

“How Should We Go About Jury Research in Scotland?”: A Response

Professor Cheryl Thomas

Director UCL Jury Project

☞ keywords to be inserted by the indexer

It is certainly important that research is carried out in Scotland to better understand the workings of the Scottish jury, especially given its unique features. In light of the call by the Post-Corroboration Safeguards Review for reliable empirical evidence about jury decision-making in Scotland, it is all the more important that this research should be done prior to and inform any changes made to the Scottish jury system.

Unfortunately the research programme set out by Chalmers and Leverick in the previous issue of the *Criminal Law Review*¹ would fail in its aims to assist this reform process because the authors are confused about how the *Review's* questions can be answered through jury research, they fail to appreciate fundamental differences in the research methods used to study juries, and as a consequence they argue for an amendment to the contempt of court act that is completely unnecessary to answer authoritatively the important questions set out by the Post-Corroboration Safeguards Review.

In this article I explain the key differences in jury research methods and, based on this, identify which methods are needed to answer the PCSR questions. I also set out the key errors of the research plan devised by Chalmers and Leverick. This is done in the sincere hope that if any jury research is conducted in Scotland to answer the *Review's* important questions it is done so based on a proper understanding of which research methods need to be used to answer each of the *Review's* questions—and an understanding of why none of this requires any amendment to the contempt of court act.

Jury research to address the issues set out by the PCSR

In this context it would be helpful to set out the six questions the Post Corroboration Safeguards Review identified as needing to be answered before reforms to the Scottish criminal process were introduced:

1. What do jurors understand to be the difference between Not Guilty and Not Proven?
2. Why do jurors choose one verdict option (Not Guilty or Not Proven) over the other?
3. Why, and to what extent, do jurors alter their position as regards Not Proven and Not Guilty as a result of/following deliberations?

¹ J. Chalmers and F. Leverick, “How Should We Go About Jury Research in Scotland” [2016] *Crim. L.R.* 607.

4. To what extent do members of a jury of 15 (as compared with a jury of 12) actually participate in deliberations?
5. What are the differences in outcome (assuming an identical factual matrix) as between a 12 person jury with only two possible verdicts and a 15 person jury with three verdicts, and what are the reasons for those differences?
6. Are there benefits in requiring the jury to attempt to reach a unanimous verdict?

This means there are three main variables with specific associated questions that would be need to be addressed in any jury research to answer the PCSR questions. These are:

Variable one: verdict options

- To what extent is jury decision-making in Scotland affected by the range of verdict options available to juries?
- If the Not Proven verdict option were removed how would it affect case outcomes?
- To what extent does deliberation shift individual jurors' decision-making between different verdict options?

Variable two: jury size

- To what extent does the size of the jury affect juror participation in deliberations?
- To what extent does the size of the jury, combined with different verdict options, affect jury decision-making?

Variable three: verdict ratios

- How is jury decision making affected by the ratios required to reach a verdict?
- Specifically how does a requirement for unanimity or different weighted majorities affect jury decision-making?

Three core jury research methods

To understand how to go about answering the questions posed by the PCSR, one has to understand the defining differences between three types of research methods used to study juries: (i) correlational; (ii) causal; and (iii) attitude/experience studies. I have provided a definition and explanation of each below. Chalmers and Leverick do not discuss or explain these three core types of empirical research with juries and the critical differences between them, and unfortunately this leads them to make crucial errors in identifying how to go about answering the six questions set out by the PCSR. This lack of understanding of how to go about researching the jury system in turn leads them to call for a change in the Contempt of Court Act 1981 that is not needed and, if instituted, would not produce the

reliable empirical evidence the Scottish government needs to decide whether and if so how to reform the jury trial process in Scotland.

Type 1: Correlational studies: large-scale quantitative analysis

A common assumption is that any research that looks at jury decisions in actual cases must be better than research that simulates such decisions, even if the simulation is conducted with real jurors at court. This assumption is incorrect because it fails to distinguish between: (1) studies that can show cause and effect between case factors and jury verdicts; and (2) those that can only show whether there are any associations (or correlations) between them—but cannot prove cause and effect.

Correlational studies can be conducted that look at the extent to which certain case factors (e.g. offence type) are associated with specific jury verdicts. Such studies obviously do not require any amendment to the Contempt of Court Act because they deal only with anonymised data and do not enquire into jury deliberations. But correlational studies do require very large datasets before any findings can even begin to suggest any connection between certain case factors and jury verdicts. In the 2010 study *Are Juries Fair?*, the UCL Jury Project carried out a correlational analysis of all jury verdicts in all Crown Courts in England and Wales between 2006–2009, looking at the relationship between jury conviction rates and a range of case factors such as court location, offence type, defendant ethnicity, number of charges and number of defendants in cases. The dataset for this study encompassed over half a million charges against all defendants, resulting in over 67,000 jury verdicts in more than 15,000 trials. According to Chalmers and Leverick, “Of 66,963 cases in the financial year 2015–16, 4,618 were prosecuted under solemn procedure but of these *only 1,058 proceeded to a trial* [emphasis added], the rest concluding with a guilty plea.” This means that to carry out a comparable analysis of actual jury verdicts in Scotland, this would require at least 15 years’ worth of data from the Scottish courts.

However, even if such data were available in Scotland, such analysis cannot answer whether certain factors actually *affect* jury decision-making. Examining verdicts in actual jury trials has the inherent problem that no two cases are identical. It is therefore not possible to disentangle the factors and influences that may have led to a jury’s decision in one case and to draw conclusions that can be extrapolated beyond that individual case. As a result, actual jury verdict analysis can only ever indicate *correlations* between case factors and case outcomes, not *causal* connections. This important limitation in correlational studies has been recognised by courts.² Almost all (five out of six) of the PCSR’s questions ask about the *effect* of changing jury size, verdict options and verdict ratios on jury decision-making. Therefore, only causal studies can answer almost all the key questions set out in the Scotland Review.

²The US Supreme Court acknowledged that actual case analysis can only identify a discrepancy that appears to correlate with specific case factors. See *McCleskey v Kemp* 481 U.S. 279, 312 (1987).

Type 2: Causal studies—case simulation

Case simulation is recognised as the most methodologically sound approach to examine *causal factors* in decision-making in the justice system.³ In contrast to correlational studies of actual jury verdicts, case simulation permits a systematic and controlled study of jury decision-making and can do what actual verdict analysis and other jury research methods cannot do: reveal how specific factors affect jury decisions. In a reliable and robust case simulation, a real trial is reconstructed on film and edited so that only a specific factor (such as verdict option, verdict ratio or size of the jury) is altered in different versions of the case simulation. This creates a systematic and controlled study of the impact of specific factors on jury decision-making.

Because juries' decisions in case simulations do not have real consequences for an actual defendant, it is crucial that the highest level of authenticity is achieved in the simulation in order to replicate the jury experience as closely as possible. Previous case simulation studies can rightly be criticised for using unrealistic materials, not providing jury instructions or allowing deliberations, and using students or volunteers as "jurors".⁴ The UCL Jury Project has pioneered the use of highly authentic case simulations run with *actual juries at court*.⁵

Simulation cases need to contain all elements of a jury trial (including evidence-in-chief and cross-examination of all witnesses, prosecution and defence closings and judge's directions to the jury). Each version of the case is shown to a large number of different real juries - made up of jurors at court who have served on trials and have just come to the end of their jury service. This is a critical element of the research. Having just experienced the trial process, these participants in the UCL Jury Project case simulations are fully immersed in the actual experience of being a juror, which is something that cannot be replicated in any other type of "mock" jury research. Chalmers and Leverick incorrectly state that my research used jurors who were summoned for jury service but were not used as serving jurors in any trials.⁶

After viewing the case simulation, each member of the jury records an initial decision and a degree of confidence in that assessment. Jurors then deliberate as a jury in an effort to reach a verdict. After deliberating jurors record their final decisions, as well as impressions of evidence and witnesses and more general views. Each variation of the case needs to be run with a large number of different juries at a representative sample of courts to ensure the reliability of results and to determine the extent to which the findings can be generalised. Any differences in jury verdicts that emerge across the case simulations can therefore be attributed to the specific factor that was varied (e.g. verdict option, verdict ratio, number of

³ S. Sommers and P. Ellsworth, "How Much Do We Really Know About Race and Juries? A Review of Social Science Theory and Research" (2003) 78(3) *Chicago-Kent Law Review* 997.

⁴ Sommers and Ellsworth, "How Much Do We Really Know About Race and Juries? A Review of Social Science Theory and Research" (2003) 78(3) *Chicago-Kent Law Review* 997.

⁵ See C. Thomas, *Diversity and Fairness in the Jury System*, 2007 Ministry of Justice research Series for details of the first case simulation study with serving jurors in England and Wales, as well as the follow up studies: C. Thomas, *Are Juries Fair?* 2010 Ministry of Justice Research Series and C. Thomas, "Avoiding the Perfect Storm of Juror Contempt" [2013] *Crim. L.R.* 483.

⁶ At p.708 Chalmers and Leverick state: "... the research Cheryl Thomas carried out in England for the Ministry of Justice, in which she tested the extent to which race influenced jury decision making *using jurors who were cited but not selected* [emphasis added] ...".

jurors). Such case simulations using real experienced jurors at court also do not require any amendment to the Contempt of Court Act because these jurors are not being asked to reveal any information about what was said or done in the deliberating room during the actual cases on which they served.

Type 3: Attitude/experience studies: post-verdict research

Post-verdict questioning of jurors in actual cases can also be beneficial if used in the correct circumstances to answer a specific type of question. This type of study has the benefit of dealing with real juries and real cases. But it too has significant limitations. Post-verdict studies rely entirely on jurors’ self-reported perceptions and recollections of their time on a jury, and it is well documented that a substantial element of decision-making can be subconscious and individuals often lack the ability to accurately identify all the factors that influence their judgement and behaviour.⁷ So this type of study is not very useful for understanding *why* juries reach verdicts (which five of the six PCSR questions are concerned with). Instead it is most useful in understanding *how* jurors experience, perceive and understand aspects of trials. Unfortunately, Chalmers and Leverick have selected this research method to try and answer almost all of the PCSR questions.

It is critical in post-verdict studies that a very high participation rate is achieved with each jury, as this helps to ensure a near complete picture of a jury’s impressions in any case. The study also needs to be carried out at court directly after jurors return a verdict and before they leave court. This ensures that jurors are not influenced by any subsequent events and there is no lack of recollection due to a time lag between the trial and the survey. The UCL Jury Project has used of this research method with juries in England and Wales on numerous occasions over the last decade.⁸ This research has been able to determine (amongst other things) the extent to which jurors understand the rules on juror contempt and internet use, as well as their views on judges’ legal directions, summing up of the facts of the case and their views on the process of jury deliberations.⁹ None of this post-verdict research in England and Wales has required any amendment to the Contempt of Court Act. Oddly, Chalmers and Leverick make no mention at all in their article about any of this post-verdict jury research conducted with real jurors in England and Wales (the most recent of which was published in this journal in 2013).¹⁰

What methods are needed to answer the PCSR’s questions?

Given these fundamental distinctions between jury research methods, it should be clear which methods are required to answer the PCSR’s six questions reliably and robustly to produce the empirical evidence that would be needed before the Scottish government proceeds to any reform of the jury system.

⁷ Nisbett & Wilson, “Telling More than We Can Know: Verbal Reports on Mental Processes” (1977) 84 *Psychology Review* 231.

⁸ See Thomas, *Diversity and Fairness in the Jury System*, 2007 Ministry of Justice research Series for details of the first case simulation study with serving jurors in England and Wales, as well as the follow up studies: Thomas, *Are Juries Fair?* 2010 Ministry of Justice Research Series and Thomas, “Avoiding the Perfect Storm of Juror Contempt” [2013] *Crim. L.R.* 483.

⁹ See Thomas, “Avoiding the Perfect Storm of Juror Contempt” [2013] *Crim. L.R.* 483.

¹⁰ See Thomas, “Avoiding the Perfect Storm of Juror Contempt” [2013] *Crim. L.R.* 483.

PCSR Questions 2, 3, 4, 5 and 6: Because these are all cause and affect questions, they all can *only* be answered authoritatively through a controlled, *causal study* of jury decision-making using the case simulation method. In their article Chalmers and Leverick propose using post-verdict surveys to answer Questions 2, 3, 4 and 6, but as the discussion above makes clear such an approach is not designed to and is incapable of producing reliable empirical answers to those questions.

PCSR Question 1: Question 1 can be explored through both case simulation and post-verdict surveys with jurors in actual cases. Therefore, the most reliable approach to answering Question 1 would be to explore this issue through *both* attitude (post-verdict surveys) and causal (case simulation) methods. Such a multi-method approach would be able to test both the *perceived* and *actual* benefits of requiring unanimity through both real world *perception* findings and simulated *causal* findings.

In what follows I explore the mistakes and misunderstandings Chalmers and Leverick make in their research plan for each of the Post Corroboration Safeguards Review questions.

Question 1: What do jurors understand to be the difference between Not Guilty and Not Proven? Chalmers and Leverick propose to survey jurors post-verdict and ask them what they understood the verdict options to mean. While it is perfectly possible to ask Scottish jurors post-verdict in actual cases about their understanding of the Not Guilty and Not Proven verdict options, as explained above, the value of doing so is limited because no two criminal cases are ever the same, and there is no way to control for how the different facts of each case may have affected juror interpretations or applications of the verdict options. In addition another major drawback in exploring this question through post-verdict studies with jurors is that there is no control from one case to another in terms of how juries are directed by the judge on this issue. While there is a standard legal explanation of the difference between verdict options, each judge will direct the jury in his or her own unique way, and there is no way to control for this variability in the research method proposed by the authors. Unfortunately, they do not seem to realise that either of these major limitations exist in their chosen research method.

As explained above, a definitive answer to Question 1 can only be achieved through case simulation research in which all juries receive the same identical judicial direction on Not Proven, Not Guilty and Guilty. While it could be helpful to ask jurors post-verdict in a wide range of actual cases to describe their understanding of each of the 3 verdict options, such post-verdict surveys could only even provide supplemental information to a case simulation study, not a definitive answer to Question 1.

In relation to the PCSR's Questions 2 and 3, Chalmers and Leverick state:

“Q2 (why the jury chose not proven over not guilty) and Q3 (why, and to what extent, do jurors alter their position as regards Not Proven and Not Guilty as a result of deliberations) could also be addressed by research with real jurors, although here the research questions could only be addressed by surveys of jurors following cases in which a not proven verdict was actually returned (or in relation to Q3 either of the two acquittal verdicts)”.

This again is incorrect, with the authors again failing to appreciate the significant problems in trying to answer those two questions by surveying jurors post-verdict. As explained above, with such an approach there is no control in terms of how the facts of each case affect juror interpretations and applications of each verdict option and no control in terms of how the juries are directed by the judge. To authoritatively answer Questions 2 and 3 researchers have to ask jurors (who are all different) to try and decide the *same* case, and then explore why Not Proven was chosen over Not Guilty and how deliberations may have affected their verdict options. Chalmers and Leverick compound their error by going on to state that: “Meaningful research into these questions would, however, require amending the Contempt of Court Act 1981.” This of course is incorrect. Meaningful research into these two questions can only be done through case simulation, which does not require any amendment to the Act.

For Question 4, Chalmers and Leverick become even more confused about how to conduct reliable jury research by claiming that:

“Q4 (the extent to which the members of a jury of 15 as compared with a jury of 12 actually participate in deliberations) could not be addressed with research into real Scottish juries alone, although research into the extent to which members of 15 person Scottish juries participate in deliberations might be compared with research into 12 person juries elsewhere.”

This is incorrect on several levels. Determining whether the number of jurors on a jury affects jury deliberations can only be validly tested through a case simulation (with real Scottish jurors)—in which large numbers of 15-member juries and large numbers of 12-member juries *try the same case*, and the level of deliberation participation is assessed. Given the uniqueness of each actual case at court, valid comparisons cannot be made between deliberation participation in *different cases*. Chalmers’ and Leverick’s additional idea that deliberations by Scottish 15-member juries would be compared to deliberations by non-Scottish 12-member juries simply compounds their error in research method. There would then be even more factors that the researchers would be unable to control for in such a multi-jurisdiction comparison (i.e. differences in legal systems, the law, jury process as well as cultural differences).

In their research plan for Question 5, Chalmers and Leverick at least recognise that a controlled case simulation study is required to answer this question, but they go on to make an astonishing suggestion that the Scottish government should in fact drop Question 5 from any jury research programme.

“The exception is Q5 (differences in outcome between differently constituted juries with identical factual matrices) Here, there are two options. First, this question could be addressed by a distinct programme of mock jury research The second option would be for the government to consider omitting this question from the programme of research altogether. While any research which adds to our stock of knowledge about jury decision-making is in principle valuable, the gains from this programme of research are likely to be marginal at best and come at significant cost given the scale of the research project required.”

The suggestion that Question 5 should be dropped from any Scottish jury research programme related to the PCSR betrays a failure to understand the central importance of Question 5 to the entire potential programme of jury reform. The overall question that needs to be addressed by any research to aid the PCSR is: *How would altering the jury size, verdict options and verdict ratios in Scotland affect jury decision-making?* Question 5 goes to the very heart of this, and answering Question 5 is therefore critically important before any reform to the Scottish jury is undertaken. The fact that the authors go on to suggest that Question 5 should be dropped altogether from any Scottish jury research programme related to the PCSR because it would be too time consuming for them to try and answer this question properly further reflects their lack of understanding of the importance of this question.

Finally in relation to Question 6, the authors demonstrate their confusion over what is and is not a research question and which PCSR questions can or should be researched:

“In terms of the questions identified by the PCSR (set out earlier in this paper), there are some that are not research questions as such, but are matters for consideration following the results of this research (and other relevant research projects and discussions). In this category are Q6 (whether there are benefits in requiring the jury to attempt to reach a unanimous verdict) ...”.

Chalmers and Leverick are incorrect in saying that Question 6 is not a research question. The question of whether a jury should be required to reach a unanimous verdict has been the subject of extensive research and legal debate for decades.¹¹ This is because it is widely recognised that there may be important benefits to jurors being asked (at least initially) to try and reach a unanimous verdict. The benefits can be to the jurors themselves, but also more widely to parties to criminal cases and the criminal justice system in terms of the validity of the verdict. The benefits to jurors in Scotland can be explored by case simulation where large numbers of different juries decide the same case and in some of these simulations juries are required to attempt to reach a unanimous verdict while in others they are not required to do so.

Contempt of Court Act

Chalmers and Leverick’s failure to understand that specific types of questions about juries can only be answered by using specific research methods, unfortunately, leads them to become confused about the need to amend the Contempt of Court Act to enable the PCSR’s six questions to be answered. All of the research methods needed to answer the PCSR’s six questions operate fully within any existing legal restrictions on discussing deliberations with jurors in actual cases. It should also be noted that all of the types of research needed to answer the Review’s questions have been conducted in England and Wales over

¹¹ See for example: M.J. Saks, *Jury Verdicts—The Role of Group Size and Social Decision Rule* (1977) DC Heath & Co; Emil J. Bove III, “Preserving the Value of Unanimous Criminal Jury Verdicts in Anti-Deadlock Instructions” (2008–2009) 97 *Georgetown Law Journal* 251; Peter J. Couglan, “In Defense of Unanimous Jury Verdicts: Mistrials, Communications, and Strategic Voting” (2000) 94(2) *American Political Science Review* 375.

the last decade fully within the bounds of s.8 the Contempt of Court Act.¹² This may be surprising to many because there is a common misunderstanding about what can and cannot be researched within the parameters of s.8. The incorrect view (to which Chalmers and Leverick fall victim) is that s.8 prevents research with real juries at court. It does nothing of the sort. All it prevents is interrogation of jurors about what was said in the deliberating room. As the discussion of jury research methods in this note makes clear, such inquiries have very limited value in actually understanding *how* juries reach verdicts and the factors that can affect decision-making.¹³

Scotland’s Post Corroboration Safeguards Review laid out a clear set of critical questions requiring empirical research before any reforms are made to the Scottish jury system. The Review’s questions demand and deserve answers that come from robust empirical research with actual Scottish juries at court. Unfortunately, the research plan set out by Chalmers and Leverick cannot deliver this.

¹² See Thomas, *Diversity and Fairness in the Jury System*, 2007 Ministry of Justice Research Series; Thomas, *Are Juries Fair?* 2010 Ministry of Justice Research Series and Thomas, “Avoiding the Perfect Storm of Juror Contempt” [2013] *Crim. L.R.* 483

¹³ For a fuller discussion of the myth that the contempt of court act prevents research with actual jurors, see C. Thomas, “Exposing the Myth” *Counsel* (April 2013).