Myths about Myths? A Commentary on Thomas (2020) and the Question of Jury Rape Myth Acceptance.

Ellen Daly, Anglia Ruskin University, UK
Olivia Smith, Loughborough University, UK
Hannah Bows, Durham University, UK
Jennifer Brown, London School of Economics, UK
James Chalmers, Glasgow University, UK
Sharon Cowan, University of Edinburgh, UK
Miranda Horvath, Middlesex University, UK
Fiona Leverick, Glasgow University, UK
Jo Lovett, London Metropolitan University, UK
Vanessa Munro, University of Warwick, UK
Dominic Willmott, Manchester Metropolitan University, UK

This commentary responds to claims that research by Cheryl Thomas ‘shows’ no problem with rape myths in English and Welsh juries. We critique the claim on the basis of ambiguous survey design, a false distinction between ‘real’ jurors and other research participants, the conflation of attitudes in relation to abstract versus applied rape myths, and misleading interpretation of the data. Ultimately, we call for a balanced appraisal of individual studies by contextualising them against the wider literature.

Introduction

In a recent article for the Criminal Law Review, Professor Cheryl Thomas (2020: 1004) asserted that ‘previous claims of widespread “juror bias” in sexual offences cases are not valid’. The conclusion was drawn from Thomas’ pioneering 2018-19 study with 771 jurors (including an unknown number from sexual offence trials) across four English and Welsh court regions. Participants were asked whether they agreed or disagreed with a series of attitudinal statements about rape immediately upon completion of their jury service. The research was ground-breaking and provided an opportunity for advancing knowledge because, while Thomas was unable to ask about deliberations or trials, she remains the only UK researcher to be given access to ‘real’ jurors in order to study rape myths.

Since its publication, Thomas’ research has been framed as the definitive answer to almost 40 years of academic inquiry, which otherwise suggested there was a problem with rape myths among both juries and the wider public (see Dinos et al, 2015; Leverick, 2020). Scottish Legal News quoted Thomas as saying, ‘hardly any jurors believe what are often referred to as widespread myths’ and in the same article, Thomas Ross QC stated that ‘Professor Thomas’ work will hopefully put an end to those arguments [about rape myths] – and allow an evidence-based approach, free of hyperbole’ (Summan, 1 Dec 2020). Similarly, Joshua Rozenberg, presenter of BBC Radio 4’s Law in Action, quoted the study using the headline: ‘Belief by juries in rape myths is a myth’ (Rozenberg, 27 Nov 2020).
Rather than celebrating a new source of important data, we therefore find ourselves delineating the limitations of Thomas’ study in order to contextualise these subsequent inflated claims. Academic research is based on principles that mirror those which a jury (should) use at trial: Each claim must be scrutinised and weighed against the wider evidential landscape. Research is therefore best understood as a mosaic, which together with other studies can build a robust picture of the legal and social world. To suggest that one study should negate decades of largely consistent research, particularly without adequately reviewing this extant literature, demonstrates a lack of understanding about scientific endeavour.

What follows is a summary of four key critiques about Thomas’ (2020) research. There is little detail about Thomas’ methodology, for example how each attitudinal statement was chosen or validated, and some of our critiques might be addressed through greater transparency about the research process. We also recognise that no study can be perfect and so would not dismiss Thomas’ findings simply because they are limited in the way that all socio-legal research is limited. On the contrary, honesty about each study allows us to build the robust evidence-base, free of hyperbole that is so-desired by the legal commentators above.

**Critique One: Survey Design**

Thomas (2020: 1005) described her methodology as ‘an anonymous and voluntary post-verdict survey before leaving court’. The survey appeared to comprise 18 statements, which were said to be based on McMahon and Farmer’s (2011) validated measure, the Subtle Rape Myth Acceptance Scale (footnote 49 of Thomas, 2020). However, when comparing the two surveys, none of the statements adopt the exact same wording and only five address broadly comparable topics. As noted by Chalmers et al (forthcoming), there is also limited repetition of each myth category within Thomas’ survey. Repetition of similar themes can reveal ambivalence that singular statements would not; for example, 93 percent of police officers in Page (2010) said they believed any woman could be raped, yet 19 percent said they were unlikely to believe a victim-survivor in marital rape and 44 percent said they were unlikely to believe a sex worker. Thomas’ added statements do provide useful insight about topics not captured by McMahon and Farmer (2011), for example delayed reporting, male complainants, and demeanour whilst giving evidence. However, this demonstrates that Thomas’ work is exploratory rather than the sole authority on rape myth acceptance and should not be framed as replicating a validated survey.

A perhaps more significant challenge is that the survey appeared to give participants a choice of ‘agree’, ‘disagree’, or ‘unsure’. This is novel, as existing research tends to use five- or seven-point Likert scales to capture the nuance involved in attitude. Whilst a Likert scale can technically contain just three points, Johns (2010) argued that data are less accurate when fewer than five points are used. Further, there is no detail about potential response effects or the format by which surveys were administered. Höhne and Krebs (2018) found that with agree/disagree scales, written surveys tended to have a primacy effect while oral ones had a recency effect. It is therefore difficult to assess validity without more information about the practicalities of survey delivery.

Thomas’ (2020) use of ‘not sure’ is also uncommon in the literature, where standard practice is ‘neither agree nor disagree’ because people often lack clear views on social issues, but tend not to use genuine non-response options (Johns, 2010). Despite this, ‘not sure’ responses comprised a sizeable proportion of Thomas’ data. This could potentially be explained by Johns (2005), who found that the
The midpoint option was used by participants who held attitudes they knew could be socially undesirable on sensitive issues.

Indeed, it is unclear how the risk of social desirability bias was mitigated in the Thomas (2020) survey. Previous research notes that participants may recognise and select the socially acceptable answer even when they do not necessarily hold such a belief (Jann, Krumpal and Wolter, 2019; McMahon and Farmer, 2011). For example, there is evidence that participants can feel pressure to choose certain answers depending on the researcher’s characteristics, perceptions of researcher authority, survey design, and degree of topic sensitivity (Krumpal, 2013). As Munro (2019: 34) argued, the context of Thomas’ participants having just discharged their legal duties ‘is perhaps reason to expect that they might have been inclined to emphasise their compliance with legal tests and deny—consciously or otherwise—biased views that may have confounded their decision-making’.

In rape literature, this was recently tested by Venema (2018), who found that social desirability was a significant influence on a range of rape myth indicators. In the study of 174 police officers, estimated prevalence of false allegations were slightly less susceptible to social desirability than rape myths themselves (Venema, 2018), so it is significant that Thomas’ strongest levels of agreement or uncertainty were found in questions about false allegations. Lüke and Grosche (2018) argued that social desirability bias can supersede demographic influences, which is also significant in light of Thomas’ (2020) finding that demographic characteristics had little impact. Of course, awareness of the socially desirable answer should cause optimism in rape myth research, as it may mean that educational programmes are working. As we argue below, however, it is not the same as believing the socially desirable answer in context.

A final element of concern over survey design is the lack of clarity on analytical strategy. Johns (2010) outlined some ways in which attitudinal surveys can ‘score’ participants across the range of statements, but Thomas (2020: 1000) provided only descriptive statistics of each measure individually, meaning there is no test of internal consistency. The study examined the relationship between attitudes and participant age, gender, religion, ethnicity, and socio-economic group; as well as the court region and type of case served on. No data or statistical analyses were reported for these tests, aside from a vague comment that ‘little to no differences were found between male and female jurors; jurors of different ages; between regions; and between jurors who tried sexual offences cases and those that did not’ (Thomas, 2020: 1004). Even if the relevant statistical tests were used and all assumptions for these tests were met, the phrasing is ambiguous as to whether there were statistically significant differences based on ethnicity and socio-economic group.

Critique Two: Distinction between ‘Real’ Jurors and Other Myth Research Participants

Notwithstanding the limitations above, Thomas’ (2020) study has a key strength in that it sampled ‘real’ jurors within a court setting. This led Thomas to argue that ‘the findings of this research demonstrate clearly for the first time why the views and decisions of real jurors who actually serve on trials can never be replicated by volunteers in mock jury studies... or by those who take part in public opinion polls’ (p.1005). We agree that jury studies in a real court setting are invaluable and should therefore be available to more researchers, but it is difficult to understand how this leads to Thomas’ conclusion that research should only involve ‘real’ jurors.

Firstly, mock jury research tends to have different aims and objectives to those of Thomas, with a strong focus on exploring group dynamics and the process of deliberation. Given the restrictions on
seeking information from ‘real’ jurors about the opinions and arguments set out in the jury room (s.74 of the Criminal Justice and Courts Act 2015), this is only possible with mock trials. Chalmers et al (forthcoming) noted that mock jury research varies in methodology and therefore rigour, but when designed robustly to increase ecological validity then it can provide important insights. Significantly, Thomas’ study is no more ecologically valid than mock jury studies or attitudinal surveys because it is an attitudinal survey rather than naturalistic study of the jury experience. Indeed, Holleman et al (2020) argued against using ‘real-world’ to describe research, because the act of data collection changes the participants’ experience unless it is achieved covertly or via deception. Participating in a survey is not part of routine social behaviour, so there is little use distinguishing between ‘real’ and ‘mock’ jury samples when it comes to ecological validity. A more compelling argument would be to distinguish between ‘real’ and ‘mock’ jury deliberations, but this is not possible in the UK.

At its core, then, Thomas’ (2020) research is a public survey with a sampling strategy that targeted participants who had recently completed jury duty in selected courts. There is no suggestion from Thomas that jury duty influenced participants’ attitudes, as no baseline data were collected before trial to test the impact of the process, and jurors served on a range of cases rather than just sexual offences. Instead, it is argued that ‘real’ juror research is preferable to other myth studies because there is less self-selection bias (Thomas, 2020). It is true that self-selection bias is a significant risk, but ‘to rely exclusively on observational schemes that are free from selection bias is to rule out a vast portion of fruitful social research’ (Winship & Mare, 1992: 328). It is also important to note that sampling from ‘real’ juries carries the same risk of selection bias as other social research, because those invited to participate self-select whether to take part (see Chalmers et al, forthcoming).

Thomas (2020) argued that her research supersedes other myth studies because it had an admirably high 99.7 percent response rate (just one person declined to participate). To put this in context, a review found that typical response rates to organisational surveys are around 40 percent, but can increase to 70 percent in particularly salient settings and when using incentives or personal contacts (Cook et al, 2000). We do not know how the research was explained to jurors, but Baruch and Holtom (2008) suggested that near 100 percent response rates should be justified by demonstrating that participants did not feel pressured to complete the survey. High response rates should therefore be treated with caution as it could indicate that participants found it hard to decline or chose to participate without due consideration of their responses (Roberts & Indermaur, 2008). This is well-established in work on criminal justice, for example prison researchers have long reflected on the need ensure truly informed and freely given consent for participation (see Field, Archer & Bowman, 2019).

Fisher (2013) wrote about the influence of social, cultural, and political contexts on the ‘voluntariness’ of research participation, even where individual researchers did not pressurise recruits. Indeed, a large body of evidence (for example, De Vries, 2004) demonstrates that social context influences both participants’ and researchers’ decision-making, and recruiting participants within an authoritative institution such as court brings unspoken power dynamics (see also Fisher, 2013). This is not an allegation of ethical mismanagement, nor of coercive sampling techniques, but rather is meant to illustrate why Rogelberg and Stanton (2007) argued that a high response rate does not automatically mean higher validity.

**Critique Three: Abstract vs Applied Myths**

A further reason that Thomas (2020) is misled in rejecting previous research is the issue of generalised versus applied statements. Triangulation means that it is helpful for Thomas to examine attitudes
framed in abstract terms, but that it remains equally important for other research to explore attitudes framed using concrete examples. This is because public opinions are known to change depending on whether a question is framed in the abstract versus being applied (see Thielo et al, 2016). For example, Smith et al (2021) found that members of the public did not hold strong views when asked generally about rape complainants’ eligibility for state compensation, but consistently disapproved of the same eligibility rules when given case studies. Similarly, a survey of 891 Canadian police officers found relatively low levels of rape myth endorsement (Page, 2010), before police case file analysis suggested high levels of rape myth support (Shaw et al, 2017). Most recently, Zidenberg et al (2021) identified a difference between the qualitative and quantitative data in a rape myth study, with participants’ (86 students and 82 members of the wider community) receiving low rape myth acceptance scores but then making myth-endorsement comments during interview.

This is a further reminder of the usefulness of mock jury research, as these studies have long found that participants reject rape myth statements in attitudinal surveys and then draw on those same ideas during deliberation (Ellison & Munro, 2010). Similarly, decades of court observation research show that in practice, rape myths are rarely deployed in plain or abstract terms by barristers (Adler, 1987; Lees, 1996; Temkin et al, 2018; Smith, 2018). This includes court observations from 2019, within the timeframe of Thomas’ data collection (Daly, forthcoming).

These observation studies suggest that rape myths are acknowledged by trial counsel in general terms, but then implied in order to scaffold critiques of the complainants’ credibility (Smith, 2018), alongside wider cultural narratives about gender, race, class, sexuality, and disability (Daly, forthcoming). This so-called concurrent validation has been demonstrated in other surveys. For example, the 2017 National Community Attitudes towards Violence Against Women Survey in Australia (Webster et al, 2018) found that beliefs about gender equality were the biggest predictors of rape myth acceptance. Thomas’ (2020) research does not speak to this known interaction of wider prejudice with rape myths (see also Thiara & Roy, 2020), nor the ways in which deliberation about specific cases might create more ambivalence than abstract statements alone.

**Critique Four: Interpretation of the Data**

Finally, we argue that the reporting of Thomas’ findings in some quarters has been misleading. For example, the research was cited in Australia to say that evidence ‘does not support the view that false preconceptions have a significant impact on jury decision-making’ (Queensland Law Reform Commission, 2020: np), but Thomas did not investigate juror decision-making nor impacts on jury verdicts. Additionally, the claim that ‘hardly any jurors believe what are often referred to as widespread myths and stereotypes about rape’ (Thomas, 2020: 1001) belies the data. Table 1 outlines the results from Thomas (2020) and shows positive recognition of, for example, male rape and the fact that sexual offences do not require physical force. These findings are welcome, notwithstanding our concerns about social desirability bias, and could be a signal that public awareness campaigns are working. However, Thomas (2020) also found evidence of relatively high endorsement about intoxication, false allegations, and a complainants’ demeanour in court. One in four people felt it was difficult to identify rape between drunk people, while almost half of participants expected the complainant to be visibly distressed when giving evidence.

Thomas acknowledged the need to address misconceptions about demeanour, but argued that the relatively high agreement with false allegation and intoxication myths were unproblematic because
there is ‘no agreed factual basis or empirical evidence to say definitely what the correct answer is’ (2020:1001). The law accounts for the role of intoxication in both consent and reasonable belief, so the statement ‘if both people are drunk, it’s hard to know if it was really rape’ can have a legally correct answer. However, there are also more ambiguous interpretations of the statement - a symptom of poor survey design. Additionally, whilst specific numbers of false allegations are not known, all evidence suggests they are not likely to be substantial (see commentary in Chalmers et al, forthcoming). Most significantly, Conaghan and Russell (2014) argued that the problem with rape myths does not come from their truth or falsity, but rather their universal and uncritical application (see also Lonsway & Fitzgerald, 1994).

<table>
<thead>
<tr>
<th>Negatively framed statements:</th>
<th>Agree (%)</th>
<th>Not sure (%)</th>
<th>Disagree (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men cannot be raped.</td>
<td>2</td>
<td>6</td>
<td>92</td>
</tr>
<tr>
<td>A rape probably didn’t happen if the victim has no bruises or marks.</td>
<td>3</td>
<td>10</td>
<td>87</td>
</tr>
<tr>
<td>A woman who goes out alone at night puts herself in a position to be raped.</td>
<td>4</td>
<td>9</td>
<td>87</td>
</tr>
<tr>
<td>If a person doesn’t physically fight back, you can’t really say it was a rape.</td>
<td>3</td>
<td>12</td>
<td>85</td>
</tr>
<tr>
<td>A woman who wears provocative clothing puts herself in a position to be raped.</td>
<td>4</td>
<td>13</td>
<td>83</td>
</tr>
<tr>
<td>If a woman sends sexually explicit texts or messages to a man, she should not accuse him of rape later on.</td>
<td>5</td>
<td>18</td>
<td>77</td>
</tr>
<tr>
<td>It is difficult to believe rape allegations that were not reported immediately</td>
<td>7</td>
<td>20</td>
<td>73</td>
</tr>
<tr>
<td>People are more likely to be raped by a stranger than someone they know.</td>
<td>5</td>
<td>31</td>
<td>64</td>
</tr>
<tr>
<td>Children often make up stories about being sexually abused</td>
<td>3</td>
<td>33</td>
<td>64</td>
</tr>
<tr>
<td>Many women who claim they were raped agreed to have sex and then regretted it afterwards</td>
<td>12</td>
<td>47</td>
<td>41</td>
</tr>
<tr>
<td>If both people are drunk, it’s hard to know if it was really rape</td>
<td>25</td>
<td>36</td>
<td>39</td>
</tr>
<tr>
<td>I would expect anyone that was raped to be very emotional when giving evidence in court about the rape.</td>
<td>43</td>
<td>35</td>
<td>22</td>
</tr>
<tr>
<td>Some people will make up allegations of sexual offences against a famous person</td>
<td>47</td>
<td>46</td>
<td>7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Positively framed statements:</th>
<th>Agree (%)</th>
<th>Not sure (%)</th>
<th>Disagree (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>There are good reasons why someone who’s been raped may be reluctant to tell anyone about it or report it to the police.</td>
<td>80</td>
<td>13</td>
<td>7</td>
</tr>
<tr>
<td>Rape within a relationship can take place over a long period of time before any complaint is made</td>
<td>77</td>
<td>19</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Statements not clearly linked to common myths:</th>
<th>Agree (%)</th>
<th>Not sure (%)</th>
<th>Disagree (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>People who make rape allegations are often not believed by the police.</td>
<td>21</td>
<td>55</td>
<td>24</td>
</tr>
<tr>
<td>The consequences of a rape conviction are more serious for a young man than for an older man</td>
<td>6</td>
<td>31</td>
<td>63</td>
</tr>
<tr>
<td>It is a hard thing to do to give evidence in court about a rape</td>
<td>77</td>
<td>20</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 1: Summary of Thomas 2020 findings
Table 1 also shows that a significant minority of participants were uncertain about the myth statements. For example, while only 7 percent of respondents believed delayed reporting to be suspicious, a further 20 percent were unsure. Further study is therefore needed to establish the potential impact on deliberations of myth-endorsing and myth-ambivalent jurors, in comparison to myth-rejecting jurors. Table 2 sets out the increased risk when acknowledging the potential influence of myth-ambivalent jurors in addition to myth-endorsing jurors. Of course, uncertain jurors might be easily swayed by the majority who refute myths; but without testing this, the Thomas study cannot claim to prove a jury containing up to nine myth-endorsing and myth-ambivalent jurors is ‘unbiased’.

<table>
<thead>
<tr>
<th>Statements</th>
<th>Prevalence of myth-endorsing jurors in a 12-person jury</th>
<th>Prevalence of myth-endorsing AND myth-ambivalent jurors in a 12-person jury</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men cannot be raped.</td>
<td>0.24 / jury</td>
<td>0.96 / jury</td>
</tr>
<tr>
<td>A rape probably didn’t happen if the victim has no bruises or marks.</td>
<td>0.36 / jury</td>
<td>1.56 / jury</td>
</tr>
<tr>
<td>A woman who goes out alone at night puts herself in a position to be raped.</td>
<td>0.48 / jury</td>
<td>1.56 / jury</td>
</tr>
<tr>
<td>If a person doesn’t physically fight back, you can’t really say it was a rape.</td>
<td>0.36 / jury</td>
<td>1.80 / jury</td>
</tr>
<tr>
<td>A woman who wears provocative clothing puts herself in a position to be raped.</td>
<td>0.48 / jury</td>
<td>2.04 / jury</td>
</tr>
<tr>
<td>There are good reasons why someone who’s been raped may be reluctant to tell anyone about it or report it to the police.</td>
<td>0.84 / jury</td>
<td>2.40 / jury</td>
</tr>
<tr>
<td>Rape within a relationship can take place over a long period of time before any complaint is made</td>
<td>0.48 / jury</td>
<td>2.76 / jury</td>
</tr>
<tr>
<td>If a woman sends sexually explicit texts or messages to a man, she should not accuse him of rape later on.</td>
<td>0.60 / jury</td>
<td>2.76 / jury</td>
</tr>
<tr>
<td>It is difficult to believe rape allegations that were not reported immediately</td>
<td>0.84 / jury</td>
<td>3.24 / jury</td>
</tr>
<tr>
<td>People are more likely to be raped by a stranger than someone they know.</td>
<td>0.60 / jury</td>
<td>4.32 / jury</td>
</tr>
<tr>
<td>Children often make up stories about being sexually abused</td>
<td>0.36 / jury</td>
<td>4.32 / jury</td>
</tr>
<tr>
<td>Many women who claim they were raped agreed to have sex and then regretted it afterwards</td>
<td>1.44 / jury</td>
<td>7.08 / jury</td>
</tr>
<tr>
<td>If both people are drunk, it’s hard to know if it was really rape.</td>
<td>3.00 / jury</td>
<td>7.32 / jury</td>
</tr>
<tr>
<td>I would expect anyone that was raped to be very emotional when giving evidence in court about the rape.</td>
<td>5.16 / jury</td>
<td>9.36 / jury</td>
</tr>
</tbody>
</table>

Table 2: Estimated prevalence of bias in a jury of 12, based on Thomas 2020 findings

Conclusion

We have outlined four areas of concern regarding Thomas’ (2020) paper in Criminal Law Review, which claimed to disprove years of research on rape myth acceptance. We argue that
the study is a missed opportunity, as it drew upon a unique sampling strategy that is not available to most researchers in the field. This does not mean that Thomas’ work should be dismissed outright, but that it is important to avoid hyperbolic claims about what a single study can reveal.

References


Rozenberg, J. (27 Nov, 2020) ‘Belief by juries in rape myths is a myth according to the only researcher who has asked the jurors themselves’ *A Lawyer Writes*, https://rozenberg.substack.com/p/belief-by-juries-in-rape-myths-is.


---

1. Two previous jury surveys were conducted by Jackson (1996) in Northern Ireland, and Zander and Henderson (1993) in England and Wales (both cited in Chalmers and Leverick’s 2016 outline of UK jury research). Neither of these studies examined sexual offences or rape myths.

2. ‘Repeated’ measures only occurred in relation to five of the myth themes and mostly involved just two broadly related statements. There is no statistical analysis provided as to how these repeat measures were used to validate and test each other. The ‘repeated’ themes were about: (a) injury / physical resistance, (b) risk from the complainant’s choice of clothing / going out alone, (c) whether it is hard to give evidence /
complainants being upset in court, (d) delayed reporting / whether there are ‘good reasons’ it might be hard to report, and (e) false allegations by children, against celebrities, or as a result of ‘regret’.

Some studies do prefer an even scale, to force participants into making decisions, but Thomas’ use of ‘unsure’ means this was not her reason for eschewing the conventional measures.