



Chalmers, J. and Leverick, F. (2016) Jury research in Scotland: a rejoinder. *Criminal Law Review*, 2016(12), pp. 924-930.

This is the author's final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

<http://eprints.gla.ac.uk/130146/>

Deposited on: 17 October 2016

Enlighten – Research publications by members of the University of Glasgow  
<http://eprints.gla.ac.uk>

*This is a pre-copyedited, author-produced version of an article accepted for publication in the Criminal Law Review. The definitive published version (J. Chalmers and F. Leverick, "Jury Research in Scotland: A Rejoinder", forthcoming in the 2016 Criminal Law Review at the time of deposit) is available online on Westlaw UK or from Thomson Reuters DocDel service.*

## **Jury Research in Scotland: A Rejoinder**

James Chalmers  
*School of Law, University of Glasgow*

Fiona Leverick\*  
*School of Law, University of Glasgow*

We are grateful to the *Criminal Law Review* for the opportunity to publish a rejoinder to Professor Thomas's response to our paper "How Should We Go About Jury Research in Scotland?"<sup>1</sup> Thomas disagrees with our conclusions about how the research questions identified by the Post-Corroboration Safeguards Review (PCSR)<sup>2</sup> would most usefully be answered. While there is more common ground between us than might initially be evident, there are a number of points which warrant a response. These can be broadly divided into general issues concerning our understanding of jury research methods and issues concerning our conclusions about the methods that should be used to address the questions identified by the PCSR.

### **General issues**

Thomas asserts that we "fail to appreciate fundamental differences in the research methods used to study juries". She identifies three different types of research method – "correlational", "causal" and "attitude/experience" studies – and objects that we "do not discuss or explain these three core types of empirical research with juries and the critical differences between them".

It is true that we did not discuss what Thomas terms "correlational studies".<sup>3</sup> Such studies conduct statistical analysis of existing trial records in an attempt to identify whether particular factors are associated with specific outcomes. We did not discuss them because they are irrelevant in the present context. None of the six PCSR questions could usefully be addressed by undertaking correlational studies, as Thomas acknowledges. For the same reason we did not discuss (except in passing)<sup>4</sup> field experimental methods. Such methods could not be used to answer any of the PCSR questions.<sup>5</sup>

It is incorrect to say, however, that we do not discuss or explain "causal" or "attitude/experience" studies. The discussion of these two research methods comprises a substantial part of our paper (pp. 703-709). The terminology that we use differs – what Thomas terms "attitude/experience studies" we discuss under the heading "research with real jurors"<sup>6</sup> and what she terms "causal studies" we discuss

---

\* We gratefully acknowledge the assistance of Emma Ainsley and the support of the Leverhulme Trust.

<sup>1</sup> [2016] Crim L.R. 697. In the remainder of this rejoinder, all unattributed page numbers refer to this article.

<sup>2</sup> The Post-Corroboration Safeguards Review, *Final Report* (2015).

<sup>3</sup> Other writers have tended to refer to them as "archival analysis": see e.g. M. Chesterman, J. Chan and S. Hampton, *Managing Prejudicial Publicity: An Empirical Study of Criminal Jury Trials in New South Wales* (2002), para. 81; D.J. Devine, L.D. Clayton, B.B. Dunford, R.P. Seying and J. Pryce, "Jury Decision Making: 45 Years of Empirical Research on Deliberating Groups" (2001) 7 *Psychology, Public Policy and Law* 622, 626.

<sup>4</sup> p. 704.

<sup>5</sup> p. 705.

<sup>6</sup> pp. 703-707.

under the heading “research with mock juries”.<sup>7</sup> We would suggest that we provide a well-balanced and comprehensive account of the advantages and drawbacks of each.

Thomas also discusses both methods, starting with mock jury studies. There is much in her discussion with which we agree. As we note, such studies “[permit] the manipulation of trial variables while leaving other factors constant and [also allow] for the observation of ‘jury’ deliberations rather than relying on self-reporting”.<sup>8</sup> Where we depart is in our recognition of the considerable drawbacks of such studies.

The main drawback of mock jury studies is that they can never replicate the trial process with complete authenticity.<sup>9</sup> Their external validity – the extent to which findings are generalisable beyond the experimental setting – is always limited. Thomas notes that when undertaking mock jury studies, 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”. The difficulty is that without near-infinite time and resources, this is simply impossible. Real trials, especially in more serious cases, commonly run over the course of several days. They will be interrupted by lunch breaks and possibly also by admissibility disputes determined in the jury’s absence. They also involve a genuine solemnity – the jurors know that they hold the fate of a real accused person in their hands in a way that may affect their behaviour. Simulations – even relatively realistic ones – cannot fully replicate these features and will always be vulnerable to the objection that jurors might not behave in a real case in the same way.

The case simulation research that Thomas has conducted is impressive in the extent to which it has attempted to address external validity. She has conducted one of the most realistic mock jury studies to date. It used a filmed trial (based on a real case) in which legal professionals and actors participated and allowed jurors to deliberate in groups of the same size as a legally valid jury.<sup>10</sup> But the simulation did not replicate the real case in several important respects. It was a video of one hour<sup>11</sup> rather than a trial of several hours (or days)<sup>12</sup> and the juries were limited to deliberating for 20 minutes before being asked for a verdict.<sup>13</sup>

In her response, Thomas objects that we are incorrect when we state that her research was undertaken with jurors who were cited but not selected, because her research used juries “made up of jurors at court who have served on trials and have just come to the end of their jury service”. Our understanding of her methodology was (as the citation we gave indicates) based on the account in her report *Are Juries Fair?*,<sup>14</sup> but we are grateful for this clarification. Although Thomas describes this as “a critical element of the research”, we would be sceptical of any claim that this improves external validity. It means that a mock jury is, unlike a real jury, necessarily comprised entirely of persons with immediate prior experience of serving on a jury.

On Thomas’s own account, this poses particular problems in the context of the PCSR questions. She argues that surveying jurors post-verdict about their understanding of the “not guilty” and “not

---

<sup>7</sup> pp. 707-709.

<sup>8</sup> p. 707.

<sup>9</sup> p. 707.

<sup>10</sup> C. Thomas, *Are Juries Fair?* (2010), p.8; C. Thomas (with N. Balmer), *Diversity and Fairness in the Jury System* (2007), pp. 42-43.

<sup>11</sup> C. Thomas (with N. Balmer), *Diversity and Fairness in the Jury System* (2007), p.43, fn. 139.

<sup>12</sup> The trial it was based on lasted for 15 hours: *ibid.*

<sup>13</sup> See *ibid.* In the real trial on which it was based the jury deliberated for 5 hours: *ibid.*

<sup>14</sup> C. Thomas, *Are Juries Fair?* (2010), p.8. (“Each jury was a panel created by Her Majesty’s Courts Service (HMCS) computerised random selection system for a trial at court. When a panel was not needed for a trial and all jurors on the panel were about to be dismissed from jury service, they were asked if they would participate in the study before leaving court.”)

proven” verdicts would be inadequate, because “there is no control from one case to another in terms of how juries are directed by the judge on this issue”. This alleged failing, however, is replicated in any research where mock jurors have received judicial directions in a different (real) case shortly or even immediately before viewing a case simulation. This is illustrative of a general problem, which is that if a mock jury is constituted in this fashion, it is difficult to attribute the research results to any specific factor varied in the experiment, as the causal effect of that factor cannot be separated out from any factors the jurors have been exposed to during the previous trial.

Thomas proceeds to discuss “attitude/experience” studies (in our terminology, real juror studies).<sup>15</sup> These involve surveying jurors (whether by questionnaire or interview) who have sat on real cases about that case. This method has been used worldwide to inform policy questions similar to those set out by the PCSR.<sup>16</sup> The best example is perhaps the research undertaken as part of the New Zealand Law Commission’s project on juries in criminal trials.<sup>17</sup> The Law Commission’s report led to a number of key changes, including the introduction of 11:1 majority verdicts<sup>18</sup> and the possibility of trial by judge alone in complex cases.<sup>19</sup> These illustrate the distinctive value of research with real jurors, because some of the issues identified by the research which led to these changes being recommended (for example, the rare “rogue juror” and the confusing and unsatisfactory way in which evidence was presented in some individual cases<sup>20</sup>) are not problems which could be reliably, if at all, evidenced via research with volunteer mock jurors deliberating after watching short case simulations.

We are, incidentally, surprised by Thomas’s claim that we “fall victim” to the “incorrect view... that s.8 [of the Contempt of Court Act 1981] prevents research with real juries at court”. In fact, we say on four separate occasions in our paper that it does not<sup>21</sup> and cite research with real jurors which has been carried out in compliance with section 8.<sup>22</sup> Section 8 only prevents research which asks jurors who have served on an actual criminal trial about the content of their deliberations. The question we address is whether such research is desirable in the context of the PCSR.

Thomas states that post-verdict questioning of real jurors can be beneficial in answering certain research questions. We agree. It provides information about real trials, avoiding the external validity

---

<sup>15</sup> Thomas also describes case simulations of the type which she has carried out as research with “real jurors”. We do not take issue with this difference in terminology, but it should be borne in mind in reading her response and our rejoinder.

<sup>16</sup> And also the approach of the courts: see *Montgomery v. H.M. Advocate*, 2001 S.C. (P.C.) 1 at 30 (discussing the approach which the courts should take to pre-trial publicity and prejudicial media coverage).

<sup>17</sup> W. Young, N. Cameron and Y. Tinsley, *Juries In Criminal Trials Part Two: A Summary Of The Research Findings* (1999), discussed in more detail in our original paper at pp. 705-706. The research informed the Law Commission’s reports: New Zealand Law Commission, *Preliminary Paper on Juries in Criminal Trials: Part I* (1998), Law. Com. PP32 and New Zealand Law Commission, *Discussion Paper on Juries in Criminal Trials: Part II* (1999), Law. Com. PP37.

<sup>18</sup> Juries Amendment Act 2008 s.19, inserting a new s.29C into the Juries Act 1981.

<sup>19</sup> Criminal Procedure Act 2011 s.107.

<sup>20</sup> See New Zealand Law Commission, *Juries in Criminal Trials* (Report 69, 2001), paras 435 and 87.

<sup>21</sup> See p. 704 (“The Act does not rule out asking real jurors about aspects of their experience as jurors other than what went on in the course of deliberations or about their understanding of legal terms or concepts.”); *ibid.* (“Moreover, the Act does not rule out asking jurors some limited questions about their deliberations”); p. 711 fn. 102 (“Thomas [argues] that it is a myth that s.8 prevents jury research. We agree entirely.”); p. 712 (“If the Scottish Government (or the Parliament) is unwilling to amend the 1981 Act, it would still be possible to carry out research with real jurors.”).

<sup>22</sup> J. Jackson, “Juror Decision-Making and the Trial Process” in G. Davies et al. (eds), *Psychology, Law and Criminal Justice* (Berlin: De Gruyter, 1996), pp.327; M. Zander and P. Henderson, *The Crown Court Study* (Royal Commission on Criminal Justice, Research Study No.19, 1993). Section 74 of the Criminal Justice and Courts Act 2015 repealed and replaced s.8 of the 1981 Act in relation to England and Wales, but s.8 remains the applicable law in Scotland.

issues inherent in mock jury studies. It has weaknesses, as we accept and discuss,<sup>23</sup> but these can be minimised, as Thomas acknowledges. We note, as Thomas does, that to minimise problems of recall it is important to survey jurors as soon as possible after the trial.<sup>24</sup> Ensuring that multiple jurors are asked about the same trial can also help to minimise issues of accurate recall and representation of events<sup>25</sup> – in the New Zealand study the researchers reported that “five or six interviews per trial were sufficient to provide an accurate and consistent picture of the deliberation process”.<sup>26</sup>

Crucially, as we made clear, mock jury research and real jury research *both* have drawbacks. Thomas’s response does not fully acknowledge this. No method of jury research is perfect. The challenge is to select the most appropriate method to answer the particular research questions and to try and address the potential weaknesses of that method in the research design.

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

Some comment is required on Thomas’s responses to our suggestions for addressing each of the PCSR questions, as follows:

*Question 1: What jurors understand to be the difference between Not Guilty and Not Proven.* Thomas argues that because “each judge will direct the jury in his or her own unique way, and there is no way to control for this variability”, a “definitive answer” to this question can only be achieved via case simulations “in which all juries receive the same identical judicial direction on Not Proven, Not Guilty and Guilty”.

We disagree: it is *because* each judge will direct the jury in his or her own unique way that case simulations cannot provide a definitive answer here. At best, case simulations will identify what jurors understand to be the difference between the verdicts when directed in a specific way in the context of a specific simulation. While such research clearly would be of value, it could not provide any “definitive answer”. Again, both real and mock jury research has benefits and limitations in this and other contexts. We would note, though, that research with real jurors has previously been employed to explore juror comprehension of judicial directions, as in the New Zealand study.<sup>27</sup>

*Question 2: Why [jurors] choose one [verdict of acquittal] over the other and Question 3: Why, and to what extent, do jurors alter their position as regards Not Proven and Not Guilty as a result of deliberations.* As with all the PCSR research questions, both methods of research have their limitations here. Case simulations would be particularly appropriate if the PCSR questions aimed to test whether specified variables affected jurors’ choice of verdicts or the deliberative process. The questions are, however, open-ended and do not do this. Instead, they invite the researchers to provide an account of jury deliberations and decision-making which recognises the variety of juror experience, and so are best answered through qualitative research with real jurors. An approach based on asking all study participants “to try and decide the same case”, as Thomas suggests, would fail to reflect the varied range of experiences of juries in different cases and provide limited insight. In

---

<sup>23</sup> pp. 706-707.

<sup>24</sup> p. 707.

<sup>25</sup> p. 706.

<sup>26</sup> W. Young, N. Cameron and Y. Tinsley, *Juries In Criminal Trials Part Two: A Summary Of The Research Findings* (1999), para. 1.7.

<sup>27</sup> New Zealand Law Commission, *Juries in Criminal Trials* (Report 69, 2001), paras 7.30-7.57. For other examples see e.g. G. Kramer and D. Koenig, “Do jurors understand criminal jury instructions? Analyzing the results of the Michigan juror comprehension project” (1990) 23 *University of Michigan Journal of Law Reform* 401; A. Reifman, S.M. Gusick and P.C. Ellsworth, “Real jurors’ understanding of the law in real cases” (1992) 16 *Law and Human Behavior* 539.

contrast, by interviewing jurors in a wide range of different cases, the New Zealand study was able to provide a rich and informative picture of the deliberation process.<sup>28</sup>

Thomas also makes the following comment:

“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.”

We did not say that. We said that “[m]eaningful research into these questions *with real jurors* would, however, require amending the Contempt of Court Act 1981”.<sup>29</sup> This should be uncontroversial: the 1981 Act clearly prohibits asking Scottish jurors about their deliberations. We would not and did not dispute that research addressing these questions could be carried out via case simulations; the question is only whether research with mock jurors or real jurors is preferable.

*Question 4: The extent to which the members of a jury of 15 (as compared with a jury of 12) actually participate in deliberations.* We are unsure why Thomas takes issue with what we said regarding this question. As we acknowledged, this question cannot be answered via real jury research, and comparison with research elsewhere will be of limited value simply because there is a lack of such research, quite aside from the difficulties of multi-jurisdictional comparison. It would, however, be possible to examine the extent of participation within the 15 person jury by way of research with real jurors.

*Question 5: The differences in outcome (assuming an identical factual matrix) as between a 12 person jury with only 2 possible verdicts and a 15 person jury with 3 verdicts, and the reasons for these differences.* Here, we agree that this question can only be answered through case simulations. Our comments about the limited utility of this question, however, remain unaffected by Thomas’s response. What proposition is to be tested by such research? If it transpires that, given a particular case simulation, the 15 person jury is more likely or less likely to convict than the 12 person jury, which system is preferable? As others have noted, “[r]esearchers cannot say whether the result reached by a jury is correct or incorrect”.<sup>30</sup> Thomas asserts a “lack of understanding of the importance of this question” on our part, but we would suggest instead that an explanation of its importance would be helpful.

One possibility, not discussed by Thomas, is that if a case simulation is seen repeatedly by different juries, the “correct” verdict might be taken to be that reached most often (the theory being that this reflects the preference of the community). This allows for a testable proposition to be identified, which is that larger juries – being more representative – may be more likely than smaller juries to reach that verdict (and so more reliable in reflecting the community’s preference).<sup>31</sup> A meta-analysis of 10 studies comparing six and twelve-person juries in the United States found no significant effects in this regard.<sup>32</sup> Despite that, the results of such a programme of research comparing the 12 and 15 person jury would be of interest. The Scottish Government will have to decide whether it has the

---

<sup>28</sup> W. Young, N. Cameron and Y. Tinsley, *Juries In Criminal Trials Part Two: A Summary Of The Research Findings* (1999), ch. 6.

<sup>29</sup> p. 710, emphasis added.

<sup>30</sup> A. Smith and M.J. Saks, “The case for overturning *Williams v. Florida* and the six-person jury: history, law, and empirical evidence” (2008) 60 *Florida Law Review* 441, 467.

<sup>31</sup> Conversely, it may be possible that above a certain size, juries deliberate less effectively and are more likely to return “wrong” verdicts.

<sup>32</sup> M.J. Saks and M.W. Marti, “A meta-analysis of the effects of jury size” (1997) 21 *Law and Human Behavior* 451, 461.

resources to fund such research;<sup>33</sup> we hope it does but recognise that the Government may have to make difficult choices as to which of the PCSR's questions to prioritise if funds are limited.

*Question 6: Whether there are benefits in requiring the jury to attempt to reach a unanimous verdict.* Here, the differences between our own views and those of Thomas may be more apparent than real. The question of whether a jury system should require a unanimous verdict (or at least an attempt to reach one) cannot be answered directly by empirical research. This is not an example of "confusion over what is and is not a research question", but recognition that some questions are ones of policy to be decided on the basis of sound empirical knowledge.

Thomas suggests that the point can be explored "by case simulation where large numbers of different juries decide the same case". This is correct, but significant practical limitations should be acknowledged. First, asking all juries to determine the same case will close off any possibility of identifying whether particular types of case pose particular difficulties in terms of achieving unanimity. Secondly, it is unlikely that researchers will be able to permit mock jurors to deliberate indefinitely,<sup>34</sup> making it difficult to determine when attempts to reach unanimity have failed. Case simulations may deem attempts to reach unanimity to have failed after a set time period, but with the result that any apparent difference between unanimity and majority decision rules identified by the research may simply be the consequence of this artificiality.<sup>35</sup>

Even if such artificiality is minimised, volunteer mock jurors are unlikely to have unlimited time or patience, and may be more inclined to compromise in deliberations (as this would accelerate the end of their commitment and their return home) than if they were deciding the fate of a real person.<sup>36</sup> The extent to which case simulation research can meaningfully inform a decision on whether a jury system should employ unanimity or majority decision rules is therefore very limited. Unsurprisingly, in examining this issue, law reform bodies have, rather than relying on data from case simulation research, placed emphasis on research with real jurors and empirical data about the rates at which actual juries have failed to reach agreement.<sup>37</sup>

## Conclusion

As we said earlier, there is considerable common ground between our original paper and Thomas's response. Our disagreement might be summed up as follows: we believe that it is important to acknowledge the limitations of *both* real jury and mock jury research, and to design any study accordingly. Thomas's response refers to the "definitive answer" which she suggests might be

---

<sup>33</sup> The question of resources is particularly acute because, unlike the U.S. research which we cite here, jury size is not the only relevant variable. A programme of research would have to consider (a) the 12 and 15 person jury; (b) the two and three verdict systems and perhaps also (c) unanimity, qualified unanimity and simple majority decision rules (cf. Q6 of the PCSR). 12 different permutations of (a), (b) and (c) are possible, meaning that if all permutations were to be considered a very high number of simulations would need to be carried out in order for the research to generate reliable results. Moreover, it is difficult to see how proposals for reform of the Scottish jury system could be based on a study employing a single factual matrix, given specific concerns relating to sexual offences and domestic violence which permeate the PCSR.

<sup>34</sup> A jury cannot return a majority verdict in England and Wales without having had at least two hours for deliberation (Juries Act 1974 s.17(3)). In New Zealand, the relevant period is four hours (Juries Act 1981 s.29C(2)(a)). As we noted above, juries in Thomas's earlier case simulation research were limited to deliberating for 20 minutes.

<sup>35</sup> V.P. Hans, "The unanimity requirement: issues and evidence", in Law Reform Commission of Canada, *Studies on the Jury* (1979) 141, 151-152.

<sup>36</sup> The difficulty being akin to that identified by M. Findlay, *Jury Management in New South Wales* (1994), p.31, where jurors raced to fill out very lengthy (104 question) post-verdict questionnaires, missing out questions.

<sup>37</sup> See e.g. Law Reform Commission of Canada, *The Jury in Criminal Trials* (1980), pp. 19-32; New Zealand Law Commission, *Juries in Criminal Trials* (Report 69, 2001), Ch. 13.

obtained to one of the PCSR's questions by carrying out case simulation research. No research methodology, however, will provide *definitive* answers to the questions raised by the PCSR. The knowledge obtained by jury research in either form will be valuable. It may provide a basis for legislative reform in due course. But it will be imperfect knowledge. The challenge for researchers is to identify, within the limits of practicality and available resources, how these imperfections can be minimised. With all of this in mind, we remain convinced that the method that would be most useful for tackling the majority of the PCSR's questions is real jury research.