)	0.157
Decision confidence	7.60 (1.95)	7.58 (2.10)	7.58 (1.83)	0.910
Perception of guilt	6.85 (2.61)	7.43 (2.36)	6.59 (2.57)	0.123
Decision ease	5.42 (2.87)	6.03 (2.65)	5.99 (2.62)	0.352
Negatively emotionally affected from giving verdict	4.91 (3.05)	5.21 (3)	4.77 (2.89)	0.668
Verdict system confusion	3.04 (2.83)	2.79 (2.58)	3.48 (2.73)	0.401
Confidence in verdict system	6.35 (2.30)	6.22 (2.50)	6.36 (2.16)	0.990

Results of logistic regression (R Output)

Call: `glm(formula = verdict ~ percept_of_guilt + verdict_sys, family = binomial, data = df)`

Deviance residuals:

Min	1Q	Median	3Q	Max
-2.04778	-0.38560	-0.05803	0.51204	2.74877

Coefficients:

	Estimate	Std.Error	z value	Pr(> z)
(Intercept)	-9.9651	1.4436	-6.903	5.09e-12***
percept_of_guilt	1.1931	0.1674	7.128	1.02e-12***
verdict_sys2	1.4381	0.4954	2.903	0.0037**
verdict_sys3	-0.1961	0.4882	-0.402	0.6880

–

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 307.96 on 226 degrees of freedom

Residual deviance: 159.08 on 223 degrees of freedom

AIC: 167.08

Pseudo R²: 1-(fit0a\$deviance/fit0a>null.deviance)

[1] 0.4834379