

Mock juror sensitivity to forensic evidence in drug facilitated sexual assaults

Regina A. Schuller ^{a,*}, Alison Ryan ^b, Daniel Krauss ^b, Gwen Jenkins ^a

^a Department of Psychology, York University, Toronto, ON, Canada

^b Department of Psychology, Claremont McKenna College, Claremont, CA, United States

ARTICLE INFO

Available online 19 February 2013

Keywords:

Sexual assault

Rape

Drugs

Negative test results

ABSTRACT

Mock jurors' reactions to variations in the quality of toxicological evidence regarding the presence of drugs in a sexual assault trial were examined. In Study 1, participants received a trial summary in which a negative test result, a negative test result plus expert testimony, or no test result was presented. The time taken by the complainant to report the alleged sexual assault was manipulated. The negative test result influenced participants' judgments, but this effect was minimized by the presence of expert testimony. The complainant's delay in reporting had little impact on judgments. In Study 2, complainant time to report was again manipulated along with the outcome of the test result (negative finding and no result). Results revealed that men were less conviction prone when the negative test result was obtained early as opposed to late. In contrast, when the test result was unavailable, men were more conviction prone when the complainant reported late as oppose to early.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

In cases of drug facilitated sexual assault (DFSA), negative forensic findings for the presence of drugs (i.e., no drugs found in the complainant's blood) can present significant challenges to a complainant's claim. Such was the case in *Regina v. Alouache*.¹ In this case, the defense introduced a negative forensic report to support the defendant's claim that he had not drugged and sexually assaulted the complainant. The prosecution, in turn, moved to introduce expert testimony to contextualize the negative forensic findings. The defense countered, arguing that the introduction of the expert testimony "would be highly prejudicial to the defense."² In its consideration of the case, the Ontario Court of Appeal upheld the prosecution's request to introduce the expert testimony. The present research is conducted against this backdrop. By varying both the quality of the forensic test result, as well as the presence of expert testimony contextualizing the evidence, the current research explores mock jurors' sensitivity to variations in the probative value of forensic evidence, their receptivity to expert testimony, and how they weigh expert testimony in reaching their decisions.

1.1. DFSA and the challenges it presents at trial

With terms such as 'date rape drugs' and 'drug facilitated sexual assault' now in the common vernacular, as well as mounting interest

from the scientific community, attention has been drawn to the contributory role of drugs and alcohol in sexual assault (e.g., Du Mont et al., 2010; Hindmarch, ElSohly, Gambles, & Salomere, 2001; Hindmarch & Brinkmann, 1999; Olszewski, 2009). The low rates of reporting, prosecution, and conviction that characterize sexual assault (e.g., Backman, 1998; see generally Temkin & Krahé, 2008), is likely even more pronounced in cases in which the victim has been surreptitiously drugged. The drugs used are fast-acting; within 15 min of ingestion (Wells, 2001) victims may experience distortions in perception, confusion, inhibition, along with an inability to offer any resistance, followed by rapid sedation and loss of consciousness (Freese, Miotto, & Reback, 2002; LeBeau et al., 1999; Wells, 2001). With victims unable to clearly recollect the circumstances surrounding the events, they may initially downplay or be unaware of the seriousness of what occurred (Fitzgerald & Riley, 2000). Accordingly, they may also delay reporting their victimization, if they report it at all (McGregor, Wiebe, Marion, & Livingstone, 2000). Indeed, in comparison to other sexual assaults, victims identified in a sample of suspected DFSA cases had longer time delays before presenting to hospital, had sustained less physical injury, and were less likely to involve the police (McGregor, Lipowska, Shah, Du Mont, & Siato, 2003), all variables that have been found to be negatively correlated with the believability of the complainant's account (Frazier & Haney, 1996; McGregor, Du Mont, & Myhr, 2002).

To further complicate matters, the victim's reluctance to report the crime can affect not only the plausibility of her claim, but it can also result in a time delay in testing for the presence of drugs in the woman's system (Hurley, Parker, & Wells, 2006; Wells, 2001), which can have serious ramifications for the accuracy of the forensic test result. Moreover, with date rape drugs often consumed alongside alcohol, their side effects may closely resemble signs of heavy alcohol intoxication (Scott-Ham & Burton, 2005), a variable that has been

* Corresponding author at: Department of Psychology, York University, 4700 Keele Street, Toronto, Ontario, Canada M3J 1P3. Tel.: +1 416 736 2100x33190; fax: +1 416 736 5814. E-mail address: schuller@yorku.ca (R.A. Schuller).

¹ [2003] CanLII 9494 (ON S.C.).

² *Supra* note 4 at Para. 19.

consistently identified as influencing third parties' perceptions and treatment of victims (e.g., Schuller & Stewart, 2000; Schuller & Wall, 1998) with law enforcement agents more dismissive of a complainant if the alleged victim has consumed alcohol (Dorandeu et al., 2006; Jordan, 2004; LeBeau et al., 1999). In turn, due to their dismissive attitudes, law enforcement officials may not stress the necessity and urgency of forensic drug testing (Hurley et al., 2006).

Lack of evidence of physical injury and inconsistencies and/or gaps in the complainant's account of the assault can render forensic evidence in a DFSA trial of critical importance. Given its import, it is imperative that jurors recognize the strengths and limitations of drug testing evidence. For instance, such evidence may be particularly probative when there is a short time delay between the alleged ingestion and the testing for drugs (although even with a short delay there are drugs that may not be effectively detected), but is far less probative when the time delay to testing is greater than the half-life of the date rape drugs tested (e.g., GHB is completely undetectable 12 h after ingestion, Olszewski, 2009; Scott-Ham & Burton, 2005).

1.2. The impact of negative forensic results in a case of DFSA

To assess the impact of a negative forensic report on mock jurors' judgments in a DFSA trial, as well as the impact that expert testimony contextualizing the negative report may have on jurors' judgments, Jenkins and Schuller (2007) conducted a juror simulation study. For some of their participants, the results of a negative forensic report (no drugs found in the complainant's blood/urine) were introduced into evidence by the defense. For another group, who also received the negative forensic report, additional testimony from the prosecution was provided by an expert witness who outlined the variety of factors that could contribute to a negative test result. For a final group, the toxicological screening was not introduced into evidence (control). Comparison across these conditions, revealed that, compared to the control condition, the presentation of the negative forensic report in the absence of expert testimony produced greater verdict leniency and evaluations more favorable to the defense. The information provided by the expert, however, negated the impact of the negative forensic report, with participants in this condition rendering judgments similar to those in the control condition. In short, when accompanied by expert testimony, the decision-makers now gave less weight to the negative test result.

How should we interpret these findings? Did the presence of the expert testimony result in the mock jurors being more accurate in their evaluation of the forensic evidence? Possibly, but it is also possible that the toxicological evidence was not given its due weight. Although the expert in Jenkins and Schuller provided information about the testing and the variables that can affect the likelihood of detection of drugs if they were present, variables impacting the sensitivity of the testing were not manipulated. The probative value of the negative forensic evidence should depend upon the accuracy of the negative test result. For example, if the screening is conducted within a reasonable time frame following the alleged ingestion, the negative finding should be evaluated as more accurate and as a result, should be more persuasive, than had the testing been delayed.

At the most basic level, in order for jurors to use expert testimony, they must evaluate and weigh the information the expert provides and appropriately apply that new information to the case at hand. When expert testimony improves juror understanding and application of the factors that are critical to evaluation of the evidence, its impact has been referred to as "sensitizing" (Crowley, O'Callaghan, & Ball, 1994; Cutler, Dexter, & Penrod, 1989; Kovera, McAuliff, & Hebert, 1999; Wells, 1986). In contrast to sensitization, however, another form of impact has also been identified. Expert testimony can also produce greater skepticism about other evidence, whereby jurors favor the side that employs the expert regardless of the relevance of the information conveyed by the expert to the evidence at

hand (Buck, London, & Wright, 2011; Cutler et al., 1989; Krauss & Sales, 2001; Levett & Kovera, 2008). Whether expert testimony will result in juror sensitization or skepticism in a DFSA case has not yet been effectively tested, and thus, the current research was designed to address this issue.

1.3. Overview of current research

In the present article, we describe the results of two studies that build upon and extend Jenkins and Schuller's initial exploration of the impact of negative forensic evidence in the context of DFSA by: (1) investigating the impact of a negative forensic report on mock jurors' decisions; (2) examining how the quality of the forensic report impacts their decisions; and (3) testing whether the presence of the expert testimony sensitizes the jurors to the factors that influence the quality (hence validity) of the forensic finding.

2. Study 1 – Method

In Study 1, participants read a simulated sexual assault trial that involved an allegation of DFSA. Judgments of participants who were not provided negative toxicological test results were contrasted with those of participants who did receive the negative test results. To assess mock jurors' sensitivity to the quality of the forensic evidence, the time frame within which the complainant initiated testing was varied. For some of the participants, the time delay was short (within 5 h of the alleged sexual assault), thus resulting in a toxicological test result that would have high diagnostic value. In a long delay condition, the complainant reported more than 24 h after the alleged assault resulting in a test with more questionable accuracy (low diagnostic value). When the negative test result was presented, half of the participants were provided with expert testimony contextualizing the negative forensic report and the other half were not.

Based upon previous research (Jenkins & Schuller, 2007), it is predicted that jurors will be less likely to believe the complainant's claim and will evidence more leniency towards the defendant when the forensic report is presented in the absence of expert testimony. Secondly, although it is not clear what effects the time delay will have on verdicts, a growing body of research suggests that through expert testimony, jurors can become more sensitive to variations in scientific evidence (e.g., Buck et al., 2011). As a result, it is hypothesized that the expert testimony will interact with a time delay in reporting, rendering the mock jurors more sensitive to the effects of the time delay manipulation on the accuracy or validity of the test result. Thus, this should result in fewer convictions in the short time delay condition but only when it is paired with expert testimony. Additionally juror characteristics, such as gender, have been shown to influence juror decision-making in cases of sexual assault (Schutte & Hosch, 1997). In line with previous research, it is predicted that men will be less likely to render guilty verdicts and will rate the complainant as less credible than women.

2.1. Participants

Participants were 208 undergraduates (87 men, 115 women, and 6 unidentified), ranging in age from 18 to 23 ($M_{\text{age}} = 19.46$, $SD = 1.11$) recruited from a consortium of small liberal arts colleges in Southern California. They received course credit for their participation.

Twenty participants were dropped prior to the data analyses because they indicated that they had been sexually assaulted ($n = 12$) or chose not to indicate whether they had been sexually assaulted ($n = 8$). Two participants who indicated that they were not jury eligible, as well as 8 who provided incomplete data, were also excluded, leaving 180 participants in the sample (100 women and 80 men).³

³ In both this as well as the subsequent study, analyses that retained these participants produced a similar pattern of results.

2.2. Procedure

Upon arrival, participants were randomly assigned to a trial condition. They were informed that they would be reading a summary of a criminal trial and were asked to assume the role of a juror for the duration of the study. After reading the trial transcript, they completed the dependent measures. After which, they were debriefed and thanked for their participation.

2.3. Stimulus case

Participants were presented with a criminal trial transcript based on materials developed by Jenkins and Schuller (2007), in which the issue at trial involved an allegation of DFSA. The trial presentation followed the format of an actual trial, with opening judicial remarks, prosecution and defense attorneys' opening statements, direct and cross-examination of witnesses, closing arguments by both the defense and prosecution, and final judicial instructions (approximately 25 pages in length). A summary of the basic facts follows:

Upon encountering each other on the metro one evening, the two decide to go to a nearby pub (they recognized each other from high school). Although the two had different accounts of who initiated plans to leave the pub, the next part of the evening takes place at the complainant's apartment. The two continue talking and looking through old high school photos, which the complainant had retrieved from her bedroom, leaving the defendant alone for a few minutes. The complainant testified that later while on the couch, the defendant kissed her and, according to her, she kissed him back initially but then resisted his advances. The complainant testified that, upon finishing her drink, she began to feel very weak and her body was unresponsive. She passed out, waking alone hours later to find signs that sexual intercourse had occurred (e.g., a broken zipper and sticky colorless fluid in her underwear) but had no recollection of the event. Suspecting that she was sexually assaulted, the complainant went to a hospital and was examined by a Sexual Assault Nurse Examiner. Physical evidence confirmed that intercourse had occurred, but there were no signs of physical injury. In contrast to this version of events, the defendant testified that although he did not dispute that sexual intercourse had occurred, he alleged that the sex was consensual and that the complainant became angry when he refused to continue the relationship with her.

Within these basic facts, six different versions of the trial were produced by systematically varying two independent variables: Complainant Time Delay (complainant reported to hospital approximately 5 or 24 h following the alleged assault) and Forensic Evidence (no forensic report provided; negative test result provided; negative test result plus expert testimony). That is, for one-third of participants, no forensic report was provided, for another third, the defense presented a negative forensic report (i.e., no drugs found), and for the remaining third, in addition to the negative forensic report, the prosecution presented expert testimony from a toxicologist to contextualize the negative forensic results. The expert testimony, loosely based on the testimony provided in the case *Alouache*,⁴ gave an account of the types of drugs that are commonly used in DFSAs, the drugs that the hospital usually tests for, as well as the drugs that were tested for in the complainant's case (alcohol, barbiturates, marijuana, cocaine, benzodiazepines, GHB, ecstasy, amphetamines, and antidepressants). The expert also testified that there are several factors that can cause erroneous negative findings including the time delay in testing and the short half-life of the drugs. He also noted that 10% of DFSA drugs could not be detected given current methods.

2.4. Dependent measures

2.4.1. Verdict and guilt assessments

Participants were asked to render a verdict (guilty, not guilty), and to rate the likelihood of the defendant's guilt on a scale from 0% to 100% with increments of 10%.

2.4.2. Case evaluations

Participants responded to a series of 10 items that assessed their acceptance of the sexual assault claim (e.g., complainant fabricated the claim (reverse coded) and complainant was drugged). These items were summed and averaged to form a composite measure of the validity of the claim ($\alpha = .87$), with higher scores denoting a stronger belief in the claim. Additionally, using 7-point scales, participants rated the credibility of both the complainant and the defendant (1 = *not at all* to 7 = *extremely*).

2.4.3. Perception of forensic evidence

Those provided with the forensic report rated the extent to which the report was convincing, influenced their verdict, their acceptance of the complainant's claim, and their acceptance of the defendant's claim (reverse coded). These ratings were subsequently summed and averaged to form a composite measure of the persuasiveness of the report ($\alpha = .75$), with higher scores indicating greater perceived persuasiveness.

2.4.4. Manipulation checks

Participants indicated whether a lab report detailing the results of the complainant's urine and blood tests was presented (yes, no), and, if they answered in the affirmative, whether drugs had been detected (yes, no). Participants also responded to a multiple choice question asking the length of the time the complainant took to report to the hospital with 5, 10, or 24 h as possible answers.

3. Results — Study 1

3.1. Manipulation checks

All but one participant in the conditions in which the negative forensic report had been presented, correctly noted its presence and identified that the results were negative. In terms of how long following the alleged assault the complainant took to report to the hospital, 90.2% indicated the correct time delay (either 5 h or 24 h).¹

3.2. Verdicts

A binary logistic regression was conducted with verdict as the dependent variable and forensic evidence (negative report alone, negative report accompanied by expert testimony, and no report/no expert testimony (all variables dummy coded)), time delay (5 h, 24 h), gender (male, female), as well as their subsequent interactions, all entered as the independent variables. This revealed a significant main effect for forensic evidence, $B = -.43$, $SE = .213$, Wald's χ^2 (1, $N = 183$) = 4.07, $p = .04$, such that those that received the negative forensic report unaccompanied by expert testimony were less likely to find the defendant guilty (25% of the participants) as compared to the other two groups (i.e., 32% in the negative forensic report with expert testimony and 37% in the no report/no expert condition). Additionally, participant gender was a significant predictor of guilt, with women compared to men more likely to render a verdict of guilt, $B = 1.64$, $SE = .388$, Wald's χ^2 (1, $N = 183$) = 19.03, $p = .0001$ (women 45%, men 15%).

3.3. Case judgments

Ratings of the case judgments (i.e., probability of guilt, defendant credibility, complaint credibility, validity of sexual assault claim) were

⁴ *Supra* note 4.

analyzed via a 3 (forensic evidence) by 2 (time delay) by 2 (participant gender) MANOVA. This analysis revealed a significant main effect for gender, Wilks' $\lambda = 6.24$, $p < .001$, and marginally significant main effects for forensic evidence, Wilks' $\lambda = 1.89$, $p = .061$, and time delay, Wilks' $\lambda = 2.28$, $p = .063$. No significant interactions were found. To examine the impact of these multivariate effects, univariate analyses were examined and are described below.

The univariate results revealed main effects for gender across all four of the judgments: defendant's probability of guilt, $F(1,168) = 20.58$, $p < .001$, $\eta^2 = .11$; validity of sexual assault claim, $F(1,168) = 23.57$, $p < .001$, $\eta^2 = .12$; defendant credibility, $F(1,168) = 8.92$, $p = .003$, $\eta^2 = .05$, and complainant credibility, $F(1,168) = 8.45$, $p = .004$, $\eta^2 = .05$. As the means in Table 1 indicate, women, in comparison to men, rated the defendant's probability of guilt higher and were more accepting of the claim. They also found the defendant less, and the complaint more, credible than did men.

A main effect was also found for forensic evidence on one of the dependent measures, probability of defendant's guilt, $F(2,168) = 5.94$, $p = .003$, $\eta^2 = .07$. Those provided with a negative test result when it was unaccompanied by expert testimony rated the defendant's probability of guilt lower ($M = 47.74$, $SD = 21.53$) compared to those provided with the expert testimony ($M = 57.14$, $SD = 24.74$), as well as those in the no report/no expert condition ($M = 60.71$, $SD = 20.17$).

With respect to ratings of defendant credibility, a main effect of time delay was found, $F(1,168) = 5.85$, $p = .017$, $\eta^2 = .03$, with those in the early delay condition rating the defendant less credible ($M = 3.97$, $SD = 1.22$) than those in the late time delay condition ($M = 4.26$, $SD = 1.24$).

3.4. Forensic evidence

A 2 (gender) by 2 (time delay) by 2 (forensic report accompanied by expert testimony or forensic report unaccompanied by expert testimony) ANOVA on participants' evaluations of the forensic evidence produced a significant main effect for forensic evidence only, $F(1,114) = 7.82$, $p = .006$, $\eta^2 = .06$. When the negative forensic report was unaccompanied by expert testimony, participants rated the report as more persuasive ($M = 5.01$, $SD = 1.16$), compared to when it was accompanied by the expert testimony ($M = 4.36$, $SD = 1.20$). Notably, the time delay of the test did not influence juror judgments.

4. Study 2 – Method

As in Study 1, participants read a simulated sexual assault trial involving an allegation of DFSA. A negative forensic report (no drugs found) was again presented, with the complainant's reporting delay manipulated. The testimony provided by the expert explained the factors that can result in false negative results, speaking directly to the accuracy of the test results and the factors that can cause

erroneous negative results. Specifically, the expert noted that results were less valid if the specimen had not been obtained within 12 h of the drug's administration. Given that the mock jurors in Study 1 rated the defendant more credible the longer the complainant delayed going to the hospital (24 h condition), a shorter delay condition that would still evidence problems with detection was tested. As evidenced by Study 1, delays in reporting can arouse suspicions in regards to the victim's credibility. Thus, the impact of complainant delay in reporting, unencumbered by the negative test result, was assessed by including an additional condition. Specifically, this condition provided no test results due to equipment malfunction. Although such evidence is unlikely to be introduced at trial, this permits an assessment of the time delay of the complainant's reporting to the hospital (early vs. late) without it being tainted by the test result information.

On the basis of Cutler et al.'s (1989) explanation of juror sensitivity, it was expected that jurors would be able to utilize the information provided in the expert testimony, integrating the information as it applied to the specific case. If the mock jurors are receptive and sensitive to the information conveyed by the expert, when the complainant reports early (as opposed to late), thereby producing more valid test results, a drop in guilty verdicts should result. The impact of the complainant's delay in reporting, unencumbered by the negative test finding, should not occur when the forensic analyses failed to produce test results (i.e., malfunction). In this condition, the shorter delay should lead to an increase in guilt judgments as the believability of the complainant's claim should be viewed more favorably.

4.1. Participants

One hundred and thirteen (51 men, 61 women, and 1 unidentified; $M_{\text{age}} = 20.08$, $SD = 3.64$) undergraduates, recruited from the research participant pool at a large Canadian university participated in exchange for partial course credit. Participants were randomly assigned to one of the possible trial conditions, then read the trial summary and completed the dependent measures. They were then debriefed and thanked for their participation. As in Study 1, participants who indicated that they had been sexually assaulted were excluded from the analyses ($n = 7$).

4.2. Materials

As in Study 1, the stimulus presentation followed the format of an actual trial. The case facts were slightly modified although still loosely based on the materials developed and used by Jenkins and Schuller (2007). The basic facts of the case are summarized below.

The complainant met the defendant at a bar after work while she was waiting for a girlfriend to arrive. Upon recognizing each other from their old high school, the defendant asked to join her. They had a couple of drinks and reminisced about old times. At one point, the defendant offered to buy another round of drinks and went up to the bar, bringing back two beers. He then asked her to go outside while he had a smoke. While outside, the defendant started kissing her. She testified that she let him know that things were moving too fast but beyond this point she had little recall; she started to feel very weak ('my body felt like it was slipping away from me') and passed out. The next thing she remembered was waking up in a nearby park, with her clothes in disarray. The complainant's girlfriend testifies that she found her friend hours later: outside, confused, disoriented, and her clothes disheveled. The friend testified that she accompanied her friend to a hospital. Physical evidence confirmed that intercourse had occurred, but that there were no signs of physical injury. While the defense did not deny that intercourse had occurred, the defendant claimed that the complainant seduced him, and led him to the park where they consensually had intercourse.

Table 1
Study 1: Summary of univariate analyses and gender effects.

	Gender (n)	Means (standard deviation)
Probability of defendant's guilt*** (out of 100%)	Men (80) Women (100)	47.12 (22.95) 61.00 (21.00)
Validity of sexual assault claim*** (1–7 scale)	Men (80) Women (100)	3.81 (1.00) 4.52 (1.02)
Complainant's credibility** (1–7 scale)	Men (80) Women (100)	4.14 (1.34) 4.57 (1.22)
Defendant's credibility** (1–7 scale)	Men (80) Women (100)	4.35 (1.14) 3.92 (1.28)

** $p < .01$.

*** $p < .001$.

Employing a 2×2 factorial design, with one variable varying the time delay between the alleged sexual assault and the complainant's reporting of the assault (5, 12 h) and the other variable varying the outcome of the test result (negative test result, no test result due to equipment failure), the two independent variables were systematically varied to produce four different versions of the case. With respect to the time delay, those in early delay condition heard testimony that the complainant immediately told her friend that she suspected she was drugged and preceded to the hospital where she was examined, while in the late delay condition, the complainant did not disclose her suspicions to her friend until the next morning, at which point they preceded to the hospital. Thus, the blood and urine samples were obtained within either 5 or approximately 12 h following the alleged assault. Within each of these conditions, half the participants were informed that the result of the drug test was negative (no drugs found), while the remaining were informed that no test result was available due to the equipment malfunctioning.

4.3. Dependent measures

Similar measures to those obtained in Study 1 were collected: manipulation checks, dichotomous measure of verdict, probability of guilt (0–100%), defendant and complainant credibility, and a composite measure of the validity of the sexual assault ($\alpha = .89$). For those who received a negative forensic report, composite measures of the persuasiveness of the report ($\alpha = .88$) and the helpfulness of the expert testimony (i.e., persuasive, informative, helpful, convincing, $\alpha = .87$) were obtained.

5. Results — Study 2

5.1. Manipulation checks

All but seven participants correctly recalled the outcome of the toxicological screen (i.e., negative test result, no results available). Additionally, all those in the early time delay condition correctly reported that the complainant reported to the hospital within 5 h, and all participants in the late time frame condition indicated that she reported approximately 12 or more hours following the alleged assault.

5.2. Verdicts

A hierarchical binary logistic regression was conducted to determine a logit model in which verdict formed the dependent variable and time delay (early, late), test outcome (negative test result, no finding due to equipment failure), and participant gender (male, female) were included as predictor variables on the first step. The two and three way interactions involving these variables were included on subsequent steps. While the main and two way interactions were not found to be significant, the three way interaction was significant, $B = 5.77$, $SE = 1.82$, Wald's $\chi^2(1, N = 183) = 10.10$, $p = .001$.

As the verdicts displayed in Fig. 1 suggest, for male participants, when the toxicological screen resulted in a negative test result, they were less likely to render a guilty verdict when the time delay was early as opposed to late. The reverse pattern was found for men when the test result was unavailable (equipment malfunction); men were then more likely to render a verdict of guilty if the complainant had reported early as opposed to late.

5.3. Case judgments

An initial three-way (Time Delay \times Test Outcome \times Gender) MANOVA was conducted on participants' case judgments (i.e., probability of guilt, validity of sexual assault, complainant/defendant

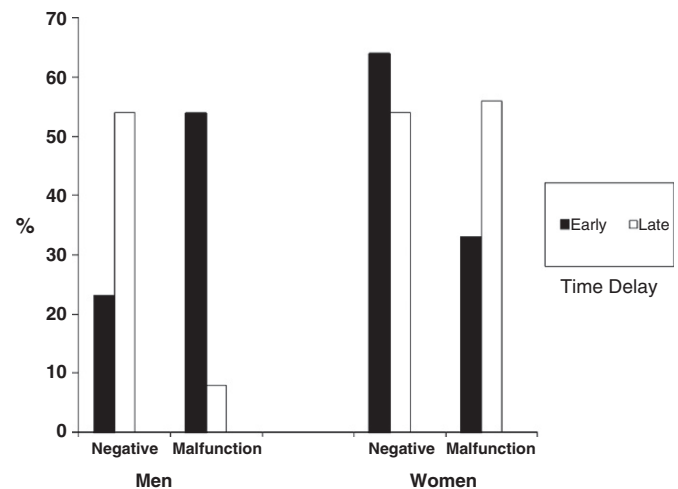


Fig. 1. Verdicts of guilt as a function of time delay and test outcome by gender.

credibility). This analysis revealed a significant main effect for gender, Wilks' $\lambda = 2.51$, $p = .047$, as well as a two way interaction involving time delay by test outcome, Wilks' $\lambda = 2.91$, $p = .026$. These effects were qualified, however, by a significant three way interaction involving these variables, Wilks' $\lambda = 3.80$, $p = .007$. To examine the impact of the significant multivariate effects, univariate analyses were examined.

At the univariate level, a main effect for gender was revealed on only one of the measures (probability of guilt, $F(1,97) = 8.42$, $p = .005$, $\eta^2 = .08$) and a significant two-way interaction involving time delay by test outcome was uncovered on one of the other case judgments (validity of sexual assault claim, $F(1,97) = 9.83$, $p = .003$, $\eta^2 = .08$). Significant three-way interactions, however, were revealed across all four of the case judgments (see Table 2, first column). The interaction was decomposed into two-way outcome by time delay simple interactions at each level of participant gender. Across all four measures, these analyses revealed that the two-way simple interaction was significant for men, but not for women (see Table 2, final column).

Simple effects performed on the interaction for the male participants revealed that, when the test result was negative, men rated the defendant as less guilty when the toxicological sample was obtained early as opposed to late (see Table 3). The obverse pattern, however, was found for men when no test result was provided (equipment malfunction); the early delay condition now received higher ratings of guilt in comparison to the late delay condition. A similar pattern emerged across the remaining judgments. When the test result was negative, men rated the claim of sexual assault less valid, the defendant more credible, but the complainant less credible, when the test was conducted early as opposed to late. The reverse pattern occurred when no test result could be provided; it was now the early test condition that resulted in less favorable judgments for the defendant.

5.4. Evaluations of forensic report and expert testimony

While rating of the persuasiveness of the report did not vary as a function of gender or the time delay manipulation, ratings of the expert testimony produced a main effect for gender, $F(1,49) = 5.56$, $p = .021$, with men ($M = 5.32$, $SD = 1.15$), compared to women ($M = 4.54$, $SD = 4.53$), rating the expert testimony more helpful.

6. Discussion and directions for future research

In Study 1, contrary to our predictions, the mock jurors seemed relatively insensitive to how reporting time might affect the quality

Table 2

Study 2: Univariate analyses for significant three-way and simple two-way interactions.

	Three-way interaction	η^2	Gender	Two-way simple interaction
Probability of defendant's guilt	$F(1,97) = 7.78^{**}$.09	Men	$F(1,97) = 10.57^{**}$
			Women	$F(1,97) = 1.12$
Validity of sexual assault claim	$F(1,98) = 14.73^{***}$.13	Men	$F(1,98) = 22.36^{***}$
			Women	$F(1,98) = .32$
Defendant's credibility	$F(1,98) = 10.12^{***}$.10	Men	$F(1,98) = 9.52^{**}$
			Women	$F(1,98) = 1.92$
Complainant's credibility	$F(1,97) = 4.74^*$.05	Men	$F(1,97) = 7.65^{**}$
			Women	$F(1,97) = .07$

Note. Three-way interaction involves participant gender, time delay, and outcome of test. Two-way simple interactions involve test outcome \times time delay at each level of participant gender.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

of the report. Results suggest that participants did not appear to realize that a 5 h versus 24 h delay in reporting was likely to result in a more valid diagnostic test and this should be given more weight in their decision making. Although the expert did mention this problem, it is possible that mock jurors' lack of sensitivity to the timing variable may be the result of the expert's testimony not drawing sufficient attention to the relationship between reporting time and the scientific veracity of the report. Consistent with finding from other jury-decision making contexts (e.g., Kovera, Gresham, Borgida, Gray, & Regan, 1997; Levett & Kovera, 2008), jurors may not be able to appropriately weigh scientific evidence unless they are motivated to do so and the expert draws a direct, focused link to the problems with the evidence. Support for this can be found in the results of Study 2, in which the expert more clearly articulated the impact that the time delay would have on the validity of the test.

The findings of the second study seem to demonstrate some mock juror sensitivity to the time delay manipulation, but only for the male participants. Across all measures, for male participants, the time delay manipulation interacted with the test outcome. When the toxicological analysis produced a negative test result, the sooner the complainant reported to the hospital, hence the earlier the blood/urine sample was collected and the more reliable the negative test result, the less believing men were of the claim: less guilt prone, more believing of the defendant, and less believing of the complainant and the validity of the sexual assault claim. When provided with a negative test result, men also found the expert testimony more persuasive than did women, suggesting that it may have sensitized the male jurors to the factors (i.e., time delay of test) that should inform their evaluation of the test result. When the toxicological screening failed to produce any results, however, the timing manipulation had the opposite effect for men. They were now more believing of the sexual assault when the complainant reported to the hospital early as opposed to late.

Taken together, the two studies provide considerable insight into the impact that negative forensic test results can have at trial, and

highlight the potential difficulties that jurors face when interpreting and applying this evidence. In Study 1, and consistent with previous research (Jenkins & Schuller, 2007; Schutte & Hosch, 1997), gender was found to be a strong predictor of mock jurors' judgments and verdicts. Women were more likely than men to render guilty verdicts and view the validity of the claim of the sexual assault as believable, rating the defendant less and the complainant more credible than men. As in Jenkins and Schuller (2007), the presence of a negative blood/urine screen did influence the mock jurors, but this effect was minimized with the presence of the expert testimony. As predicted, the presence of expert testimony detailing the weaknesses of a negative forensic report had a significant impact on verdicts and guilt assessments. Participants assigned to the no negative test result condition (hence no expert testimony) estimated the defendant's guilt to be the highest, followed by those who received the negative forensic report that was accompanied by expert testimony. When the report was presented in the absence of expert testimony, however, far fewer guilty verdicts were rendered. The effects in Study 1, however, were not moderated by the reporting delay manipulation and provided no support for mock juror sensitivity to a variable that is of legal relevance (i.e., probative value) to the evidence (i.e., the validity of the negative forensic test result).

Results of Study 2, in which the expert testimony more explicitly articulated the limitations of a negative forensic report, provide some support for juror sensitivity to the relationship between the complainant's delay in reporting and its impact on the validity of the test result. However, these results were confined to only the male participants; women were uninfluenced by the complainant's delay in reporting, regardless of whether or not a negative test result was provided. When the forensic analyses resulted in a negative test result (no drugs found), men were less conviction prone when the negative test result was more likely to be valid (i.e., the complainant reported early as opposed to late). This pattern was not evidenced, however, when the forensic analyses failed to produce test results. The complainant's delay in reporting was now associated with greater skepticism of the validity of the sexual assault.

Although in the case of a negative test result, the impact of the timing manipulation suggests sensitivity to a legally relevant factor in men's decision making, its impact when no test result was provided is troubling as complainants' delay in reporting has little bearing on the veracity of a their claim. As previous research has repeatedly demonstrated (Spohn & Spears, 1996; see Temkin, 2010), the longer it takes a rape victim to report the crime, the less believing others (e.g., police, prosecution, and jurors) are of the validity of the sexual assault. These results are consistent with the growing body of work documenting the influence of people's expectations regarding victim behavior following sexual assault (e.g., Temkin, 2010).

It is possible that the differences found in Study 2 reflect important gender differences in the way in which men and women approach the evidence in sexual assault trials. That is, might these differences reflect

Table 3

Case judgments decomposing the two-way interaction by gender of participant.

	Negative test result (n)		Malfunction (no test result)	
	Early (13)	Late (13)	Early (13)	Late (12)
Men				
Probability of guilt	36.15	58.39	68.31	42.92
Validity of sexual assault claim	3.50	4.85	5.30	3.56
Defendant credibility	4.31	3.46	2.31	4.08
Complainant credibility	3.42	4.69	4.88	3.67
Women				
Probability of guilt	78.38	58.85	69.00	58.94
Validity of sexual assault claim	4.99	4.55	4.72	4.63
Defendant credibility	2.86	3.54	3.58	3.12
Complainant credibility	4.36	3.92	4.33	4.12

differential cognitive biases in seeking and evaluating evidence? Consistent with such an interpretation, men found the expert testimony more persuasive than did women. In contrast, women, who may see themselves as more similar to the complainant, may be more supportive of evidence that points to the defendant's guilt. As such, the time delay may have a weaker impact on female jurors because they are aware that delays in reporting are not indicative of guilt by itself. When this is paired with the negative test result, it is possible that female jurors are still hesitant to discount the possibility that even the early more valid test failed to detect the drug because the expert also noted that certain drugs may be undetectable even in the shorter time frame. At this point, these explanations are purely speculative and await future research. Given that the complainant's delay in reporting has little bearing on the veracity of her claim (i.e., an extra-legal variable influencing decisions), understanding the differential impact that the timing manipulation had on men and women jurors will be of critical importance.

In closing, caveats of the current research must be noted. As with all research conducted on college students, the generalizability of the results can be called into question. Although few differences in decision-making between college students and nonstudents have been found (Bornstein, 1999), the need for replication with more representative samples is warranted (Wiener, Krauss, & Lieberman, 2011). Perhaps of greater significance, is the failure of the current research to include juror deliberations, which given the pervasiveness of gender differences in this domain, may take on considerable importance.

Finally, replication using different and varied trial scenarios is also warranted. The case after which the stimulus materials in the current study were modeled was based on an alleged surreptitious drugging. Horvath and Brown's (2007) examination of rape cases that were reported to the police in which the victim was under the influence of drugs/alcohol revealed few cases in which drugs had been administered surreptitiously. Rather the majority were more likely to be characterized by circumstances in which the victim's alcohol use, particularly alcohol combined with various other drugs, was voluntary. This is likely to be of considerable relevance in the adjudication of sexual assault as is highlighted by the recent work of Girard and Senn (2008) in which the voluntariness of administration (voluntary and involuntary) and substance employed (e.g., GHB and alcohol) was found to have a powerful impact on decision makers' attributions of blame and responsibility. Additionally, future research should also consider the effects of race and other victim and defendant characteristics (e.g., a male-on-male sexual assault) on our juror findings. These factors may also have a significant impact on decisions.

While our primary goal in this research was to investigate the impact of negative forensic evidence on decision-makers in DFSA, the findings have broader implications for decision makers' evaluations of scientific evidence more generally. Examining other forms of evidence, Kovera et al. have investigated the ability of various decision-makers (i.e., judges, attorneys) to detect basic internal threats to the validity of expert psychological evidence. Not only did judges (i.e., the gatekeepers to the testimony's admissibility) appear to have difficulty detecting methodologically flawed expert psychological evidence used by the expert (Kovera & McAuliff, 2000), attorneys (Kovera, Russano, & McAuliff, 2002) as well as mock jurors (McAuliff & Duckworth, 2010) fared no better. In light of this work, it is perhaps not surprising that many of the mock jurors in the current research failed to adequately understand or weigh the relationship between delay in reporting and likely validity of test results in their decisions. Even with explicit guidance from the expert to direct them to this problem, research suggests that jurors often show difficulties in adequately utilizing this information (Levett & Kovera, 2008). Future research exploring the question of whether jurors are sensitive to variations in the probative value and application of various forms of evidence, including negative forensic results, is clearly warranted. Such research is particularly timely. Results from such a research agenda would not only provide decision-makers with a more informed database for determinations of the admissibility

of this type of evidence, but would also provide direction for ways in which to eliminate its potential negative impact in trials of sexual assault.

References

- Backman, R. (1998). The factors related to rape reporting behaviour and arrest. *Criminal Justice and Behavior*, 25, 8–29. <http://dx.doi.org/10.1177/0093854898025001002>.
- Bornstein, B. H. (1999). The ecological validity of jury simulations: Is the jury still out? *Law and Human Behavior*, 23, 75–91. <http://dx.doi.org/10.1023/A:1022326807441>.
- Buck, J. A., London, K., & Wright, D. B. (2011). Expert testimony regarding child witnesses: Does it sensitize jurors to forensic interview quality? *Law and Human Behavior*, 35, 152–164. <http://dx.doi.org/10.1007/s10979-010-9228-2>.
- Crowley, M. J., O'Callaghan, M. G., & Ball, P. J. (1994). The juridical impact of psychological expert testimony in a simulated child sexual abuse trial. *Law and Human Behavior*, 18(1), 89–105. <http://dx.doi.org/10.1007/BF01499146>.
- Cutler, B. L., Dexter, H. R., & Penrod, S. D. (1989). Expert testimony and jury decision making: An empirical analysis. *Behavioral Sciences & the Law*, 7, 215–225. <http://dx.doi.org/10.1002/bsl.2370070206>.
- Dorandeu, A., Pages, C., Sordino, M., Pepin, G., Baccino, E., & Kintz, P. (2006). A case in south-eastern France: A review of drug facilitated sexual assault in European and English-speaking countries. *Journal of Clinical Forensic Medicine*, 13, 253–261. <http://dx.doi.org/10.1016/j.jcfm.2005.11.013>.
- Du Mont, J., MacDonald, S., Bainbridge, D., Asllani, E., Smith, M., & Cohen, M. M. (2010). Drug-facilitated sexual assault in Ontario, Canada: Toxicological and DNA findings. *Journal of Forensic and Legal Medicine*, 17, 333–338. <http://dx.doi.org/10.1016/j.jflm.2010.05.004>.
- Fitzgerald, N., & Riley, K. J. (2000). Drug-facilitated rape: Looking for the missing pieces. *National Institute of Justice Journal* (Retrieved May 4th, 2004, from <http://www.ncjrs.org/pdffiles1/jr000243c.pdf>)
- Frazier, P. A., & Haney, B. (1996). Sexual assault cases in the legal system: Police, prosecutor, and victim perspectives. *Law and Human Behavior*, 20, 607–628. <http://dx.doi.org/10.1007/BF01499234>.
- Freese, T. E., Miotto, K., & Reback, C. J. (2002). The effects and consequences of selected club drugs. *Journal of Substance Abuse Treatment*, 23(2), 151–156. [http://dx.doi.org/10.1016/S0740-5472\(02\)00267-2](http://dx.doi.org/10.1016/S0740-5472(02)00267-2).
- Girard, A. L., & Senn, C. Y. (2008). The role of new "date rape drugs" in attributions about date rape. *Journal of Interpersonal Violence*, 23(1), 3–20. <http://dx.doi.org/10.1177/0886260507307648>.
- Hindmarch, I., & Brinkmann, R. (1999). Trends in the use of alcohol and other drugs in cases of sexual assault. *Human Psychopharmacology*, 14, 225–231. [http://dx.doi.org/10.1002/\(SICI\)1099-1077\(199906\)14:4<225::AID-HUP92>3.0.CO;2-3](http://dx.doi.org/10.1002/(SICI)1099-1077(199906)14:4<225::AID-HUP92>3.0.CO;2-3).
- Hindmarch, I., ElSohly, M., Gambles, J., & Salamone, S. (2001). Forensic urinalysis of drug use in cases of alleged sexual assault. *Journal of Clinical Forensic Medicine*, 8, 197–205. <http://dx.doi.org/10.1054/jcfm.2001.0513>.
- Horvath, M., & Brown, J. (2007). Alcohol as drug of choice: Is drug assisted rape a misnomer? *Psychology, Crime & Law*, 13, 417–427. <http://dx.doi.org/10.1080/10683160601061117>.
- Hurley, M., Parker, H., & Wells, D. (2006). The epidemiology of drug facilitated sexual assault. *Journal of Clinical Forensic Medicine*, 13, 181–185. <http://dx.doi.org/10.1016/j.jcfm.2006.02.005>.
- Jenkins, G., & Schuller, R. A. (2007). The impact of negative forensic evidence on mock jurors' perceptions of a trial of drug-facilitated sexual assault. *Law and Human Behavior*, 31, 369–380. <http://dx.doi.org/10.1007/s10979-006-9068-2>.
- Jordan, C. E. (2004). Intimate partner violence and the justice system: An examination of the interface. *Journal of Interpersonal Violence. Special Issue: Toward a National Research Agenda on Violence Against Women – Part 2*, 19(12), 1412–1434. <http://dx.doi.org/10.1177/0886260504269697>.
- Kovera, M. B., Gresham, A. W., Borgida, E., Gray, E., & Regan, P. C. (1997). Does expert psychological testimony inform or influence juror decision making? A social cognitive analysis. *The Journal of Social Psychology*, 82, 178–191. <http://dx.doi.org/10.1037/0021-9010.82.1.17>.
- Kovera, M. B., & McAuliff, B. D. (2000). The effects of peer review and evidence quality on judge evaluations of psychological science: Are judges effective gatekeepers? *The Journal of Applied Psychology*, 85, 574–586. <http://dx.doi.org/10.1037/0021-9010.85.4.574>.
- Kovera, M. B., McAuliff, B. D., & Hebert, K. S. (1999). Reasoning about scientific evidence: Effects of juror gender and evidence quality on juror decisions in a hostile work environment case. *The Journal of Applied Psychology*, 84(3), 362–375. <http://dx.doi.org/10.1037/0021-9010.84.3.362>.
- Kovera, M. B., Russano, M. B., & McAuliff, B. D. (2002). Assessments of the common-sense psychology underlying Daubert: Legal decision makers' abilities to evaluate expert evidence in hostile work environment cases. *Psychology, Public Policy, and Law*, 8, 180–200. <http://dx.doi.org/10.1037/1076-8971.8.2.180>.
- Krauss, D. A., & Sales, B. D. (2001). The effects of clinical and scientific expert testimony on juror decision making in capital sentencing. *Psychology, Public Policy, and Law*, 7(2), 267–310. <http://dx.doi.org/10.1037/1076-8971.7.2.267>.
- LeBeau, M., Andollo, W., Hearn, W. L., Baselt, R., Cone, E., Finkle, B. (1999). Recommendations for toxicological investigations of drug-facilitated sexual assault. *Journal of Forensic Science*, 44, 227–230.
- Levett, L. M., & Kovera, M. B. (2008). The effectiveness of opposing expert witnesses for educating jurors about unreliable expert evidence. *Law and Human Behavior*, 32, 363–374. <http://dx.doi.org/10.1007/s10979-007-9113-9>.
- McAuliff, B., & Duckworth, T. (2010). I spy with my little eye: Jurors' detection of internal validity threats in expert evidence. *Law and Human Behavior*, 34, 489–500. <http://dx.doi.org/10.1007/s10979-010-9219-3>.

- McGregor, M. J., Du Mont, J., & Myhr, T. L. (2002). Sexual assault forensic medical examination: Is evidence related to successful prosecution? *Annals of Emergency Medicine*, 39, 639–647. <http://dx.doi.org/10.1067/mem.2002.123694>.
- McGregor, M. J., Lipowska, M., Shah, S., Du Mont, J., & Siato, C. (2003). An exploratory analysis of suspected drug-facilitated sexual assault seen in a hospital emergency department. *Women & Health*, 37, 71–80. <http://dx.doi.org/10.1023/A:1005519028528>.
- McGregor, M. J., Wiebe, E., Marion, S. A., & Livingstone, C. (2000). Why don't more women report sexual assault to the police? *Canadian Medical Association Journal*, 162, 659–660.
- Olszewski, D. (2009). Sexual assaults facilitated by drugs or alcohol. *Drugs: Education, Prevention & Policy*, 16, 39–52. <http://dx.doi.org/10.1111/j.1471-6402.1998.tb00177>.
- Schuller, R. A., & Stewart, A. (2000). Police responses to sexual assault complaints: The role of perpetrator/complainant intoxication. *Law and Human Behavior*, 24, 535–551. <http://dx.doi.org/10.1023/A:1005519028528>.
- Schuller, R. A., & Wall, A. M. (1998). The effects of defendant and complainant intoxication on mock jurors' judgments of sexual assault. *Psychology of Women Quarterly*, 22, 555–573. <http://dx.doi.org/10.1111/j.1471-6402.1998.tb00177>.
- Schutte, J. W., & Hosch, H. M. (1997). Gender differences in sexual assault verdicts: A meta-analysis. *Journal of Social Behavior and Personality*, 12, 759–772.
- Scott-Ham, M., & Burton, F. C. (2005). Toxicological findings in cases of alleged drug-facilitated sexual assault in the United Kingdom over a 3-year period. *Journal of Clinical Forensic Medicine*, 12, 175–186. <http://dx.doi.org/10.1016/j.jcfm.2005.03.009>.
- Spohn, C., & Spears, J. (1996). The effect of offender and victim characteristics on sexual assault case processing decisions. *Justice Quarterly*, 13, 401–431. <http://dx.doi.org/10.1080/07418829600093141>.
- Temkin, J. (2010). "And always keep a hold of nurse, for fear of finding something worse": Challenging rape myths in the courtroom. *New Criminal Law Review*, 13, 710–734. <http://dx.doi.org/10.1525/nclr.2010.13.4710>.
- Temkin, J., & Krahé, B. (2008). *Sexual assault and the justice gap: A question of attitude*. Oxford: Hart Publishing.
- Wells, G. L. (1986). Expert psychological testimony: Empirical and conceptual analyses of effects. *Law and Human Behavior. Special Issue: The Ethics of Expert Testimony*, 10, 83–95. <http://dx.doi.org/10.1007/BF01044560>.
- Wells, D. (2001). Drug administration and sexual assault: Sex in a glass. *Science & Justice*, 41(3), 197–199. [http://dx.doi.org/10.1016/S1355-0306\(01\)71890-4](http://dx.doi.org/10.1016/S1355-0306(01)71890-4).
- Wiener, R., Krauss, D., & Lieberman, J. (2011). Differences between college student and more representative jurors. *A special issue of Behavioral Sciences and the Law*, 29, (pp. 325–479).