Cabinet Secretary for Justice and Home Affairs
Angela Constance MSP

F/T: 0300 244 4000
E: scottish.ministers@gov.scot

Scottish Government
Riaghaltas na h-Alba

Convener
Criminal Justice Committee
The Scottish Parliament
Edinburgh
EH99 1SP

Justice.committee@parliament.scot

16 February 2024

Victims, Witnesses, and Justice Reform (Scotland) Bill

Dear Convener,

I welcomed the opportunity to give evidence to the Committee earlier this month on Parts 4 to 6 of the Victims, Witnesses, and Justice Reform (Scotland) Bill, and to set out why I believe it is so vitally important that we take the opportunity these reforms offer to transform the experiences of people going through our justice system.

During the session I undertook to provide the Committee with further information on several areas in which Members had expressed an interest.

Not proven – meta analysis by Jackson et al.

The Committee asked about whether the abolition of the not proven verdict should be a stand-alone reform, or whether it should be accompanied by an increase in the majority required for conviction. Both the Lord Advocate and Rape Crisis Scotland had previously set out concerns in their evidence to the Committee that moving to a qualified majority could make it harder to secure convictions, particularly for sexual crimes.

I take this view very seriously, as I know the Committee do, and I am very keen to ensure we listen to the full range of views. As you are aware, the Senators, legal academics, and the defence community argue that to maintain a simple majority in a two verdict system could lead to an increased risk of miscarriage of justice – the latter two groups suggesting that we should therefore consider setting a higher threshold of ten out of twelve for a conviction.

However, as I said at Committee, independent evidence suggests that moving to two verdicts will lead to an increase in convictions in finely balanced cases. I referenced the jury research which has received much focus in the Committee, and also a more recent meta-analysis by Jackson et al (2024) (which includes the jury research findings in its analysis) and I have attached separately for your consideration. As with all research, there are limitations with the work and care should be taken when considering the extent to which the

findings may be reflected in actual cases. However, the Jackson et al findings are clear across a range of different offences including physical assault, rape and homicide:

"there is a statistically significant effect towards lower conviction rates under the Scottish three-verdict system than under an Anglo-American two-verdict system".

The question for juries to answer, “has the Crown proved the accused’s guilt beyond a reasonable doubt?”, would not change as a result of a move to a two verdict system. However the independent evidence referred to, indicates that changing the structure of the system and the verdicts available is likely to have an impact on case outcomes.

This is also consistent with concerns about the use and understanding of the not proven verdict. Some jurors may use the not proven verdict as a compromise verdict or as a way to ‘sit on the fence’. The evidence suggests, that when this option is removed and jurors are required to make a binary decision between two verdicts, the behaviour of jurors will be affected. This would have an impact in the cases which currently result in a not proven verdict, and also in the range of cases where the existence of the not proven verdict makes a not guilty verdict more likely. For example, a 15 person jury which is split 7 guilty, 5 not guilty, and 3 not proven, would currently result in a not guilty verdict. Under a two verdict system that continued to require a simple majority (8 jurors from 15) for a conviction, a conviction would be returned, if just one of those jurors who would have previously voted for “not proven”, instead voted for “guilty”.

It also noteworthy that there are no other similar legal systems with two verdicts, where they also use a simple majority. Qualified majorities are common in England and Wales, Ireland, New Zealand and Australia, while unanimity is required in Canada, and most states in the USA.

In developing the Bill, we have been mindful that any changes to verdicts and juries would apply to all types of crime, so reforms to the jury system cannot be considered only in the context of particular types of offending.

Much of the rest of the Bill contains ambitious provisions that focus specifically on improving the system for complainers in sexual offence cases. I encourage Members to work with me on making those aspects of the Bill as strong as possible, to ensure that the justice system is compassionate, effective and trauma-informed to support complainers to give their best evidence and to hold perpetrators to account.

As part of our proposals to abolish the not proven verdict, we also propose providing the additional safeguard of a qualified majority that many stakeholders consider critical to maintaining a fair system; without going as far as the thresholds typical in other countries, to reflect the unique additional safeguards that we have in Scotland such as the corroboration rule.

There is no doubt that these are complex, finely-balanced issues upon which there are a range of views, and the responsibility falls to Parliamentarians to navigate through these to find the most appropriate resolution. I look forward to hearing your views in due course.

Rights of audience in the Sexual Offences Court

I also committed to give members further information on the proposed rights of audience for the Sexual Offences Court and to provide clarity on the considerations that have informed our position on this.

Currently, the ability of a defence agent to appear in court to represent an accused is determined by whether they hold the rights of audience to appear in the particular court in which the case is being prosecuted. Solicitors may appear in the Justice of the Peace and sheriff courts but unless they have extended rights of audience (i.e. they are a solicitor advocate) they are unable to appear in the High Court. Accordingly, solicitors without extended rights cannot appear in cases which involve an offence of rape or murder, as well as certain other offences which are within the exclusive jurisdiction of the High Court.

By contrast, the extended rights of audience granted to advocates and solicitor advocates allow them to appear in any of Scotland's criminal courts including the High Court. Advocates and solicitor advocates may also appear in the lower courts and an accused can, subject to means testing, apply to the Scottish Legal Aid Board (SLAB) for funding to employ counsel to represent them in the lower courts.

The Committee may wish to note that the rules relating to rights of audience do not apply to prosecutors. The Lord Advocate, in their role as independent head of the prosecution service, is not restricted in who they can appoint to prosecute a case on their behalf. The independence of the Lord Advocate in discharging their prosecutorial functions is enshrined in section 48(5) of the Scotland Act 1998. It is therefore not within the legislative competence of the Scottish Parliament to seek to place any obligations or restrictions on those who prosecute on behalf of the Lord Advocate.

Lady Dorrian's Review recommended that only advocates and solicitor advocates, who have completed the requisite trauma informed training, should have rights of audience to appear in the Sexual Offences Court.

As my officials outlined at Committee, the decision to depart from this recommendation was informed by the findings of a cross justice-sector working group established to consider the operation of the Court. The group had concerns with the proposal to restrict solicitors from appearing in the Court. The jurisdiction of the Court, which brings together all solemn level sexual offending prosecuted in the sheriff and jury courts and the High Court would require advocates or solicitor advocates to appear in a far greater number of cases than at present (the Financial Memorandum refers to over 300 additional cases per year). This could not be easily absorbed by the current cadre of advocates and solicitor advocates and would undoubtedly lead to delays in cases coming to trial.

A further concern identified by the Group was that this approach would restrict opportunities for solicitors *'who might otherwise have conducted certain sexual offence trials before a solemn sheriff and jury court, to maintain and develop their skills'* impacting on the capacity of solicitors, some of whom will become the next generation of advocates, solicitor advocates, and judges, from building their knowledge and experience of sexual offence cases.

I share the concerns raised by the Working Group. In my view, it would be counterproductive to risk delays to trial by removing solicitors' rights to appear in cases for which they currently

provide an appropriate level of representation. I also see merit in having solicitors in the Court, specially trained in trauma informed practice, building skills and experience and developing their practice in dealing with these types of cases.

The Bill therefore enables advocates, solicitor advocates and solicitors, all subject to completion of the requisite trauma-informed training, to appear in the Court. The exception to this is in cases involving an offence of rape or murder which remain restricted to advocates and solicitor advocates. This approach reflects that rape and murder are the only offences within the jurisdiction of the Court that must currently be prosecuted in the High Court.

I do, however, appreciate the concerns raised by members at Committee and recognise the importance of ensuring that accused in the Sexual Offences Court have access to the same level of representation as under existing court structures.

I recognise that the High Court currently hears cases involving serious sexual offences which do not include rape or murder. The complexity in identifying those cases for the purpose of restricting rights of audience is that almost all of the other offences that appear in the High Court do so following the exercise of prosecutorial discretion and therefore the actual offence itself does not provide a definitive or meaningful indication of whether it will be prosecuted in the High Court or sheriff and jury court. Prosecutors acting in the public interest will carefully consider all of the relevant facts and circumstances of a case before making a decision as to whether prosecution is appropriate and, if so, which court it should be prosecuted in.

We are working with justice partners to identify a suitable mechanism to ensure that an appropriate test is developed that would allow, as far as possible, for 'otherwise High Court cases' to be identified and made eligible for Legal Aid for advocates or solicitor advocates. I have asked my officials to explore whether this mechanism can be placed in primary legislation to provide the reassurance Members are seeking in respect of this important aspect of the Court. I will, of course, give careful consideration to any recommendations that the Committee may make on this issue in its Stage 1 Report.

Pilot of single judge trials

Operational detail on the pilot

As the Committee will be aware, in 2022 a cross-sector Working Group was established to consider how to take forward Lady Dorrian's recommendation to hold a pilot of single judge rape trials. The Group developed a proposed model for the pilot, published in December 2022, which I have enclosed with this letter. The model sets out how the Group envisaged that the pilot would operate at each stage of the court process, from the determination of whether a case met the criteria for the pilot until the disposal of the case.

A key principle of the proposed model is that changes to existing processes for rape cases should be kept to a minimum. The intention is that hearings and trials in the pilot would follow existing, familiar procedures as far as possible, with changes only made where they are necessary to reflect the fact that the fact finder and decision maker is a judge rather than a jury. When we are evaluating the pilot, this approach will enable us to better understand the impact of changing the decision maker.

The Working Group recognised that further detailed consideration would be needed to refine its proposals, but I hope that the draft model helps to give the Committee a more concrete idea of what the pilot would look like in practice. We will continue to collaborate with partners from across the justice sector to develop the operational detail of the pilot - as I emphasised during my evidence session, I am keen to build as much consensus on that as possible. While it is appropriate that procedural details be set out in regulations rather than in primary legislation, as our analysis and our engagement with partners progress I will undertake to amend the Bill to include more information on the pilot, to provide MSPs with greater clarity on key issues like the case criteria. I also look forward to hearing the Committee's views on the pilot, which will inform the approach we take.

Evaluation of the pilot

The Working Group also explored how the pilot should be evaluated. As I set out during my evidence, it identified that the pilot's core objectives should be to:

1. assess **how the process of conducting a single judge trial is perceived** by those involved in the trial process. For example, do practitioners find there are differences in the way evidence is led, in cross-examination, or in how often judges intervene? Do accused people and complainers find written reasons for verdicts beneficial? How do complainers experience examination in chief and cross-examination?

2. explore the impact of single judge trials on the **effectiveness and efficiency** of managing rape trials. For example, are trials shorter? Are there any differences in frequency of adjournments, or in numbers of s.275 applications?

3. consider the impact of single judge trials on **outcomes**. For example, how often are guilty pleas entered, and when in the process? What is the conviction rate and what are the themes in reasons for convictions and aquittals, as set out in written reasons? What are the main reasons for any appeals lodged?

The Working Group was clear that evaluation of the pilot should be weighted towards Objective 1. Its report noted that, "*the primary (measurable) motivation behind the piloting of single judge trials is to improve the experience of complainers in sexual offence cases. While Objectives 2 and 3 are important, these should be considered as being less significant in determining whether the pilot should be considered a success".*

An annex to the Group's report set out in detail an indicative series of questions and methods that could be used to evaluate the pilot against each of its research objectives. I have enclosed a copy of the Annex with this letter, and I hope that it gives Members a helpful idea of the potential scope of the evaluation, and the kind of qualitative and quantitative data that could be gathered. We plan to work with analysts and with stakeholders to develop the final research specification for the pilot, to ensure that it is designed as robustly as possible and reflects views from across the sector. While it would not be appropriate to include detailed research questions in primary legislation, I am happy to consider whether we could amend the Bill to make explicit that the review of the pilot under section 66 would need to cover those three themes identified.

Yours sincerely,

ANGELA CONSTANCE MSP

St Andrew's House, Regent Road, Edinburgh EH1 3DG
www.gov.scot

Routledge
Taylor & Francis Group

Psychiatry, Psychology and Law

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tppl20

The effect of verdict system on juror decisions: a quantitative meta-analysis

Elaine Jackson, Lee Curley, Fiona Leverick & Martin Lages

To cite this article: Elaine Jackson, Lee Curley, Fiona Leverick & Martin Lages (11 Jan 2024): The effect of verdict system on juror decisions: a quantitative meta-analysis, Psychiatry, Psychology and Law, DOI: 10.1080/13218719.2023.2272912

To link to this article: https://doi.org/10.1080/13218719.2023.2272912

Psychiatry, Psychology and Law, 2024
Vol. 0, No. 0, 1–19, https://dx.doi.org/10.1080/13218719.2023.2272912

The effect of verdict system on juror decisions: a quantitative meta-analysis

Elaine Jackson[a]*, Lee Curley[b], Fiona Leverick[c] and Martin Lages[a]

[a]*School of Psychology and Neuroscience, University of Glasgow, Glasgow, UK;* [b]*School of Psychology, The Open University, Milton Keynes, UK;* [c]*School of Law, University of Glasgow, Glasgow, UK*

We study the effect of the Scottish three-verdict system (guilty, not guilty, not proven) and the Anglo-American two-verdict system (guilty, not guilty) on juror decisions by combining data sets from 10 mock trials reported in suitable studies. A logistic regression with random effects uses the exact number of convictions and acquittals in 10 mock trials from a total of 1778 jurors to reliably estimate the effect of verdict system. We found a statistically significant verdict effect suggesting that the odds for a conviction by a juror are about 0.6 times or 40% lower under the three-verdict system than under a conventional two-verdict system. Possible explanations and implications of this verdict effect are discussed. This finding helps to better understand juror decision making in the context of the current reform of the Scottish three-verdict system into a two-verdict system.

Keywords: cognitive bias; conviction rate; juror decisions; logistic regression; not-proven verdict; random effects; three-verdict system; two-verdict system.

Article History: Received 28 May 2023; Accepted 28 September 2023

Introduction

Scotland has a separate and distinct legal system from the rest of the UK. One of the key differences is in respect of its criminal verdict system. Like many other legal systems built on the Anglo-American model, it uses juries to determine guilt in criminal cases. But rather than the two-verdict system (*guilty* and *not guilty*), utilised by the majority of Anglo-American countries, the Scottish system (*guilty*, *not guilty* and *not proven*) offers jurors the choice of three verdict options (Chalmers et al., 2021a, 2022). The *not proven* verdict operates as a second acquittal verdict, with no legal definition, and judges are dissuaded from trying to define the verdict to the jury (Curley et al., 2022). Legally the *not proven* verdict has exactly the same effect as the *not guilty* verdict, and there is no distinction in law between the two verdicts of acquittal, with both verdicts being given when the prosecution has failed to prove their case beyond reasonable doubt. Jurors are simply told that they have two alternative options with the same consequences (Judicial Institute for Scotland, 2022). Furthermore, the Scottish jury system uses juries with 15 members (rather than 12), and it reaches a jury verdict via a simple majority rule, rather than requiring (almost) unanimity amongst jurors (Chalmers et al., 2020).

Correspondence: Martin Lages, School of Psychology and Neuroscience, University of Glasgow, 62 Hillhead Street, Glasgow, G12 QB, Scotland, UK. E-mail: Martin.Lages@glasgow.ac.uk
*Additional affiliation: School of Business and Creative Industries, University of West of Scotland, Paisley, UK.

In recent years, there has been increased scrutiny of the *not proven* verdict, particularly by charitable organisations and activists (Rape Crisis Scotland, 2021) who claim that the verdict is confusing and that it leads to a disproportionate number of acquittals in sexual violence trials. In 2022, the Scottish Government announced that it would introduce legislation to end the use of the *not proven* verdict (Scottish Government, 2022). This comes at a time when there is increased interest in the *not proven* verdict elsewhere, with proponents of the Scottish verdict system arguing that *not proven* – or something similar – should be introduced into legal systems that operate without it (e.g. Allan, 2017; Phalen, 2018; Picinali, 2022). One of the key difficulties in assessing the case for the *not proven* verdict is that until recently, there had been relatively few empirical studies that have investigated the effect of verdict system on juror decision making.

Now, however, a sufficient number of experimental studies comparing the two verdict systems is available. The aim of the current study is therefore to conduct a quantitative meta-analysis on the results from carefully selected studies that investigated the differential effect of verdict systems on juror decisions – specifically, the impact of the third verdict option of *not proven* on juror preferences for conviction. Although there is evidence from individual studies that the Scottish verdict system reduces convictions, other studies suggest that the availability of the *not proven* verdict does not significantly affect conviction rates. This provides a compelling rationale for undertaking a meta-analysis that addresses controversies arising from conflicting or ambiguous claims.

Background

The *not proven* verdict has a long history in the Scottish legal system. The original verdicts in Scottish criminal cases were essentially *guilty* and *not guilty*. However, there was a lack of consistency in what they were called, with guilt being declared through terms such as 'had done wrangis' (scots) or 'convictus' (latin), and innocence being delivered through names such as 'made qwyt' or 'clene and sakles' (scots; Barbato, 2004). In the early seventeenth century, a change in procedure meant that jurors did not give general verdicts on criminal cases, meaning they no longer made decisions in relation to the guilt of the accused. Instead, they were asked to declare whether various facts were proven or not proven (in other words, they gave special verdicts), and then the judge would declare the accused innocent or guilty on the basis of these factual findings. This procedure eventually disappeared, and juries once again became the body that pronounced a general, or overall, verdict on the case. But the terminology of *not proven* remained, and juries started to use this as a general verdict, alongside *guilty* and *not guilty* (Chalmers et al., 2021a; Curley et al., 2022; Willock, 1966).

Although Scotland is unusual in having a differentiated verdict system, it is not unique. In Israel, there are also two acquittal verdicts – a full acquittal and a 'doubt-based acquittal', in which doubt exists regarding the acquitted person's innocence (Vaki & Rabin, 2021). There was also a case where a Washington state judge gave a *not proven* over a *not guilty* verdict (Bray, 2005). Until a major reform of the Code of Criminal Procedure for Italy in 1988, the Italian criminal justice system differentiated between full acquittals and acquittals 'for lack of evidence' (Gebbie et al., 1999; Picinali, 2022).

There has been much debate over the merits of the *not proven* verdict in the Scottish verdict system. Its proponents argue that it is a safeguard against wrongful conviction in borderline cases that do not quite reach the legal threshold of proof beyond reasonable doubt (Allan, 2017; Phalen, 2018). In such cases, jurors can opt for the *not proven* verdict, whereas faced with an identical case in a two-verdict system, they may be tempted to convict (Curley et al., 2022). Opponents of the *not proven* verdict argue that it leaves an unjust

stigma on an acquitted person (Hope et al., 2008) and allows jurors and juries to use it as a compromise verdict in difficult cases, rather than fully discussing the evidence (Chalmers et al., 2022). Debate in recent times has also focused on sexual offence cases, with a campaign to abolish the *not proven* verdict being conducted by Rape Crisis Scotland, on the basis that it causes unnecessary distress to those whose allegations of sexual assault conclude with a *not proven* verdict being returned (Chalmers et al., 2022).

For a long time, empirical studies of the effect of the *not proven* verdict were few and far between. However, several research teams have empirically investigated whether the *not proven* verdict, as a third option, has an effect on juror decision making. Smithson et al. (2007) report that the *not proven* verdict significantly reduced the frequency of *not guilty* verdicts; however, this had little effect on the frequency of *guilty* verdicts returned. Hope et al. (2008) support this finding and suggest in two separate studies that the presence of the *not proven* verdict reduced the frequency of *not guilty* verdicts but did not reduce the frequency of *guilty* verdicts (save in one experimental condition where the evidence against the accused was moderately strong). Similarly, Curley et al. (2019) report that jurors assigned to the three-verdict system returned significantly lower levels of *not guilty* verdicts than those assigned to the two-verdict system whereas convictions were not significantly different between verdict systems.

Ormston et al. (2019) conducted a large-scale multifactorial experiment, investigating in a between-subjects counter-balanced design firstly, the effect of verdict system (three vs. two), secondly, the effect of jury size (15 vs. 12) and the majority rule (simple vs. unanimous) on (pre- and post-deliberation) juror and jury verdicts in a physical assault and rape mock trial. For the purposes of the current analyses, only the effect of verdict systems on pre-deliberation juror decisions in corresponding experimental conditions are considered, because the remaining two factors only influenced post-deliberation verdicts. Ormston et al. (2019) report that the proportion of *guilty* juror verdicts was reduced for the three-verdict system compared to the two-verdict system. However, this was only statistically significant for the physical assault mock trial but not for the rape mock trial. When differentiating between the two acquittal verdicts, it was found that jurors preferred the *not proven* option to that of the *not guilty* option.

Finally, the study by Curley et al. (2022) suggests that jurors returned significantly fewer *guilty* verdicts within the three-verdict system than within the two-verdict system. Interestingly, a test of an experimental verdict system (with *proven* and *not proven* verdict) indicated that jurors returned significantly fewer convictions than the two-verdict English system, whereas the number of convictions in this *proven – not proven* verdict system was comparable to that in the current Scottish system.

It seems likely that the studies by Ormston et al. (2019) and Curley et al. (2022) give more accurate results because in these studies the stimulus material was more realistic than in the earlier studies by Curley et al. (2019), Hope et al. (2008) and Smithson et al. (2007). The studies by Ormston et al. (2019) and Curley et al. (2022) are also more ecologically valid according to the six criteria outlined by Willmott et al. (2021), with the former study meeting five of the six criteria and the latter study meeting four. They do not meet all criteria because Ormston et al. (2019) and Curley et al. (2022) did not collect participants from the electoral roll, and Curley et al. (2022) also omitted deliberations due to the Covid-19 pandemic. However, individual studies, no matter how ecologically valid they are, can only provide limited evidence, whereas a quantitative meta-analysis on conviction rates can establish a more precise and robust summary estimate of the verdict effect.

There are a number of possible reasons as to why the availability of the *not proven* verdict may influence the frequency by which

guilty verdicts are given. Firstly, the availability of the *not proven* verdict may polarise the other available options (*guilty* and *not guilty*), leading jurors to use the *not proven* verdict as a compromise (Chalmers et al., 2022; Curley et al., 2021; Hope et al., 2008). Secondly, there may be differences between legal meaning and lay semantic interpretation of the *not guilty* and *not proven* verdicts – namely, as a distinction between 'truth' and 'proof' (Jackson, 1998). Terms such as 'proven' may encourage jurors to focus on the negatives or weaknesses of the Crown's (prosecution) evidence since the burden of proof lies with them to demonstrate the guilt of the accused beyond reasonable doubt, thus leading jurors to be more likely to acquit (Curley et al., 2021, 2022; Hope et al., 2008; Jackson, 1998; McKenzie, 1985). Finally, introducing a third option may be related to the so-called 'decoy effect' in preference and consumer behaviour (Huber et al., 1982; Kahneman & Tversky, 1984). Although plausible, it remains unknown whether passing a verdict invokes similar decision heuristics as preferential choices (Hope et al., 2008).

Further, research has shown that jurors may have preferences for particular verdicts as the trial starts, which influence the verdicts they decide upon and how they evaluate evidence (Carlson & Russo, 2001). The availability of the *not proven* verdict may cause some jurors to favour this verdict option over *guilty* and *not guilty* verdicts (as shown in the research by Ormston et al., 2019), decreasing the chances of jurors favouring either the prosecution or defence evidence, thus leading to a lower frequency of *guilty* verdicts in the three-verdict system. Similar findings have been reported in more recent literature (see e.g. Curley et al., 2018, Lilley et al., 2022; Willmott et al., 2018).

Nevertheless, if the availability of the *not proven* verdict does decrease conviction rates, as suggested by some studies, then this may imply that the criminal justice system fails to provide justice for complainants. As there is a debate in the literature about the extent and significance of the effect of the *not proven* verdict on juror conviction rates, a quantitative meta-analysis on existing data sets from suitable studies seems timely and fitting.

To summarise, each of the studies highlighted above demonstrates that different verdict systems can influence how jurors reach their decisions (Curley et al., 2022). However, there is disagreement on the effect: for example, solely decrease *not guilty* verdicts (Curley et al., 2019; Hope et al., 2008; Smithson et al., 2007) or decrease both *guilty* and *not guilty* verdicts (Curley et al., 2022; Ormston et al., 2019). In other words, faced with an identical stimulus in the form of a mock trial, the propensity for jurors to convict may be affected by the number of verdicts available to them. However, these studies have different sample sizes, use different mock trials and present material in different ways. This makes it difficult to quantify and test the overall effect of verdict system in a conventional analysis, with some evidence to suggest that verdict choice is influenced by verdict systems (Ormston et al., 2019) and by crime type (Curley et al., 2022; Ling et al., 2021; Walker & Woody, 2011). Furthermore, since Smithson et al. (2007) published the first empirical paper on the *not proven* verdict, the body of research on this specific topic has grown. Still, there has been no attempt to combine findings across suitable studies in a quantitative meta-analysis to estimate the effect of verdict system. The results may inform the Scottish Government, which has announced plans to change the three-verdict system to a two-verdict system.

We can improve and clarify the effect of verdict system on juror decisions by considering not only juror verdicts from a single study or mock trial but juror verdicts from a range of studies that use the same experimental design. Despite the growing interest in meta-analyses, specifically in the context of juror decisions, most articles are either based on systematic literature reviews (e.g. Eatley et al., 2018, on the

Crime Scene Investigation (CSI) effect; Hudspith et al., 2023, on rape myths) and scoping reviews (e.g. Leverick, 2020, on rape myths). Fewer studies have collated effect size measures across studies to provide improved inferential statistics (e.g. MacCoun & Kerr, 1988, on the leniency effect; Bystranowski et al., 2021, on the anchoring effect) but so far none has targeted the effect of verdict systems on conviction rates.

Here, we use the exact numbers of convictions and acquittals as reported in studies with matched mock trials to estimate the effect of verdict system. Introducing 'verdict system' as a fixed and 'mock trials' as a random factor in a logistic regression already establishes a meta-analysis (Cooper et al., 2019; Harbord & Whiting, 2009; Simmonds & Higgins, 2016) that can increase reliability and generalisability of estimated effects even if the number of studies is small (Yarkoni, 2020).

Method

Secondary data analyses

The selection of reports and studies for this meta-analysis followed PRISMA guidelines (Page et al., 2021). Full eligibility criteria and rationale are provided in Supplementary Materials (Table S1, https://osf.io/ybvpz). In short, we included quantitative studies that compare the number of juror convictions and acquittals in matched trials under the English/Anglo-American two-verdict and the Scottish three-verdict system. All relevant studies were identified through a comprehensive search of 11 databases in February 2022 using combinations of the keywords *Juror bias** *Juror research** *Jury research** *Mock Juror trial** *Juror Simulation * Scottish verdict system** *English verdict system** *Scottish jury research** *Cognitive bias in juries* (Table S2 and S3 in Supplementary Materials).

The authors also carried out a 'snowballing' literature search to identify additional studies by searching the reference lists of publications in English. The studies and mock trials identified here are, to the best of our knowledge, the full extent of available data sets in this domain.

Inclusion/exclusion criteria

The inclusion criteria for studies and juror decisions were: (a) a random assignment of mock jurors to matched mock trials under two verdict systems, comparable to the English/Anglo-American two-verdict and the Scottish three-verdict system, (b) exact reporting of the number of mock juror decisions in terms of convictions and acquittals. Any studies or juror decisions that violated one of these criteria were disregarded from the meta-analysis.

Only reports in English were included. In the first main searches, 4,385 records were identified (see flow diagram in Figure S1, Supplementary Materials). After removal of duplicates and ineligible reports, 127 records were screened, and 103 of these could be excluded. The remaining 24 full texts were assessed, and a further 19 reports were excluded: 13 did not utilise verdict systems comparable to the English/Anglo-American two-verdict and the Scottish three-verdict system, and six did not report mock juror verdicts in terms of convictions and acquittals.

A total of 10 studies and mock trials from five reports were identified that used matched stimulus material in a between-subjects design. In each of the studies the same mock trial was presented to jurors under different verdict systems using a transcript, audio recording or full mock-trial video recording. Details of the mock trials are summarised under Materials and in Table 1 together with the number of convictions (*guilty*) and acquittals (*not guilty*; *not proven*) as reported for mock jurors randomly assigned to an English/Anglo-American two-verdict and the Scottish three-verdict system.

Materials

In each of the 10 studies, participants were randomly assigned to different verdict

Table 1. Mock trials for mock jurors.

				2-Verdict		3-Verdict	
Mock trial	Design	Presentation	Crime	Convictions	Acquittals	Convictions	Acquittals
Smithson et al., 2007	B/W	Transcript	Homicide	9	43	5	47
Smithson et al., 2007	B/W	Transcript	Death by negligence	26	26	33	19
Smithson et al., 2007	B	Audio clip	Homicide	6	18	2	22
Hope et al., 2008	B	Transcript	Sexual assault[a]	27	17	31	29
Hope et al., 2008	B	Transcript	Physical assault	24	45	16	57
Ormston et al., 2019	B	Mock trial video	Rape	95	119	75	140
Ormston et al., 2019	B	Mock trial video	Physical assault	69	147	43	175
Curley et al., 2021	B/W	Vignette	Homicide	27	37	14	50
Curley et al., 2021	B/W	Vignette	Homicide	15	49	12	52
Curley et al., 2022	B	Mock trial video	Homicide	46	32	28	51

Note: Mock jurors were assigned to the English/Anglo-American two-verdict and the Scottish three-verdict system, where each mock juror was assigned to only one verdict system in each mock trial but made decisions in two different mock trials. B = between-subjects design, B/W = mixed between/within-subjects design; convictions = guilty; acquittals = not guilty, not proven. The odds ratios of each mock trial are not shown but correspond to the estimates presented in the forest plot of Figure 2.

[a]This sexual assault trial may be better classified as a rape trial due to the description of the vignettes in Myers et al. (2003).

systems but sometimes participants served as jurors in more than a single mock trial or gave more than one verdict. After following the evidence presented in the mock trial, each randomly assigned juror convicted or acquitted the accused or defendant using the two-verdict or the three-verdict system. Subsequent decisions by the same juror in a given trial were discounted to avoid carry-over effects.

In Mock Trials 1 and 2 (Smithson et al., 2007, Study 1) a total of $N = 104$ participants read a transcribed scenario of a criminal trial (homicide) and of a civil trial (death by negligence) and gave a verdict under a two-verdict as well as a three-verdict system in a mixed

design with the order of mock trials and verdict systems counterbalanced across participants (order/verdict system – between subjects; mock trial – within subjects). We only considered the first and ignored the second verdict of each juror in a trial so that each participant gave a verdict under a single verdict system only. We treated the first verdict in each trial as independent observations although the same participant made juror decisions in each of the two mock trials.

In Mock Trial 3 (Smithson et al., 2007, Study 2) a total of $N = 72$ participants were assigned to three different groups. Two groups of 24 participants were assigned to an English/Anglo-American two-verdict system and the Scottish three-verdict system following a homicide trial. A further 24 mock jurors were assigned to a fictional verdict system and were therefore disregarded from the analyses. A third study (Smithson et al., 2007, Study 3) with a total $N = 96$ participants was disregarded for the same reason.

In Mock Trial 4 (Hope et al., 2008, Study 1) a total of $N = 104$ participants were assigned to the two-verdict system and three-verdict system in a between-subjects design passing their verdicts in a sexual assault trial. The description of the trial (Myers et al., 2003) suggests that according to the Sexual Offences Act (Scotland) 2009 the criminal case would be classified as a rape rather than sexual assault.[1]

In Mock Trial 5 (Hope et al., 2008, Study 2) a total of $N = 142$ participants were assigned to a two-verdict and three-verdict system in a physical assault trial. The juror verdicts in Study 2 were accumulated across three trial versions that featured weak, moderate and strong evidence.

In Mock Trial 6 (Ormston et al., 2019) a total of $N = 429$ and in Mock Trial 7 a total of $N = 434$ participants watched a video of proceedings in a rape and a physical assault trial, respectively. After watching the video they provided individual juror verdicts before (and after) deliberating in a jury. We only used juror verdicts before deliberation and disregarded 106 participants from both mock trials because they were not clearly assigned to one of the experimental conditions in a counterbalanced design. We treated juror decisions before deliberation as independent of jury size (12, 15) and voting rule (majority, unanimity) because both factors had no effect on these juror verdicts.

In Mock Trials 8 and 9 (Curley et al., 2019, 2021) two transcripts of different court trials (Vignettes 1 and 2) were presented to the same $N = 128$ participants in a mixed design. The order of vignettes and verdict systems was counterbalanced across participants. In the analysis of Curley et al. (2019) the juror verdicts were treated as independent observations by only using the verdicts from the first vignette presented to each participant. In an alternative analysis, Curley et al. (2021) used the verdicts from both vignettes.

In Mock Trial 10 (Curley et al., 2022) a total of $N = 227$ participants watched a video recording of a re-enacted court case on physical assault. In a between-subjects design three groups of participants decided as mock jurors under different verdict systems (English, Scottish, Experimental). We disregarded the verdicts of the 70 jurors who were assigned to the experimental verdict system that featured only *proven* and *not proven* as verdict options.

Results

The data and analysis code are available at https://osf.io/ybvpz. Data were analysed using R, Version 4.0.0 (R Core Team, 2020), R-package *metafor* Version 2.4–0 (Viechtbauer,

[1]This is because the Sexual Offences (Scotland) Act 2009 defines rape as 'penetration of the vagina, anus or mouth of another person by the penis without consent'. The trial clearly deals with penetration, and as all rape trials are a type of sexual assault trial but not all sexual assault trials are a type of rape trial, rape is a more legally accurate trial type to explain the vignettes used in the Myers et al (2003) and Hope et al (2008) studies.

Table 2. Number and percentage of convictions and acquittals for the Scottish three-verdict and Anglo-American two-verdict system.

	Conviction		Acquittal		Total	
Juror decision verdict system	*N*	%	*N*	%	*N*	%
3-Verdict	259	15	642	36	901	51
2-Verdict	344	19	533	30	877	49
Total	603	34	1175	66	1778	100

Note: Convictions = guilty; acquittals = not guilty, not proven. Juror verdicts are pooled across mock trials.

2010). Since this is a secondary data analysis of previously published data no ethical approval or pre-registration is required.

We were interested in whether verdict system (Scottish, Anglo-American) in a between-participants design significantly affects conviction rates when combining data of 10 mock trials and different studies. Table 2 features the total number of convictions (*guilty*) and acquittals (*not guilty*; *not proven*) under the Scottish three-verdict and English two-verdict systems of all mock jurors pooled across trials.

A Fisher's exact test as well as Pearson's chi-squared test with Yates' continuity correction (Prescott, 2019) indicates that the odds ratio (*OR*) significantly deviates from 1 ($p < .00001$), suggesting that verdict system (three verdicts, two verdicts) and juror decisions (conviction, acquittal) are not independent. The *OR* estimated by maximum likelihood equals 0.625, 95% confidence interval, CI [0.51, 0.77]. This estimate uses pooled data and suggests that the odds of being convicted by jurors under the Scottish three-verdict system are 37.5% lower than those under the Anglo-American two-verdict system.

This odds ratio can be computed from the entries in Table 1. The odds for a conviction under the Scottish three-verdict system P(Conviction|Three Verdicts)/(1 − P(Conviction|Three Verdicts)) = (259:901)/(642:901) = 0.403 are divided by the odds for a conviction under the English verdict system P(Conviction|Two Verdicts)/(1 − P(Conviction|Two Verdicts)) = (344:877)/(533:877) = 0.645, resulting in an odds ratio $OR = 0.403/0.645 = 0.625$.

An alternative measure, based on the ratio of probabilities rather than odds, is the relative risk or risk ratio (*RR*). The numbers in Table 1 indicate a conviction rate of 344:877 or 39.2% under the two-verdict system and a lower conviction rate of 259:901 or 28.7% under the three-verdict system. The risk ratio between the verdict systems is $RR = 0.287/0.392 = 0.732$. According to this measure, jurors under the Scottish three-verdict system are 26.8% less likely to convict than jurors under the Anglo-American two-verdict system. An even more intuitive measure provides the risk difference $RD = 0.392 - 0.287 = -0.105$. This indicates a 10.5% reduction of the likelihood to convict when comparing the Scottish to the Anglo-American verdict system.

However, these results are based on pooled data and do not take into account different odds ratios and sample sizes of mock trials. We therefore conducted a logistic regression with random effects using the observed number of convictions and acquittals from each mock trial and verdict system. This method works well for small and stratified samples and leads to more precise estimates and statistical inference (Cooper et al., 2019). More conventional meta-analyses rely on effect sizes in less controlled studies with more complex designs and therefore require a larger number of studies (e.g. Milkman et al., 2021).

Logistic regression with random effects

For the main analysis each individual verdict of a mock juror was categorised as a

conviction or an acquittal, and the binary data served as the dependent variable. Studies with mock trials (labelled 1–10) are treated as a random effect. Introducing mock trials as a random effect is appropriate because they vary unsystematically across studies.

This approach makes weak assumptions about the underlying distributions and variance estimation but should give more robust and accurate results because it can accommodate extreme values, as well as varying and unbalanced sample sizes (Cooper et al., 2019; Harbord & Whiting, 2009; Simmonds & Higgins, 2016). In the following, we report the results of a logistic regression model with random effects, using restricted maximum likelihood estimation (REML), that is specifically adapted for meta-analyses and is implemented in the function rma ( ) of the R-package *metafor* (Viechtbauer, 2010).

By taking into account the different number of convictions and acquittals for each verdict system across the 10 mock trials, the estimated log odds for verdict system is −0.52, 95% CI [−0.77, −0.27]. This value is statistically significant ($z = -4.10$, $SE = 0.13$, $p < .0001$), and the corresponding odds ratio is $OR = 0.59$, 95% CI [0.46, 0.76]. This means that for all mock trials the odds of a conviction are reduced under the Scottish three-verdict system by a factor of 0.593, amounting to a change in odds by 40.7% compared to the Anglo-American two-verdict system. Note that the *OR* estimate is slightly lower and has a narrower confidence interval than that from Fisher's exact test reported above ($OR = 0.63$, 95% CI [0.51, 0.77]).

In the funnel plot of Figure 1 each data point represents a mock trial with the standard error plotted against the odds ratio. The symmetric appearance of data points with a single outlier (Mock Trial 2) suggests no selection bias, and Kendall's rank test for funnel plot asymmetry is not significant ($\tau = .02$, $p = 1.0$). A test of heterogeneity is also not significant, Cochran's $Q(9) = 12.28$, $p = .198$, $I^2 = 22.2\%$. The single outlier to the right of the funnel denotes the only 'civil trial' (Smithson et al., 2007, Study 2). In this mock trial a city council rather than a person was sued for death by negligence, which may explain the increased odds ratio. Excluding the data of Trial 2 had only a small effect on the results ($OR = 0.55$; 95% CI [0.44, 0.68]). The data point for Mock Trial 3 (Smithson et al., 2007, Study 2) at the bottom of the plot has the lowest sample size ($N = 48$) and the largest standard error among the 10 mock trials.

In Figure 2 the results are summarised in a forest plot. The size of each black square corresponds to the sample size in the corresponding mock trial. The horizontal position of the squares indicates the estimated odds ratio (on a log scale), and the horizontal whiskers describe the 95% confidence intervals. Three out of 10 mock trials (Trials 7, 8 and 10) suggest odds ratios that are significantly lower than 1.0, whereas the estimates for the other mock trials tend to be more uncertain and therefore ambiguous. Only Trial 2 suggests an odds ratio higher than 1.0. The polygon or diamond at the bottom of the plot describes the summary estimate. The centre of the polygon corresponds to the point estimate of $OR = 0.59$, and the left and right edges indicate the 95% CI [0.46, 0.76]. Thus, the summary estimate is significantly lower than 1.0.

In an ancillary analysis each individual verdict of a mock juror was categorised into *not guilty* and *other*, with the latter category consisting of *guilty* and *not proven* verdicts for the three-verdict system. The resulting binary variable informs about changes in the number of *not guilty* verdicts when *not proven* is available or not. As before, verdict system (three verdicts, two verdicts) is entered as a fixed effect whereas trial (1 to 10) is treated as a random effect.

By taking into account the number of *not guilty* and *other* verdicts for each verdict system across the 10 mock trials, the estimated log odds for verdict system is −2.13, 95% CI [−2.82, −1.45]. This estimate are also statistically significant ($z = -6.14$, $SE = 0.35$,

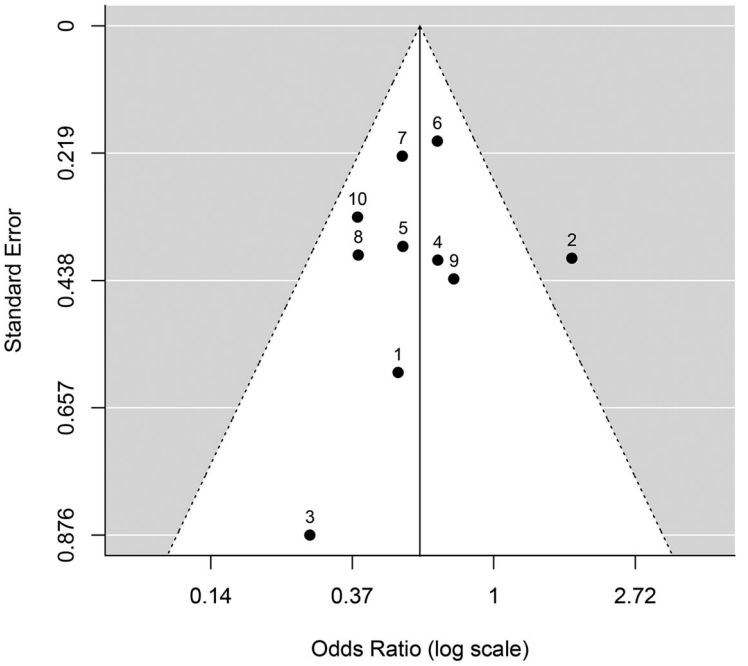


Figure 1. Funnel plot of odds ratios (log scale) centred on the odds ratio estimate 0.59 (vertical line) with standard errors (*SE*) of estimates on the *y*-axis. The number above each black data point refers to the corresponding mock trial.

$p < .0001$). The corresponding odds ratio of $OR = 0.12$, 95% CI [0.06, 0.23], suggests a reduction of the odds for a *not guilty* verdict by a factor of 0.12 or 88% when comparing the three-verdict system to the two-verdict system.

Discussion

In Scotland, jurors in criminal trials have three verdicts available to them (*guilty*, *not guilty* and *not proven*), whereas in England and Wales, and other common law systems such as the United States of America and Australia, there are only two verdicts (*guilty* and *not guilty*). The issue of whether juror verdicts might be affected by the number of verdict options available to them has been of some interest recently, in terms of both policy making and academic research. It is next to impossible to establish the true impact that different verdict systems might have on conviction rates in Scotland and England, for example, by simply comparing the reported conviction rates of trials in the two countries. The reason for this is that in addition to the different verdict systems, the Scottish and English legal systems have many other dissimilarities, in terms of jury size, majority rules for the jury and the way in which criminal offences are defined and admitted to court. For example, Scotland has a rule requiring corroboration of evidence, meaning that cases cannot proceed unless there are two sources of evidence of every 'crucial fact' (the identity of the perpetrator and the key offence elements). England does not have such a rule. This is likely to affect the number and type of cases that make it to trial in the first place, meaning that simply comparing actual conviction rates between countries would be misleading. For the same reason it would be problematic to include studies in the present meta-analysis that do not randomly assign jurors to different verdict systems and therefore have non-matching mock trials.

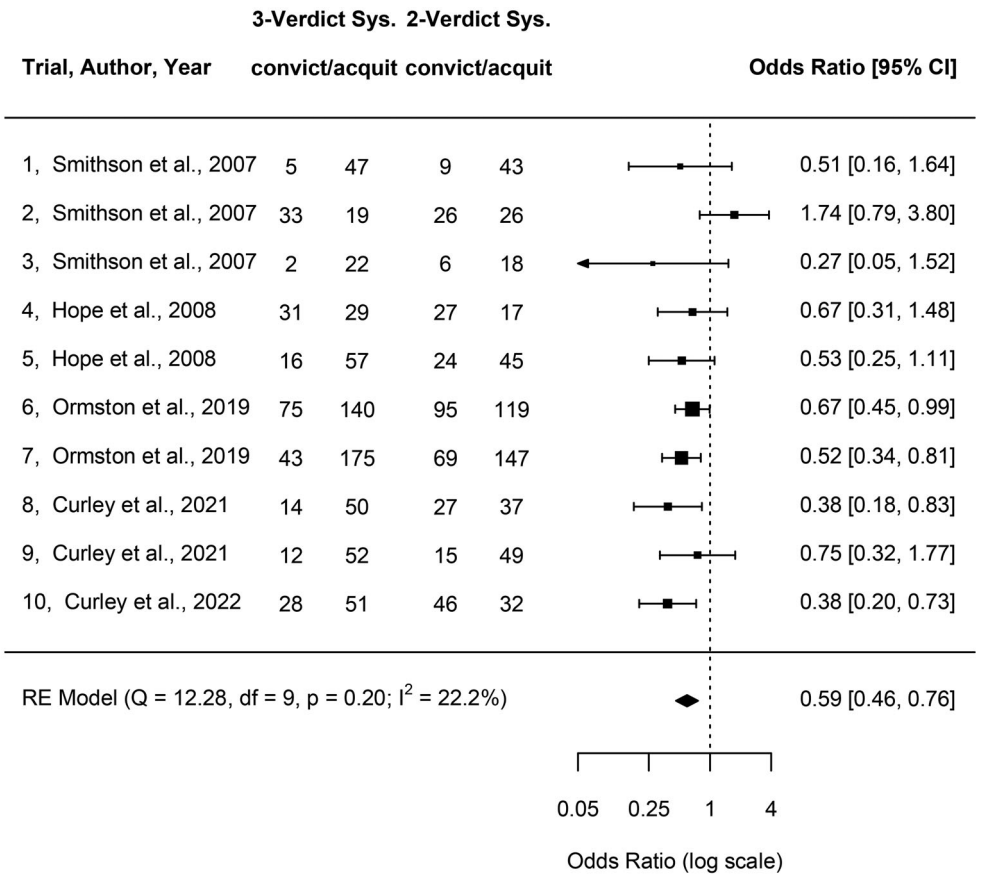


Figure 2. Forest plot of estimated odds ratios for 10 studies and mock trials based on a logistic model with random effects (Viechtbauer, 2010). CI = confidence interval. An odds ratio below 1.0 indicates a lower number of convictions for the 3-verdict system than for the 2-verdict system (see text for details).

In order to investigate the effect of the three-verdict system compared to the two-verdict system it is therefore necessary to employ controlled experiments in which all features of the trial are held constant save for the verdict system. Such experiments typically use mock trials as a stimulus for jurors to reach a verdict, and a sufficiently large number of these experimental studies are now available to evaluate the results across studies and mock trials. In the present paper, we performed a quantitative meta-analysis on several experimental studies and found that despite considerable variability across studies and mock trials in terms of sample size, stimulus material, type of crime and conviction rates the results are quite unambiguous: there is a statistically significant effect towards lower conviction rates under the Scottish three-verdict system than under an Anglo-American two-verdict system. We estimate that under the Scottish compared to the Anglo-American verdict system the odds of a conviction are reduced by a factor of 0.593 (40.7%).

An equivalent analysis on *risks* or probabilities suggests a factor of 0.73, with the risk for a conviction reduced by 27% under the Scottish verdict system. The risk difference (*RD*) for convictions between the two verdict systems approaches $RD = 11\%$.

This summary estimate appears to be the first of its kind in the domain of juror decisions

and establishes a highly significant effect of verdict system on the odds of convictions or *guilty* verdicts. This is accompanied by significantly lower odds for *not guilty* verdicts than for *guilty* and *not proven* verdicts, which are reduced by a factor of 0.12 (88%) when comparing the Scottish three-verdict system with the Anglo-American two-verdict system.

Taken together, the results demonstrate a robust and significant effect of verdict system on juror decisions across mock trials, ranging from death by negligence to physical assault, rape and homicide. Introducing mock trials as a random factor in a logistic regression improves estimation because it helps to explain variability in the data (Cooper et al., 2019). This increases robustness and generalisability of overall estimates and statistical power (Yarkoni, 2020).

There are a number of possible psychological explanations for the effect of verdict system. One general explanation comes from the field of bounded rationality, which states that decision making of an individual is shaped by both their cognition and the decision environment (Simon, 1956). Bounded rationality is a concept that suggests that decision-making capabilities of individuals are limited by the information they can collect, their cognitive limitations of information processing and the finite amount of time they have to reach a decision. It suggests that people make decisions based on 'satisficing' rather than strict maximising principles.

Applied to decision making by jurors, bounded rationality suggests that the decision-making process has internal constraints (e.g. processing limits, both in terms of information storage and processing speed), as well as external constraints (e.g. how complex the external information is and/or how costly it is to search for information). Additionally, time constraints may prevent jurors from fully considering and integrating all the evidence and arguments before rendering a decision.

As a result of these constraints, jurors may employ heuristics or mental shortcuts to make decisions. For example, in extreme cases jurors may rely on stereotypes, emotions or prior beliefs to form a verdict. These factors can influence the decision-making process and may lead to decisions that are not considered fully rational (Curley et al., 2021). Similarly, the process during jury deliberation may also be constrained by the fact that jurors often try to reach a consensus where some jurors may be more persuasive and influential than others (Clark et al., 2007). Group dynamics and social pressure may also affect the final jury verdict. Furthermore, as the verdict system is changed from a two-verdict system to a three-verdict system, the decision environment changes, potentially introducing more complexity and ambiguity to the decision process and therefore increasing the use of heuristics, which ultimately affects conviction rates (Curley et al., 2022).

Another reason for why the availability of the *not proven* verdict relates to a lower conviction rate may be due to attention being on 'proof' rather than 'truth'. In the three-verdict system, there are two acquittal verdicts, one with a focus on proof, which may raise scepticism of jurors towards the prosecution's evidence, thus tilting the scales of justice in favour of the defence. The removal of the *not proven* verdict may also stop jurors using it as a convenient compromise. Due to the verdict falling in the middle of *guilty* and *not guilty*, its absence forces jurors to 'pick a side', with some jurors choosing *not guilty* and some choosing *guilty*, causing an increase in both when compared to a three-verdict system (Hope et al., 2008; Smithson et al., 2007). Although the odds are larger for *not guilty* versus *other* than for *guilty* versus *other* (acquittals), it still runs contrary to the legal perspective that if jurors fail to reach a decision beyond the threshold of reasonable doubt in a three-verdict system, then they should choose a *not guilty* verdict in a two-verdict system.

In their seminal paper, Carlson and Russo (2001) showed that jurors tend to settle on

verdicts early on in a trial, which causes them to favour evidence that supports their initial verdict and distort or disregard evidence that does not confirm their initial verdict. The availability of the *not proven* verdict implies that jurors are more likely to favour the 'middling verdict' from early on in a trial. This means that they are less likely to be biased to the arguments of either the prosecution or the defence team, decreasing their chances of reaching a *guilty* and *not guilty* verdict and thus leading to a lower number of convictions in a three-verdict system.

The statistically significant effect of verdict system on conviction rates indicates an obvious decision bias in jurors. Although trials and legal consequences remain the same, introducing a second label for acquittal reduces conviction rates of individual jurors. Therefore, juror decisions must be exposed to some form of cognitive bias. More specifically, under both verdict systems jurors are asked to assign different labels to what is essentially a binary choice: to convict or acquit. Under an Anglo-American verdict system *guilty* and *not guilty* verdicts map directly onto conviction and acquittal. Under the Scottish verdict system *guilty* also maps onto conviction whereas *not guilty* and *not proven* both map onto acquittal. The third option *not proven* simply offers an alternative label to acquit the defendant/accused but a rational decision maker should not be affected by this third option (Chalmers et al., 2022; Hope et al., 2008).

In preference tasks and consumer behaviour (Huber et al., 1982; Kahneman & Tversky, 1984) introducing a third option or 'decoy' can drive a decision maker who is indifferent about two options toward one of the options. A successful decoy facilitates 'asymmetric dominance', which requires attributes on two dimensions with indifferent information between two of the options but a clear dominance of one option over the decoy. In the context of juror decisions, the two dimensions may be described as the 'presumption of innocence' of the defendant and 'proof beyond reasonable doubt' against the defendant, not unlike the two concepts of 'truth' and 'proof'. If a juror is indifferent between a *guilty* and *not guilty* verdict, then the introduction of *not proven* as a third verdict option may shift jurors' decision towards an acquittal because *not proven* reminds jurors that without evidence beyond reasonable doubt the defendant should be acquitted even if they are believed to be guilty.

The aspiration that a verdict in a court case is objective and rational or unbiased and unprejudiced appeals to everyone whereas the presence of bias or prejudice that may affect jurors' decisions seems wrong and harmful (Sherrod, 2019). However, the presumption of 'innocence until proven guilty' and providing 'proof beyond reasonable doubt' may trigger different predispositions in jurors' decision making. Therefore, attempting to eradicate cognitive bias is quite challenging and may conflict with instructions given to jurors. Verification of a verdict is also extremely difficult to achieve because this requires not only identification and validation of previously unknown facts in a court case (e.g. guilt or innocence of the accused, validity of existing evidence and witnesses, new evidence), but also application of the same normative rules and standards (e.g. criminal proceedings, juror selection, interpreting laws and societal norms), and even prediction of future events (e.g. consequences of a verdict on victim(s), the accused, and society in general).

Limitations

There are a number of limitations in the present analyses that need to be acknowledged. An exhaustive search identified a total of only 10 studies and mock trials with varying numbers of jurors and odds ratios for conviction as well as ecological validity.

Ecological validity refers to the extent to which a mock trial reflects the features of a real criminal trial. For example, does the mock trial use: (a) accurate legal directions, (b)

realistic evidence presented by trained actors and (c) a representative (e.g. community) sample of mock jurors rather than relying on student participants (Bornstein, 1999)? The higher the ecological validity of the experiment, the more likely it is that the findings can be replicated in real criminal trials. Some of the studies that are included here had relatively low ecological validity and, for example, used trial transcripts or audio vignettes. This is not surprising because creating a realistic audio-visual stimulus of proceedings in court is expensive and time consuming. Nevertheless, the largest data sets included in the meta-analysis – that of Ormston et al. (2019) and of Curley et al. (2022) – had the highest ecological validity. Ormston et al. used hour-long trial videos, scripted in collaboration with legal professionals and performed by actors in a real courtroom. Curley et al. used a 53-minute-long video featuring legal professionals and a real judge, also in a real courtroom, with actors playing the roles of the witnesses. In both studies jurors were given the same legal directions as they would be given in a real trial and were drawn from the local population to reflect the demographic make-up of real juries.

Another possible limitation would be the presence of publication bias (van Aert et al., 2019). The meta-analysis is based on the number of convictions and acquittals from 1778 jurors from 10 studies published in journal articles between 2007 and 2022. Publication bias may have prevented the publication of null effects or contradicting results. However, we are not aware – through searches of theses, conference abstracts and personal communications with researchers in the field – of unpublished studies in this domain, and the funnel plot in Figure 1 does not suggest an asymmetry.

Another limitation of the present analysis may be the degree to which each of the studies varied in relation to crime type (death by negligence, rape, homicide and physical assault). Research has shown that verdicts can vary for different crime types (Ormston et al., 2019). Increasing the number of mock trials would have enabled us to introduce a moderator variable such as crime type. However, we only identified 10 studies in the literature that systematically manipulated verdict system for matching mock trials. Despite this limitation, the present meta-analysis informs the Scottish Government about the consequences of abolishing the *not proven* verdict. Combining the findings from a small but controlled body of studies on the *not proven* verdict makes it possible to establish a robust statistically significant effect of verdict system on juror decisions.

Finally, the strong effect of verdict system on juror decisions does not necessarily translate into an equivalent effect on jury verdicts. In real criminal trials, verdicts are delivered by juries after deliberation, usually in a group of 12 jurors (or, in Scotland, a group of 15 jurors). In most jurisdictions, verdicts have to be (near-)unanimous, although in Scotland juries can deliver majority verdicts (Chalmers et al., 2020). This means that individual juror verdicts may change over the course of the deliberations in order to facilitate agreement, questioning whether jurors' pre-deliberation verdicts reliably predict jury verdicts. In Ormston et al. (2019) jury deliberations were included in the research design, and jury verdicts were recorded for a total of 64 juries in a counterbalanced design (2 verdict systems, 2 trial types, 2 jury sizes and 2 majority rules). In only one instance (simple majority required, three-verdict system, 15-person jury) mock jurors were significantly more likely to give a *guilty* verdict post-deliberation than pre-deliberation. This suggests that deliberation may have a leniency effect on juror decisions. The exclusion of deliberations in jury research may decrease the ecological validity of the studies in the current analysis. However, only one jury study that involved jurors deliberating collectively in a jury was available, which is a more general issue (Curley & Peddie, in press; Diamond, 1997). Future research into the *not proven* verdict, and other relevant factors,

should include jury deliberations and jury verdicts.

Implications

In 2023, the Scottish Government introduced legislation to abolish the *not proven* verdict, which, at the time of writing, was being consulted on. If this legislation gains parliamentary approval and becomes law, this would bring Scotland in line with the vast majority of legal systems worldwide in having only two verdicts that can be returned after a criminal trial – *guilty* and *not guilty*. The results of the present meta-analysis suggest that this may lead to a significant increase in convictions. This is, of course, on the assumption that the results of the experimental studies can be translated into the context of real-world trials. As we noted above, this is not an assumption that can be made, given the differences between the experimental studies and real criminal cases and trials (Curley & Peddie, in press; Diamond, 1997). Nevertheless, what can be concluded with a reasonable degree of certainty is that if the abolition of the *not proven* verdict does have an effect on conviction rates, and without any further changes being made to the legal system (juror size, majority rule), it will be in the direction of an increased number of convictions across most trial types.

This is a change that would be welcomed by groups who argue that justice is not being secured for those who have experienced sexual assault. In recent years, much of the criticism of the *not proven* verdict has come from Rape Crisis Scotland, who have advocated for its abolition (Rape Crisis Scotland, 2021). Rape and sexual offence cases have a far lower conviction rate than trials for other serious offences, and the *not proven* verdict is used disproportionately in such cases (Chalmers et al., 2021a). Research has suggested that jurors sometimes use the *not proven* verdict as a compromise verdict to bring deliberations to a close rather than resolve disagreements or uncertainty (Chalmers et al., 2022). This is especially concerning in rape cases and sexual offence cases, where it is well documented that jurors exhibit other biases such a reliance on rape myths (Chalmers et al., 2021b; Leverick, 2020).

There is also the danger that an increase in conviction rates might be an increase in wrongful convictions – the conviction of the factually innocent. This seems unlikely to occur in the arena of sexual offences, where conviction rates are low (see Chalmers et al., 2021b), but could be a genuine problem in other contexts – for example, when the evidence is primarily based on eyewitness identification or a confession (Chalmers et al., 2022). The *not proven* verdict has traditionally been seen as one of the major protections against wrongful conviction in Scotland, and removing it without putting other safeguards in place may well increase the risk of wrongful convictions.

All of these points are necessarily speculative. It is near impossible to identify the 'correct' rate of convictions that the criminal justice system should be returning. Criminal trials are contested versions of reality, and even the parties directly involved might not recall or know everything that happened in the context of an alleged crime. As changes are made, however, policy makers should be mindful that other protections may need to be implemented in order to reduce the risk of wrongful convictions. These may include changes to the majority rule for conviction (whereby a *guilty* verdict can be returned if only 8 of the 15 jurors favour *guilty*) or to the rules relating to particular types of evidence, such as confessions or eyewitness identification evidence.

Conclusion

The highly significant effect of verdict system on conviction rates does not simply confirm earlier findings from single studies but establishes for the first time a reliable estimate and statistical test of the verdict effect across mock trials. Thus, the main result from the present

meta-analysis is that, across studies and trial types, the Scottish three-verdict system reduces the odds of a conviction by 40.7% and the probability of a conviction by 10.5% compared to an Anglo-American two-verdict system.

The more fundamental question of whether abolishing the *not proven* verdict and the likely increase in convictions may benefit or harm defendants, victims and the wider society is far more difficult to address because this requires validation of verdicts in criminal legal trials as 'correct' or 'true'. Such a validation is very difficult, if not impossible, because it relies on gathering new information, sometimes long after a first trial (e.g. through technological advances, confessions). Nevertheless, only such a validation would allow determining whether a change in conviction rates constitutes a successful adaptation of the legal system or not.

Ethical standards

Declaration of conflicts of interest

Elaine Jackson has declared no conflicts of interest.
Lee Curley has declared no conflicts of interest.
Fiona Leverick has declared no conflicts of interest.
Martin Lages has declared no conflicts of interest.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Supplemental data

Supplementary data and analysis code can be accessed at https://osf.io/ybvpz

Funding

This work was supported by the EU under Erasmus+ Grant QHELP 2019-1-EE01-KA203-051708.

ORCID

Lee Curley http://orcid.org/0000-0001-5829-4740
Fiona Leverick http://orcid.org/0000-0001-5245-9874
Martin Lages http://orcid.org/0000-0002-4962-3116

References

Allan, J. (2017). A democratic legal positivist's defence of Scotland's bastard verdict. *Australian Journal of Legal Philosophy*, *42*(1), 214–227.

Barbato, J. M. (2005). Scotland's bastard verdict: Intermediacy and the unique three-verdict system. *Indiana International & Comparative Law Review*, *15*(3), 543–582. https://doi.org/10.18060/17848

Bornstein, B. H. (1999). The ecological validity of jury simulations: Is the jury still out? *Law and Human Behavior*, *23*(1), 75–91. https://doi.org/10.1023/A:1022326807441

Bray, S. L. (2005). Not proven: Introducing a third verdict. *University of Chicago Law Review*, *72*(4), 1299–1330.

Bystranowski, P., Janik, B., Próchnicki, M., & Skórska, P. (2021). Anchoring effect in legal decision-making. A meta-analysis. *Law and Human Behavior*, *45*(1), 1–23. https://doi.org/10.1037/lhb0000438

Carlson, K. A., & Russo, J. E. (2001). Biased interpretation of evidence by mock jurors. *Journal of Experimental Psychology: Applied*, *7*(2), 91–103. https://doi.org/10.1037/1076-898X.7.2.91

Chalmers, J., Leverick, F., & Munro, V. E. (2021a). A modern history of the not proven verdict. *Edinburgh Law Review*, *25*(2), 151–172. https://doi.org/10.3366/elr.2021.0692

Chalmers, J., Leverick, F., & Munro, V. E. (2021b). The provenance of what is proven: exploring (mock) jury deliberation in Scottish rape trials. *Journal of Law and Society*, *48*(2), 226–249. https://doi.org/10.1111/jols.12287

Chalmers, J., Leverick, F., & Munro, V. E. (2022). Beyond doubt: The case against not

proven. *The Modern Law Review*, *85*(4), 847–878. https://doi.org/10.1111/1468-2230.12707

Chalmers, J., Leverick, F., Munro, V. E., Murray, L., & Ormston, R. (2020). Three distinctive features, but what is the difference? Key findings from the Scottish Jury Project. *Criminal Law Review*, 983–1003.

Clark, J., Boccaccini, M. T., Caillouet, B., & Chaplin, W. F. (2007). Five factor model personality traits, jury selection, and case outcomes in criminal and civil cases. *Criminal Justice and Behavior*, *34*(5), 641–660. https://doi.org/10.1177/0093854806297555

Cooper, H., Hedges, L. V., & Valentine, J.V. (Eds.). (2019). *The handbook of research synthesis and meta-analysis* (3rd ed.). Russell Sage Foundation.

Curley, L. J. & Peddie, N. (in press). The strengths and limitations of the experimental study of juror and jury decision making. In N. Monaghan (Ed.), *Challenges in the jury system: UK Juries in comparative perspective*. Routledge.

Curley, L. J., MacLean, R., Murray, J., Laybourn, P., & Brown, D. (2019). The bastard verdict and its influence on jurors. *Medicine, Science, and the Law*, *59*(1), 26–35. https://doi.org/10.1177/0025802418811740

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*, *40*(3), 452–466. https://doi.org/10.1002/bsl.2568

Curley, L. J., Murray, J., MacLean, R., Laybourn, P., & Brown, D. (2018). Faith in thy threshold. *Medicine, Science, and the Law*, *58*(4), 239–250. https://doi.org/10.1177/0025802418791062

Curley, L., Munro, J., Frumkin, L., & Turner, J. (2021). Informing reform: The views of legal professionals on the unique aspects of Scottish Law. *Medicine, Science, and the Law*, *61*(4), 256–265. https://doi.org/10.1177/0025802421992913

Curley, L., Murray, J., MacLean, R., Munro, J., Lages, M., Frumkin, L., Laybourn, P., & Brown, D. (2022). Verdict spotting: Investigating the effects of juror bias, evidence anchors, and verdict system in jurors. *Psychiatry, Psychology, and Law : An Interdisciplinary Journal of the Australian and New Zealand Association of Psychiatry, Psychology and Law*, *29*(3), 323–344. 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. https://doi.org/10.1023/A:1024831908377

Eatley, G., Hueston, H. H., & Price, K. (2018). A meta-analysis of the CSI effect: The impact of popular media on jurors' perception of forensic evidence. *Politics, Bureaucracy & Justice*, *5*(2), 1–10.

Gebbie, G. C., Jebens, S. E., & Mura, A. (1999). Not proven as a juridical fact in Scotland, Norway and Italy. *European Journal of Crime, Criminal Law and Criminal Justice*, *7*(3), 262–276. https://doi.org/10.1163/15718179920518853

Harbord, R. M., & Whiting, P. (2009). Metandi: Meta-analysis of diagnostic accuracy using hierarchical logistic regression. *The Stata Journal: Promoting Communications on Statistics and Stata*, *9*(2), 211–229. https://doi.org/10.1177/1536867X0900900203

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

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, *9*(1), 90–98. https://doi.org/10.1086/208899

Hudspith, L. F., Wager, N., Willmott, D., & Gallagher, B. (2023). Forty years of rape myth acceptance interventions: a systematic review of what works in naturalistic institutional settings and how this can be applied to educational guidance for jurors. *Trauma, Violence & Abuse*, *24*(2), 981–1000. https://doi.org/10.1177/15248380211050575

Jackson, B. S. (1998). Truth or proof?: The criminal verdict. *International Journal for the Semiotics of Law*, *11*(3), 227–273. https://doi.org/10.1007/BF01110409

Judicial Institute for Scotland. (2022). *Jury manual*. https://judiciary.scot/home/media-information/publications/judicial-institute-publications

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, *39*(4), 341–350. https://doi.org/10.1037/0003-066X.39.4.341

Leverick, F. (2020). What do we know about rape myths and juror decision making? *The*

International Journal of Evidence & Proof, *24*(3), 255–279. https://doi.org/10.1177/1365712720923157

Lilley, C., Willmott, D., & Mojtahedi, D. (2022). Juror characteristics on trial: Investigating how psychopathic traits, rape attitudes, victimization experiences, and juror demographics influence decision-making in an intimate partner rape trial. *Frontiers in Psychiatry*, *13*, 1086026. https://doi.org/10.3389/fpsyt.2022.1086026

Ling, S., Kaplan, J., & Berryessa, C. M. (2021). The importance of forensic evidence for decisions on criminal guilt. *Science & Justice : Journal of the Forensic Science Society*, *61*(2), 142–149. https://doi.org/10.1016/j.scijus.2020.11.004

MacCoun, R. J., & Kerr, N. L. (1988). Asymmetric influence in mock jury deliberation: Jurors' bias for leniency. *Journal of Personality and Social Psychology*, *54*(1), 21–33. https://doi.org/10.1037//0022-3514.54.1.21

McKenzie, J. P. (1985). The Editorial Notebook; Between guilt and innocence. *The New York Times*, A.

Milkman, K. L., Gromet, D., Ho, H., Kay, J. S., Lee, T. W., Pandiloski, P., Park, Y., Rai, A., Bazerman, M., Beshears, J., Bonacorsi, L., Camerer, C., Chang, E., Chapman, G., Cialdini, R., Dai, H., Eskreis-Winkler, L., Fishbach, A., Gross, J. J., … Duckworth, A. L. (2021). Megastudies improve the impact of applied behavioural science. *Nature*, *600*(7889), 478–483. https://doi.org/10.1038/s41586-021-04128-4

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Altman, D., Antes, G., Atkins, D., Barbour, V., Barrowman, N., & Berlin, J. A. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Myers, B., Rosol, A., & Boelter, E. (2003). Polygraph evidence and juror judgments: The effects of corroborating evidence. *Journal of Applied Social Psychology*, *33*(5), 948–962. https://doi.org/10.1111/j.1559-1816.2003.tb01933.x

Ormston, R., Chalmers, J., Leverick, F., Munro, V. E., & Murray, L. (2019). *Scottish jury research: Findings from a large-scale mock jury study.* Scottish Government.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ (Clinical Research ed.)*, *372*(71), n71. https://doi.org/10.1136/bmj.n71

Phalen, H. (2018). Overcoming the opposition to a third verdict: A call for future research on alternative acquittals. *Arizona State Law Journal*, *50*(1), 401–430.

Picinali, F. (2022). *Justice in-between.* Oxford University Press.

Prescott, R. J. (2019). Avoid being tripped up by statistics: Statistical guidance for a successful research paper. *Gait & Posture*, *72*, 240–249. https://doi.org/10.1016/j.gaitpost.2018.06.172

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Rape Crisis Scotland. (2021, April 8). *End not proven*. Blog post. https://www.rapecrisisscotland.org.uk/news/blog/end-not-proven/

Scottish Government. (2022). *A stronger and more resilient Scotland: The programme for government 2022 to 2023.* Scottish Government.

Sherrod, D. R. (2019). *The jury crisis: What's wrong with jury trials and how we can save them*. Rowman & Littlefield.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*(2), 129–138. https://doi.org/10.1037/h0042769

Simmonds, M. C., & Higgins, J. P. T. (2016). A general framework for the use of logistic regression models in meta-analysis. *Statistical Methods in Medical Research*, *25*(6), 2858–2877. https://doi.org/10.1177/0962280214534409

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

Stijnen, T., Hamza, T. H., & Ozdemir, P. (2010). Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. *Statistics in Medicine*, *29*(29), 3046–3067. https://doi.org/10.1002/sim.4040

The Sexual Offences Act Scotland. (2009). *Published online in The National Archives*.

https://www.legislation.gov.uk/asp/2009/9/body

Vaki, Y., & Rabin, Y. (2021). Two kinds of acquittals – different kinds of doubts. *Criminal Law Forum*, *32*(1), 97–123. https://doi.org/10.1007/s10609-020-09407-9

van Aert, R. C. M., Wicherts, J. M., & van Assen, M. A. L. M. (2019). Publication bias examined in meta-analyses from psychology and medicine: A meta-meta-analysis. *PLoS One*, *14*(4), e0215052. https://doi.org/10.1371/journal.pone.0215052

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://www.jstatsoft.org/v036/i03. https://doi.org/10.18637/jss.v036.i03

Walker, C. M., & Woody, W. D. (2011). Juror decision making for juveniles tried as adults: The effects of defendant age, crime type, and crime outcome. *Psychology, Crime & Law*, *17*(8), 659–675. https://doi.org/10.1080/10683160903493471

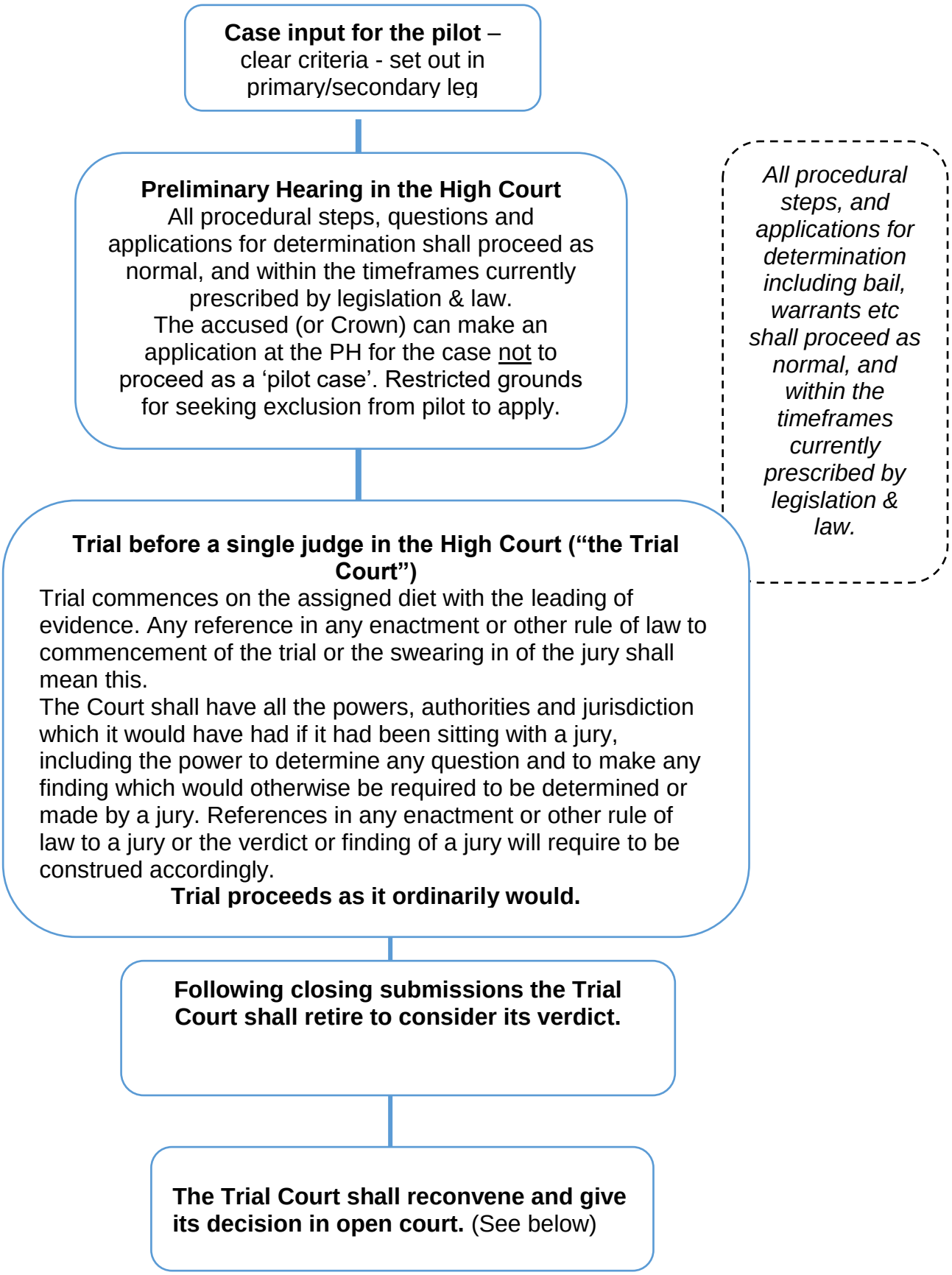
Willmott, D., Boduszek, D., Debowska, A., & Hudspith, L. (2021). Jury decision making in rape trials: An attitude problem. *Forensic Psychology*, 94–119.

Willmott, D., Boduszek, D., Debowska, A., & Woodfield, R. (2018). Introduction and validation of the juror decision scale (JDS): an empirical investigation of the story model. *Journal of Criminal Justice*, *57*, 26–34. https://doi.org/10.1016/j.jcrimjus.2018.03.004

Willock, I. D. (1966). *The origins and development of the jury in Scotland.* Stair Society.

Yarkoni, T. (2020). The generalizability crisis *The Behavioral and Brain Sciences*, *45*, E1. https://doi.org/10.1017/S0140525X20001685

Annex B: proposed model for a single judge pilot for rape cases



The Trial Court reconvenes and give its decision in open court.

Guilty verdict.
The Trial Court shall, at the time of conviction or as soon as practicable thereafter, give a judgment in writing stating the reasons for the conviction.
The extent of detail of the reasons given is something that will have to be considered along with the judiciary.

Acquittal verdict.
The Trial Court shall, in open court, give a short oral judgment stating its reasons.
The principles applying to written reasons for guilty verdicts should apply.
The extent of detail of the reasons given is something that will have to be considered along with the judiciary.

The Trial Court shall pass sentence on the day, or as more commonly occurs, the court may continue the cause for sentence in open court to an assigned date to allow the collation of reports.

Appeals- see below

Appeal against the verdict of Court

Appeals against conviction and/or sentence shall be available with leave of the court by the accused, under the present law as contained in s106, 107 and s108 (Lord Advocate's right against disposal) of the Criminal Procedure (Scotland) Act 1995 (1995 Act) subject to amendments. The appeal in respect of conviction/conviction and sentence shall be before 3 judges. Sentence appeals will have a minimum quorum of 2 judges.

Grounds of appeal: Section 106(3) of the 1995 Act, which makes provision for a right of appeal against conviction by a jury on grounds of a miscarriage of justice shall apply to a verdict of the Trial Court subject to the substitution of references to the 'trial court' in place of/in addition to references to the jury. Namely:

3) … a person may bring under review of the High Court any alleged miscarriage of justice, which may include such a miscarriage based on—

(a) subject to subsections (3A) to (3D) below, the existence and significance of evidence which was not heard at the original proceedings; and

*(b) the jury's/**the trial court** having returned a verdict which no reasonable jury, properly directed,/ trial court could have returned.*

In an appeal under s106(3) alleging a miscarriage of justice, the initial question for the appeal court, applying the dicta of court in *Megrahi* v *HM Advocate* 2002 JC 99 at [26], will be "*whether in arriving at its verdict the trial court misdirected itself in law or as to a matter of fact so that it took a course which it was not entitled to do or failed to do what it should have done. If and to the extent that this has been shown, the further question would be whether a miscarriage of justice has resulted*".

Disposal

The appeal shall be disposed of in accordance with the terms of section 118, in so far as is applicable.

The procedural requirements and timescales for appeals from the decision of a jury trial shall remain and apply. In so far as applicable, steps will include intimation of intention to appeal in accordance with s109, lodging of a note of appeal in terms of s110(1)(a); decision on leave to be considered and made in terms of s107, in so far as applicable. Production of a report in terms of s113. S111 & 112 (admission to bail); s114; s115 (applications-oral/writing); s116 (abandonment) & 117(presence of appellant) to apply.

Annex C - Potential questions and methods for evaluating single judge trial pilot against objectives

	Objective 1 - To assess how the process of conducting a single judge trial for rape cases is perceived by those involved in the trial process	
	The Experience of those Involved in the Trial Process	**The Experience of those involved in examination/questioning of witnesses**
Potential evaluation questions	Trial Judges[1]: • Are there any differences in the evidence led or the way it is presented by counsel in single judge trials compared to jury trials? • What advantages and disadvantages have you identified associated with the requirement and process of providing written reasons for verdicts in criminal solemn cases as standard? • Do you have any other observations on the trial process comparative to Jury Trials?	Trial Judges: • Did the judges noticed in difference in the number of times they had to intervene? interventions? • Do you have any other observations on the cross-examination process comparative to Jury Trials?
	Prosecution Counsel: • What are the key differences in presenting evidence to a single Judge compared to a Jury during the pilot and how do you adapt your approach to take account of this? • How do your preparations for a single judge trial differ from a Jury Trial? • Do you have any other observations on the trial process during the pilot comparative to Jury Trials? • Are there any differences in the evidence led or the way it is presented by defence counsel in single judge trials compared to jury trials?	Defence Counsel/solicitors: • Would the number, focus and presentation of questions you asked during the examination in chief and cross-examination have changed had the trial been taking place in front of a Jury? • How does the examination in chief and cross-examination process compare as a mechanism for presenting evidence/arguments to a Judge comparative to a Jury? • Do you have any other observations on the examination in chief and cross-examination process comparative to Jury Trials? • Are there any differences in the evidence led or the way it is presented by prosecution counsel in single judge trials compared to jury trials?
	Defence Counsel/solicitors:	Prosecution Counsel

[1] It is envisaged that the views of those judges and legal practitioners that have been involved in the pilot will be gathered through semi-structured interviews or questionnaires at the conclusion of the pilot. We are conscious of the extent to which the views of judges can be sought and further consideration would require to be given to what questions could be asked of Judges.

	• What are the key differences in presenting evidence to a Judge compared to a Jury during the pilot and how do you adapt your approach to take account of this? • How do your preparations for a single judge trial differ from a Jury Trial? • Do you have any other observations on the trial process during the pilot comparative to Jury Trials?	• Would the number, focus and presentation of questions you asked during the cross-examination have changed had the trial been taking place in front of a Jury? • How does the cross-examination process compare as a mechanism for presenting evidence/arguments to a Judge comparative to a Jury? • Do you have any other observations on the cross-examination process comparative to Jury Trials?
	Complainer[2]: • How did the trial process compare to your expectations of the process? Have you had any experience of the criminal justice system prior to this? • Did you give live evidence at the trial or was your evidence captured via another process e.g. a written statement or pre-recorded evidence. • Was the trial process longer or shorter than you anticipated? • Was being provided with written reasons for a verdict helpful in understanding how and why a judge arrived at a particular verdict? • Any other observations on the trial process	Complainer: • How did examination in chief and cross-examination compare with your expectations of the process? Did this take place during the trial or via some alternative means e.g. evidence by commissioner? • Do you feel that the questions during the cross-examination process were relevant to and focused on the specifics of the case? Do you have any other observations on the examination in chief cross-examination process ?
	Accused: • How did the trial process compare to your expectations of the process? • Was the trial process longer or shorter than you anticipated • Was being provided with written reasons for a verdict helpful in understanding how and why a judge arrived at a particular verdict? • Any other observations on the trial process	Use of special measures as set out in the Criminal Procedure Act 1995 compared to Jury Trials
		Length of examination in chief and cross-examination compared to Jury Trials?

[2] It is envisaged that the views of those complainers and accused involved in the pilot will be gathered through semi-structured interviews or questionnaires

	Objective 2 - To explore the impact of single judge trials on the effectiveness and efficiency of managing rape trials				
	Trial Duration	**Trial Costs**	**Delays & Adjournments**	**Section 275 Applications**	**Other relevant measures relating to case mgmt**
Potential evaluation questions	What is average length of time from indictment to trial compared to Jury Trials?	Where the accused is in receipt of Legal Aid, the average cost to Legal Aid compared to Jury Trials	How many adjournments were there on average compared to Jury Trials?	What is the average number of applications lodged under Section 275 of the Criminal Procedure Act compared to Jury Trials	Is there any difference in the use of standard or non-standard special measures compared to jury trials, including both number and outcomes of applications?
	What is the average length of time from Preliminary Hearing to Trial compared to Jury Trials?	What is the average cost to COPFS compared to Jury Trials?	How many continued Preliminary Hearing Diets compared to Jury Trials?	-	What are the views of the Judge on the efficiency and effectiveness of the trial process compared to Jury Trials.
	What is the average length of a Single Judge Rape Trial compared to a Jury Trial?	What is the average cost of a Jury Rape Trial? (baseline only)	How many adjourned trial diets compared to Jury trials?	-	
	What is the average length of time from conclusion of the leading of evidence and conclusion of parties submissions to delivery of the verdict by the Judge compared to Jury Trials	What is the average cost to the Court system i.e. court staffing	-	-	What are the views of prosecution counsel on the efficiency and effectiveness of the trial process compared to Jury Trials.
	-	-	-	-	What are the views of defence counsel on the efficiency and effectiveness of the trial process compared to Jury Trials.

Objective 3 - To consider the effect of Single Judge Trials on outcomes			
Conviction rates/verdicts across cases with different characteristics	**Proportion of guilty pleas**	**Prevalence and substance of appeals against verdicts and the success of those appeals**	**Impact on other areas of the justice system**
What is the conviction rate during the course of the pilot?	What is the proportion of guilty pleas compared to Jury Trials?	What is the number of appeals lodged in accordance with Section 106(1)(a) of the Criminal Procedure Act 1995 as compared to jury trials?	What impact, if any, does the pilot have on the time it takes for other cases to come to trial from e.g. PH to trial?
What are the primary reasons for conviction/ acquittal as set out in reasons for verdict?	What proportion of guilty pleas are entered at different stages of the trial process compared to Jury Trials? • Preliminary Hearing • Pre-Trial • During Trial	What are the primary reasons for appeals lodged as set out at Section 106(3) of the Criminal Procedure Act as compared to jury trials?	-
-	-	What proportion of cases are granted leave to appeal by the High Court in accordance with Section 107 of the Criminal Procedure Act as compared to Jury trials?	-
-	-	What proportion of convictions are quashed following an appeal lodged in accordance with Section 106(1)(a) or s106(f) as applicable of the Criminal Procedure Act 1995 compared to Jury trials?	-
-	-	How many appeals are lodged under section 107A of the Criminal Procedure Act 1995 as compared to jury trials?	-
-	-	What are the primary reasons for appeals lodged under section 107A of the Criminal Procedure Act 1995 as compared to jury trials?	-