Validation of the factor structure and predictive validity of the Forensic Evidence Evaluation Bias Scale for robbery and sexual assault trial scenarios

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Psychology, Crime and Law</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>GPCL-2012-0140.R1</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Empirical Study</td>
</tr>
<tr>
<td>Keywords:</td>
<td>Juror decision making, Forensic, evidence, Decision making, attitudes</td>
</tr>
</tbody>
</table>
Abstract

In order to investigate the role of pre-trial attitudes about forensic science in juror decision making, a previous study demonstrated the predictive validity of the Forensic Evidence Evaluation Bias Scale (FEEBS), using a murder trial scenario, which featured ambiguous prosecution DNA evidence. The current study validates the FEEBS using two new crime types and the conditions include a manipulation of the presence of DNA evidence in the trial scenario. The FEEBS successfully predicted mock jurors’ perceptions of the probative value of DNA evidence for both robbery and sexual assault trials. The two subscales of the FEEBS were demonstrated to have different predictive ability depending on the presence or absence of DNA evidence. A confirmatory factor analytic technique was used to validate the underlying two-factor structure of the FEEBS, as previously proposed. These results are discussed with reference to the CSI Effect literature, and the potential for improvement to less empirically supported voir dire questioning techniques.
The right to a trial by jury is considered by some to be a celebrated feature of many adversarial legal systems, for instance in the United States and some Commonwealth countries (e.g. UK, Canada, Australia). One premise on which this judicial feature relies is the court’s ability to summon representative and impartial jurors (Levesque, 2006). Pre-trial bias has been a focus of legal psychological research for many years, as the prevailing cognitive model for juror decision making (the story model; Pennington & Hastie, 1986, 1992) highlights the contribution of prior beliefs and knowledge to the decision making process. This idea that jurors do not enter a courtroom as ‘blank slates’, and that their previous beliefs and knowledge play an important role in their verdict decisions, raises the question of whether there exists such a thing as an impartial juror.

In an effort to identify individual differences which influence juror decision making, early research focused on personality characteristics which might affect jurors’ verdict preferences, such as Locus of Control (Rotter, 1990), Authoritarianism (Jensen, 1957), and Belief in a Just World (Rubin & Peplau, 1975). Studies of this sort were relatively unsuccessful at demonstrating a clear link between personality constructs and verdict preference, with the possible exception of Authoritarianism which was moderately successful at identifying jurors who would be more punitive in their judgments. However Authoritarianism did not always reliably predict jurors’ decisions across various crime types and a meta-analysis concluded that legally relevant measures of authoritarianism (such as the Legal Attitudes Questionnaire by Boehm, 1968 and revised by Moran & Comfort, 1982) demonstrated better predictive validity than more general measures (Narby, Cutler, & Moran, 1993).

The conclusion that legally relevant measures of Authoritarianism were better predictors than general measures is not exclusive to the juror decision making literature. Research literature investigating the link between attitudes and behaviour has consistently demonstrated that specific attitudes are more predictive of behaviour than general attitudes, when the
behaviour being predicted is specific (Ajzen & Fishbein, 1977). Such findings provided theoretical support for the development of more juror-specific, and legally relevant, attitude measures which are likely to be more effective for identifying biases which may have an impact on verdict decisions. The first such measure to be developed was the Juror Bias Scale (JBS; Kassin & Wrightsman, 1983), which was designed to measure attitudes relevant to how jurors cognitively reach verdict decisions. The JBS measures attitudes on two continuums, based on the assumption that two separate assessments are made by jurors, represented as (1) the likelihood that the defendant committed the crime, and (2) the reasonable doubt threshold required to vote guilty (Kassin & Wrightsman, 1988; Myers & Lecci, 1998).

Early studies using the JBS to predict mock jurors’ verdict preferences found some support for the scale’s predictive ability, as jurors scoring high on the JBS (defined by a median split), which is coded in the direction of a prosecution bias, were more likely to vote guilty (Kassin & Wrightsman, 1983). There were some inconsistencies found across a number of studies, most notably the fact the JBS did not consistently predict verdict preference in rape trial scenarios (Kassin & Wrightsman, 1983; Myers & Lecci, 1998; Lecci & Meyers, 2002). This finding was hypothesised by Kassin and Wrightsman to be due to crime-specific biases operating in these cases, such as victim blaming attitudes and levels of rape myth acceptance.

It has been suggested by Narby, et al. (1993) that research investigating the relationship between attitudes and verdict preference has failed to produce consistent findings because of the lack of attention given to potential moderating variables, such as the strength of trial evidence. Scales such as the JBS were created and tested before the introduction of advanced forensic science techniques which are now commonplace in the courts and often form an important part of the prosecution’s case. The exact prevalence of forensic evidence in criminal trials is difficult to estimate, however for some crime types (such as murder) it is relatively rare for the prosecution not to rely on some scientific evidence.
Indeed it may be that the use of forensic science in trials has steadily increased in recent years, as prosecutors become more reluctant to proceed to trial without such evidence. In order to investigate how attitudes about forensic evidence might influence juror verdicts the Forensic Evidence Evaluation Bias Scale (FEEBS) was recently developed (authors’ names deleted for review). The FEEBS consists of ten items, which were derived based on pre-trial attitudes about forensic science hypothesised as being problematic by previous literature concerned with identifying the existence of the so-called CSI Effect (Tyler, 2006; Podlas, 2006; Cole & Dioso-Villa, 2007; Schweitzer & Saks, 2007).

The CSI Effect literature investigates the, largely anecdotal, claim that popular forensic science-based television programs are influencing jurors’ behaviour, and proposes two main effects (Cole & Dioso-Villa, 2007). One version of this effect is the proposition that jurors have unrealistically high expectations about the inclusion of forensic science in the prosecution’s case and are therefore unwilling to convict if this type of evidence is not presented (known as the anti-prosecution effect). The second effect is the claim that jurors are likely to perceive forensic science evidence as more probative than it actually is, due to its portrayal as conclusive and infallible in fiction (referred to as the defendant’s effect). Despite a growing number of recent empirical studies investigating the CSI Effect, most have not found a direct link between television viewing habits and juror decision making (Podlas, 2006; Shelton, Kim, & Barak, 2006; Schweitzer & Saks, 2007; Smith, Bull, & Holliday, 2010). However there is some empirical evidence to suggest that there might be an indirect link between viewing habits, expectations for forensic evidence, and subsequent verdict preference (Kim, Barak, & Shelton, 2009).

The FEEBS items were developed to measure attitudes held by mock jurors about forensic science evidence, as proposed in the CSI Effect literature, in order to test the predictive validity of these attitudes in relation to verdict preference. In the first two studies, a principal
components analysis of the FEEBS suggested that there were two distinct clusters of attitudes being measured (authors’ names removed for review). The first represents what were termed ‘pro-prosecution’ attitudes (subscale referred to as the FEEBpp), which correspond to beliefs about the infallibility and conclusiveness of forensic science. The second cluster was referred to as ‘pro-defence’ attitudes (FEEBpd), and these measured high expectations for the presence of forensic science evidence and a reliance on this type of evidence for conviction.

The predictive validity of the FEEB subscales was tested using a murder trial scenario (authors’ names removed for review) which was presented to 159 jury-eligible student participants, along with the JBS and the General Belief in a Just World scale (GBJW; Dalbert, Montada, & Schmitt, 1987). The FEEBpp subscale was found to be significantly (but not strongly) correlated with the JBS ($r = .24, p < .05$) and the GBJW scale ($r = .17, p < .05$) while the FEEBpd subscale was not significantly correlated with either scale. The murder trial scenario used in that study was designed to contain ambiguous forensic DNA evidence, as cases with strong evidence are less likely to lead to jurors relying on their prior attitudes or biases (Kassin & Wrightsman, 1983; De La Fuente, De La Fuente, & Garcia, 2003).

The main findings from the mock juror task in the initial FEEBS study suggested that, although the forensic DNA evidence was not the only evidence presented by the prosecution (there was circumstantial and witness evidence also presented), participants’ perception of the probative value of the DNA evidence was the only significant predictor of their verdict preference. Participants who perceived the ambiguous DNA evidence to be of high probative value were more likely to perceive the defendant to be guilty. The FEEBpp subscale was the only attitude measure which significantly predicted the perceived strength of the ambiguous DNA evidence ($\beta = .48, p < .001$), and thus the conclusion reached was that these pro-prosecution attitudes about forensic science evidence predict juror verdict preference via the mediator variable of perceived strength of the forensic evidence. This finding is congruent with
previous attitude-behaviour research which would predict more specific attitudes to be better predictors of behaviour compared with more generic attitudes (e.g. for JBS, $\beta = .02$, $p > .05$).

The first FEEBS study, described above, demonstrated the potential relationship between pro-prosecution attitudes about forensic evidence and juror verdict preferences in cases with ambiguous forensic DNA evidence. The FEEBS was a superior predictor of juror verdict over more general attitude measures, which is to be expected given previous, robust findings that general attitude measures do not predict behaviour as well as more specific, relevant attitudes (Ajzen & Fishbein, 1977). However, there are a number of questions which were not addressed by this initial study. Firstly, the pro-defence subscale of the FEEBS did not successfully predict any aspect of the mock juror task. It was concluded that this may be due to the fact that the pro-defence attitudes refer to an unwillingness to convict when there is no forensic evidence presented by the prosecution, which was not a scenario which was tested in the study.

Secondly, the only crime type used in the previous study was murder, and the predictive validity of the FEEBS should be tested with other crime types in order to determine whether this measure suffers from similar variability across different crimes to the JBS and Authoritarianism measures (Narby, et al., 1993; Kassin & Wrightsman, 1983). In particular, previous research suggests that in trials for sexual offences, general juror bias measures are less successful predictors of verdict due to other, more specific juror attitudes which can be involved in these types of trials (Kassin & Wrightsman, 1983; Myers & Lecci, 1998; Lecci & Myers, 2002). For example, victim blaming and rape myth acceptance attitudes may be contributing to the decision making process in these types of trials, resulting in a less direct relationship between general attitudes and verdict preference (Sleath & Woodhams, 2011).

In order to address some of these issues, the current study had a number of aims and hypotheses. The first is to apply a confirmatory factor analytic approach to the FEEBS items, which is an important step in validating the underlying structure of this novel attitude measure.
Based on the previous results from exploratory analyses, it was hypothesised that the FEEBS will best fit a two-factor model. The second aim was to test the predictive validity of the FEEBS for two additional crime types, robbery and sexual assault. Due to the increased specificity of the attitudes measured by the FEEBS, it was hypothesised that it would be a significant predictor for these new crime types. Finally, the inclusion of a trial scenario in which the prosecution does not present any forensic science evidence allowed the hypothesis that the pro-defence subscale of the FEEBS would predict verdict preference in the absence of forensic evidence to be tested. It is hoped that addressing these issues will further our understanding of how pre-trial attitudes about forensic science may influence juror decision making and verdict preferences.

Method

Participants

The target population for this study was jury-eligible members of the general public, and therefore participants met some basic criteria for jury service as described by the Criminal Justice System for England and Wales (www.cjonline.gov.uk). For the purposes of this research this meant that participants were members of the general public, between the ages of 18-69 years, and reported not ever being convicted of a criminal offence. The study was advertised through a press release issued by the University’s press office, and was advertised on various websites and newspapers throughout the England and Wales.

The final sample consisted of 200 participants, of which 71 (35.5%) were male, and 129 (64.5%) female. The age of participants ranged from 18-69 years ($M_{\text{age}} = 33.72$, $SD = 13.20$). Participants were also asked to indicate their highest level of education completed, and 25% of participants reported completing secondary school, 58% completed undergraduate degrees, and 17% held postgraduate qualifications.
Materials and procedure

Each participant read through one version of a trial, which included opening remarks by the prosecution and defence, a series of eyewitness and expert witness statements, and a summary by the prosecution and defence. The trial scenarios consisted of two versions of a robbery trial (with and without prosecution DNA evidence) and two versions of a sexual assault trial (with and without prosecution DNA evidence), resulting in a 2x2 between-subjects factorial design. Participants were randomly assigned to complete one of the four trial conditions. The scenarios were approximately 1600 words in length, and all versions contained nearly identical testimony from an eyewitness who saw the perpetrator briefly (but did not witness the crime occur), the responding police officer, and the arresting police officer. The responding police officer presented testimony relating to the questioning of the witness and victim, and the arresting officer testified about information obtained from interviewing the defendant which provided circumstantial evidence implicating the defendant in the crime.

The only trial evidence manipulated between conditions was the prosecution’s DNA evidence, which was either present (but of weak probative value) or absent. In conditions where this DNA evidence was presented, the testimony of the laboratory analyst was included in the trial materials and the evidence presented was of a weak probative standard, similar to the evidence in the previous FEEBS study involving a murder trial scenario (authors names deleted for blind review). The probative value of the DNA evidence was considered to be low due to the fact that the DNA evidence recovered could quite easily be explained without requiring the defendant to have committed the crime in question (Taroni, Aitken, Garbolino, & Biedermann, 2006). For instance, a cigarette end recovered from the street outside a crime scene with the defendant’s DNA found on it does not prove that the defendant committed the crime, and could plausibly be explained innocently.
The trial materials were adapted into an online format using Zoomerang, which is an online survey tool available by subscription. This allowed participants to complete the mock juror task, without requiring them to attend in person. By visiting an initial webpage set up for the study, participants were then randomly forwarded to only one of the four experimental conditions containing one version of the trial materials. Participants initially completed some demographic information to confirm basic jury eligibility followed by the FEEBS (authors’ names deleted for blind review), the JBS (Kassin & Wrightsman, 1983), and GBJW (Dalbert, Montada, & Schmidt, 1987) scale before moving on to read the trial scenario materials and complete the mock juror task.

Each trial scenario began with a summary of the case (outlining the charges faced by the defendant), after which participants indicated how probable they thought it was that the defendant was guilty of the crime (expressed as a percentage). This provides a quantitative estimate of each participant’s pre-trial perception of guilt. Testimony from each of the key witnesses followed on separate pages, with participants being asked to indicate how useful they perceived each piece of evidence to be for determining the guilt of the defendant on a 5 point Likert scale (ranging from irrelevant/useless to conclusive/definite). After reading all of the testimony, closing arguments were presented by the prosecution and defence, and participants were once again asked to indicate the probability of the defendant’s guilt as well as their verdict decision (guilty or not-guilty).

Results

Confirmatory factor analysis of the FEEBS

Previously, the FEEBS had only been subjected to exploratory factor analysis, which is a useful tool from a hypothesis generating perspective. It is however important with any new attitude measure to progress to use of a hypothesis testing method in order validate the measure.
(Mulaik, 2010). To this end, confirmatory factor analysis (CFA) was carried out on the hypothesised 2-factor model of the FEEBS, based on the previous research conducted using principal components analysis (abbreviated versions of the FEEBS items provided in Table 1). The CFA was conducted using AMOS version 16, and a summary of the goodness-of-fit statistics for the model are provided in Table 2.

The $\chi^2$ value for the model (39.06) was non-significant ($p = .25$), which indicates that the hypothesised model is a good fit with these data (DiStefano & Hess, 2005), however it has been noted that the $\chi^2$ statistic can be misleading with large sample sizes and therefore it is important to also interpret the other available goodness-of-fit statistics (Oakman, Van Ameringen, Mancini, & Farvolden, 2003; Bentler, 1990). The most common statistic to report in CFA is the Goodness of Fit Index (GFI), which should be greater than .90 if the model fits the data adequately (Oakman, et al., 2003), although some authors suggest values of .95 or greater should be achieved (DiStefano & Hess, 2005). Using either of these suggested threshold values, the GFI for the current CFA indicates a model that fits the data quite well. The final two statistics which are presented are the Delta 2, or Incremental Fit Index (IFI), and the Root Mean Squared Error of Approximation (RMSEA). The IFI value should be greater than .90 if the proposed model is a good fit to the data (Byrne, 2010), which indeed the current value is. The RMSEA takes into account the complexity of the model and should be less than .05 if the model is a close fit (Oakman, et al., 2003), which is true of the current model.

In addition to comparing the two-factor model with the independence model, which assumes a zero correlation between items and latent factors (reported above), it is also important to compare the hypothesised model with a single factor model (Mulaik, 2010). Therefore the factor analysis was run again forcing a solution with one single factor using the generalised least squares method, and the resulting chi-square goodness-of-fit test confirmed that this solution
was not an adequate fit to these data ($\chi^2 = 86.07, p < .001$). Figure 1 summarises the CFA results in a path diagram.

**Comparison of attitude measures**

The correlations between the pre-trial attitudes measured are provided in Table 3, and these correlations are similar to those found in the previous study (authors’ names deleted for review). The very small correlations between the FEEBS and other existing attitude measures (JBS PC and RD subscales, GBJW) provides support for the view that the attitudes represented by the FEEBS items are fundamentally different from those included in previous, more general pre-trial attitude measures.

**Probability of guilt estimates and verdict decisions**

Participants provided two probability of guilt estimates, once after reading only a summary of the case containing no evidence (pre-trial probability), and once after reading all of the trial testimony and the closing statements (final probability). The mean pre-trial probability of guilt estimates in each of the four conditions ($M_{overall} = 51.08$) did not differ significantly from a test value of 50% ($t = 1.85, p = .07$), indicating, as expected, that the scenario summaries were not perceived as favouring the prosecution or defence before the evidence was presented. Table 4 summarises the verdict decisions by condition, and participants who voted guilty had significantly higher final probability of guilt estimates ($M_{robbery} = 90.24, M_{sexual assault} = 89.62$) than those who voted not guilty ($M_{robbery} = 57.31, M_{sexual assault} = 65.83$), for both crime types ($t = 13.97, p < .001$) as would be expected.

**Relationship between attitude measures and probability of guilt estimates for DNA present conditions**
Having demonstrated that participants who voted guilty reported significantly higher probability of guilt estimates, multiple regression analysis was used to determine which pieces of evidence in the trial scenario predicted the reported probability that the defendant was guilty. The results are summarised in Table 5. The results replicated the previous findings for the murder trial scenario ($\beta = .46, p < .001$ in previous murder trial study), as the DNA evidence was the only significant predictor of probability of guilt in the robbery scenario, and the strongest predictor in the sexual assault scenario (although the eyewitness testimony was also a significant predictor in this condition).

In order to determine whether any of the pre-trial attitude measures predicted the perceived strength of the DNA evidence (which is of weak/ambiguous value to the case), another multiple regression was carried out. The results of this regression analysis are summarised in Table 6. Similar to previous findings, the only significant predictor of participants’ strength ratings for the DNA evidence was their score on the FEEBpp subscale ($\beta = .48, p < .001$ in previous murder trial study). This relationship is positive, indicating that stronger pro-prosecution beliefs about forensic evidence (attitudes relating to absolute faith in the capabilities of forensic science) predicted higher ratings of probative value for the ambiguous/weak DNA evidence in these scenarios.

The relationship between pro-prosecution attitudes about forensic science (FEEBpp), perceived strength of DNA evidence, and final probability of guilt estimates for both the robbery and sexual assault conditions were found to satisfy the conditions of a partial mediation model (depicted in Figure 2). For both crime types, the Sobel test (as outlined in Preacher & Hayes, 2004) confirmed a significant indirect effect of the mediator variable (perceived strength of DNA evidence) on probability of guilt ($Z = 3.17, p < .05$ for robbery; $Z = 2.81, p < .05$ for sexual assault). Both of these models produced similar effect sizes using Fairchild, MacKinnon, Taborga, and Taylor’s (2009) $R^2$ effect size measure which expresses the amount of variance in
final probability of guilt estimates which is accounted for by the mediator variable; for robbery 
$R^2 = .40$ and for sexual assault $R^2 = .34$.

Relationship between attitude measures and probability of guilt estimates for DNA 
absent conditions

As this is the first study to test the predictive validity of the FEEBS with trial scenarios in 
which no forensic evidence is included in the prosecution’s case, multiple regression analysis 
was used to investigate which pieces of trial evidence predicted probability of guilt estimates for 
robbery and sexual assault. The results are presented in Table 5. The only significant predictor of 
perceived probability of guilt was the testimony from the arresting officer, for both robbery and 
sexual assault, which seems reasonable as this was the only testimony to provide substantial 
circumstantial evidence for the defendant’s guilt.

A further multiple regression analysis was conducted to determine whether any of the 
pre-trial attitude measures predicted the perceived strength of the arresting officer’s testimony. 
The results are presented in Table 6. The only significant predictor, for both robbery and sexual 
assault, was the participant’s score on the FEEBpd subscale. This relationship was negative, 
indicating that stronger pro-defence attitudes about forensic science (e.g. high expectations for 
the presence of forensic evidence in criminal trials) predicted lower strength ratings for the 
circumstantial evidence (as provided by the arresting officer’s testimony). This would be 
predicted by the CSI Effect literature, which hypothesises that jurors with unrealistically high 
expectations for the presence of forensic evidence are less likely to be persuaded by other types 
of non-scientific evidence, resulting in an unwillingness to convict in these cases.

Similar to the ‘DNA present’ conditions, a partial mediation model was proposed for the 
relationship between FEEBpd scores (representing unrealistically high expectations for the 
presence of forensic evidence in an investigation), perceived strength of the circumstantial
evidence, and final probability of guilt estimates. Only in the robbery condition was the mediation model demonstrated to be significant, yielding an effect size of $R^2 = .25$ (Fairchild, et al., 2009). This relationship is depicted in Figure 3.

**Discussion**

In this paper a confirmatory factor analytic approach was used to confirm the underlying structure of the FEEBS, drawing on the results of exploratory principal components analyses described in a previous study (authors’ names deleted for review) for further empirical and theoretical support. The hypothesised two-factor model of the FEEBS was found to be a good fit to the data collected, which lends support to the use of the FEEBS as a measure of pre-trial beliefs about forensic science evidence held by jury eligible members of the public. By confirming this factor structure further evidence is provided for the existence of reliable and distinct subscales, which represent pro-prosecution (FEEBpp) and pro-defence (FEEBpd) clusters of attitudes about forensic science evidence. These factors are conceptually similar to the problematic beliefs anecdotally described in much of the CSI Effect literature, described by Cole and Dioso-Villa (2007) as the ‘Strong Prosecutor’ and ‘Defendant’ effects.

In a previous study (authors’ names deleted for review) there were no significant correlations between the FEEBpd subscale and the JBS or GBJW measures, which seemed surprising as it could be argued that a small negative correlation might be expected. However, such negative correlations were found in the present study. It is not clear why these correlations differ between the studies, although it could be due to the fact that the previous study sample consisted of jury-eligible student participants while the current study sampled jury-eligible members of the general public. It is well documented that student samples differ in many respects from community-based samples (both in demographic characteristics and attitudes and behaviour), and have been shown to display lower correlations between legal attitude measures.
overall compared with general members of the public or jury samples (Narby, et al., 1993),
which is indeed the trend found here.

Previously, the usefulness of the FEEB subscales for predicting participants’ perceptions
of weak DNA evidence in a criminal trial scenario was only tested using a murder scenario, and
therefore it was important in the current study to present trial material for different crime types
(robbery and sexual assault). The inclusion of multiple crime types is crucial as some previous
juror bias measures have been shown to have inconsistent findings across different crime types,
particularly in the case of sexual assault (Kassin & Wrightsman, 1983, 1988; Lecci & Myers,
2008). In addition, trial stimuli in which no forensic DNA evidence was presented by the
prosecution were included in order to determine whether the FEEBpd subscale scores are
associated with perceptions of cases where no forensic evidence is included.

In the robbery scenario which included weak DNA evidence the perceived strength of
this evidence was the only significant predictor of probability of guilt estimates made by
participants, which is consistent with the previous study and also with the existing literature
(Skolnick & Shaw, 2001). For the sexual assault scenario the weak DNA evidence was also the
most significant predictor of probability of guilt, however the perceived strength of the
eyewitness testimony was also found to be a significant predictor. Given that the eyewitness in
the scenario was not able to identify the perpetrator, it may seem strange that participants found
this testimony to be important for determining guilt in the sexual assault trial. However some
research has found that people’s perceptions of the credibility of rape allegations were dependent
upon whether a witness could provide evidence that the victim had suffered an attack, for
instance by coming across the victim after the attack had occurred (Sleath & Woodhams, 2011).
Therefore, it could be the case that participants in the current study associated the eyewitness
testimony with increased victim credibility, which in turn may have impacted on their estimates
of the probability of the defendant’s guilt.
For both the robbery and sexual assault scenarios the only pre-trial attitude measure which significantly predicted the perceived strength of the weak DNA evidence was the FEEBpp subscale. The relationship was such that the stronger the pro-prosecution beliefs about forensic evidence, as measured by the FEEBpp scale, the higher the strength rating for the weak DNA evidence in the scenario. This finding is consistent with the results in the previous study involving the murder trial scenario, and this provides some further evidence for the reliability of the FEEBpp subscale for predicting how mock jurors will perceive weak DNA evidence in the context of a criminal trial. This relationship was shown to operate as a similar partial mediation model to that which was demonstrated in the previous research.

In the robbery and sexual assault scenarios in which no DNA evidence was included as part of the prosecution’s case, the testimony from the arresting police officer was the only significant predictor of participants’ probability of guilt estimates. This finding seems reasonable as, in the absence of any physical evidence, participants were relying on the circumstantial evidence present in this testimony for their decisions about how likely it was that the defendant committed the crime. The only pre-trial attitude measure which was shown to be associated with the perceived strength of this circumstantial evidence was the FEEBpd subscale. This relationship was negative which indicated that the stronger a participant’s pro-defence beliefs about forensic science (e.g. unrealistically high expectations for the presence of forensic evidence in criminal trials) the less likely they were to find the circumstantial evidence compelling, and perceive the defendant as guilty, which seems to correspond to the so-called ‘anti-prosecution’ version of the CSI effect (Podlas, 2006). It was demonstrated here that participants who have unrealistically high expectations for the presence of forensic evidence are less willing to convict when such evidence isn’t presented, as hypothesised by the CSI Effect literature.
The findings for both the DNA present and DNA absent versions of the scenarios seem to be consistent with what would be predicted by the CSI effect hypothesis, in that the beliefs held about forensic science predicted how evidence (both forensic and circumstantial) was perceived, and in turn the verdict decision made by participants. However, the CSI effect hypothesis would go on to predict that these beliefs come from exposure to crime fiction dramas on television. The failure to find this direct link between television viewing and decision making in the previous literature has led some authors to conclude that no such effect exists (Podlas, 2006; Schweitzer & Saks, 2007; Shelton, et al., 2006).

However, it has been demonstrated here that these beliefs about forensic science did contribute significantly to participants’ perceptions of evidence strength, defendant guilt and verdict decisions. The rationale for the development of the FEEBS was based on the argument that pre-trial attitudes about forensic science evidence held by potential jurors may have an impact on how ambiguous forensic evidence is perceived in the context of criminal trials. The source of these pre-trial attitudes in the popular media has been the pursuit of the previous CSI effect literature, but is of less concern in the study described in this paper. However, what has been demonstrated is that the pre-trial beliefs hypothesised by the CSI effect literature can be measured and used to predict mock juror perceptions of forensic evidence in cases where this evidence is of a low probative value or absent altogether.

Although some of the limitations identified in previous studies were improved upon in this paper (e.g. using a sample from the general public, including a no DNA condition), there are still some methodological issues which limit the applicability of these findings to juries in actual courtrooms. Firstly, the presentation format of the trial materials has been limited in an attempt to control for extraneous variables which might impact on the decision making process, however the result is trial scenarios which are lacking much of the case detail that would be presented to real juries. The use of more realistic material presentation format has been found to impact on
participants’ responses (Martin, De La Fuente, De La Fuente, & Garcia, 2007) and so the
simplistic format of the materials in the current study is likely to limit the generalisability of the
results.

Secondly, a real jury verdict decision depends on the outcome of group deliberations and
in the study described within this paper the participants were not participating in group decision
making. It is unclear how the process of deliberation might impact on decisions about forensic
evidence and verdicts, however for the purposes of this research it is the individual juror pre-trial
attitudes which are of interest and for this reason no deliberation has been included in the
methodology. Some previous research has demonstrated that the deliberation process
exacerbates the impact of some biases, such as pre-trial publicity bias (Shaw & Skolnick, 2004),
so this would be an interesting area to investigate in relation to the attitudes identified by the
FEEBS. This may be of particular relevance as juries may be increasingly likely to be
homogenous with respect to attitudes and pre-trial beliefs as a result of peremptory challenges
exercised by the prosecution and defence attorneys (Hastie, 1991).

The practical implications of this research could include the use of the FEEBS items to
inform voir dire questioning techniques, as currently courts are reporting an increasing use of
questions such as ‘do you watch CSI?’ which are believed anecdotally to identify potentially
biased jurors. However the existing literature would suggest that questions such as these are
unlikely to provide any guidance to the judiciary regarding potential juror biases as they have
not been shown to predict juror verdict decisions. Therefore by demonstrating clusters of
attitudes which can successfully predict verdict preference (such as those measured by the
FEEBS), the validity of voir dire questioning may be improved significantly. Future research
could investigate this further by comparing the predictive validity of standard voir dire questions
with more empirically-derived questions using standardised attitude measures.
References


Table 1

*Forensic evidence evaluation bias scale (FEEBS) items*

<table>
<thead>
<tr>
<th>Pro-prosecution subscale (FEEBpp)</th>
<th>Pro-defence subscale (FEEBpd)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forensics always identifies the guilty person (FEEB6)</td>
<td>If there is no forensic evidence, the jury should not convict (FEEB7)</td>
</tr>
<tr>
<td>Forensic evidence alone is enough to convict (FEEB4)</td>
<td>No forensic evidence means investigators didn’t look hard enough (FEEB5)</td>
</tr>
<tr>
<td>Science is always the most reliable way to identify perpetrators (FEEB10)</td>
<td>If no forensic evidence is recovered then the defendant is probably innocent (FEEB2)</td>
</tr>
<tr>
<td>Forensics always provides a conclusive answer (FEEB8)</td>
<td>Police should not charge someone without forensic evidence (FEEB9)</td>
</tr>
<tr>
<td>Every crime can be solved with forensic science (FEEB1)</td>
<td>Every criminal leaves trace evidence at every scene (FEEB3)</td>
</tr>
</tbody>
</table>

*Note.* These are abbreviated versions of the items presented to participants. The item number in parentheses refers to the same item numbers cited in Figure 1.
Table 2

**Summary of Goodness-of-Fit statistics for 2-factor model of the FEEBS**

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$</th>
<th>df</th>
<th>GFI</th>
<th>IFI/Delta 2</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-factor model</td>
<td>39.06</td>
<td>34</td>
<td>.96</td>
<td>.96</td>
<td>.03</td>
</tr>
<tr>
<td>Independence model</td>
<td>170.10</td>
<td>45</td>
<td>.83</td>
<td>0.00</td>
<td>.12</td>
</tr>
</tbody>
</table>

*Note. N = 200*
Table 3

Correlations between pre-trial attitude scale measures

<table>
<thead>
<tr>
<th></th>
<th>FEEBpp</th>
<th>FEEBpd</th>
<th>JBS-PC</th>
<th>JBS-RD</th>
<th>GBJW</th>
</tr>
</thead>
<tbody>
<tr>
<td>FEEBpp</td>
<td>-</td>
<td>-.49**</td>
<td>.12</td>
<td>.06</td>
<td>.18*</td>
</tr>
<tr>
<td>FEEBpd</td>
<td></td>
<td>-</td>
<td>-.17*</td>
<td>-.12</td>
<td>-.23*</td>
</tr>
<tr>
<td>JBS-PC</td>
<td></td>
<td></td>
<td>.29**</td>
<td>.15*</td>
<td></td>
</tr>
<tr>
<td>JBS-RD</td>
<td></td>
<td></td>
<td></td>
<td>.13</td>
<td></td>
</tr>
</tbody>
</table>

Note. *p < .05, **p < .001; N = 200.
Table 4

*Frequency of verdict decisions by condition*

<table>
<thead>
<tr>
<th></th>
<th>Guilty</th>
<th>Not guilty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robbery no DNA</td>
<td>20 (40%)</td>
<td>30 (60%)</td>
</tr>
<tr>
<td>Sexual assault no DNA</td>
<td>16 (32%)</td>
<td>34 (68%)</td>
</tr>
<tr>
<td>Robbery with DNA</td>
<td>21 (42%)</td>
<td>29 (58%)</td>
</tr>
<tr>
<td>Sexual assault with DNA</td>
<td>26 (52%)</td>
<td>24 (48%)</td>
</tr>
<tr>
<td>Total</td>
<td>83 (41.5%)</td>
<td>117 (58.5%)</td>
</tr>
</tbody>
</table>

*Note. N = 200; N = 50 in each condition.*
Table 5

*Multiple regression analyses predicting final probability of guilt in the robbery and sexual assault for DNA present and DNA absent evidence conditions*

<table>
<thead>
<tr>
<th>Predictor</th>
<th>DNA Present</th>
<th></th>
<th></th>
<th></th>
<th>DNA Absent</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Robbery (N=50)</td>
<td>Sexual assault (N=50)</td>
<td>Robbery (N=50)</td>
<td>Sexual assault (N=50)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>31.43</td>
<td>19.18</td>
<td>8.95</td>
<td>26.58</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>15.72</td>
<td>11.06</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eye witness</td>
<td>-2.13</td>
<td>5.20</td>
<td>-6.38</td>
<td>2.65</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Responding police</td>
<td>-1.36</td>
<td>2.36</td>
<td>1.41</td>
<td>-1.51</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arresting police</td>
<td>-4.00</td>
<td>-1.31</td>
<td>18.96</td>
<td>10.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>.56</td>
<td>.52</td>
<td>.34</td>
<td>.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* *p < .05, **p < .001
Table 6

Multiple regression analyses predicting perceived strength of incriminating evidence (DNA or arresting officer testimony) in the robbery and sexual assault conditions

| Predictor | DNA Evidence | | | | Arreasting Officer testimony | | | |
|-----------|--------------|---|---|---|-----------------|---|---|---|---|
|           | Robbery (N=50) | Sexual assault (N=50) | Robbery (N=50) | Sexual assault (N=50) | | | | |
| Constant  | 1.67 | 1.48 | 2.46 | 1.37 | 3.41 | 1.22 | 4.99 | 1.14 |
| FEEBpp    | .10 | .04 | .45* | .21 | .05 | .60** | .02 | .03 | .09 |
| FEEBpd    | -.03 | .04 | -.14 | -.01 | .07 | -.03 | -.09 | .03 | .44** |
| JBS-PC    | .02 | .03 | .08 | -.02 | .04 | -.12 | -.01 | .03 | -.05 |
| JBS-RD    | .05 | .03 | .19 | -.01 | .03 | -.03 | .03 | .03 | .17 |
| GBJW      | -.02 | .02 | -.09 | -.01 | .03 | -.03 | .03 | .02 | .18 |
| R²        | .35 | | .38 | | .39 | | .14 | | |

Note. * p < .05, ** p < .001
Response to reviewers’ comments

We would like to thank the reviewers and editor for their thoughtful and helpful comments, and have addressed them with the following changes to the manuscript.

1. P.19 – Grammar in sentence corrected.
2. Spacing corrected to follow APA format (main body and references)
3. Hypotheses added to the introduction (final paragraph)
4. Added justification for expecting the FEEBS to be better able to contribute to predicting juror verdicts that other, more general, attitude measures (p. 6)
5. One reviewer suggested that if the CFA isn’t adding any ‘new’ knowledge, it could be a shorter section of the paper. We have left the details of the CFA in their current form, as the amount of detail provided is the minimum required for the reader to interpret the results. However, we have added a sentence on page 10, which highlights why conducting/reporting the CFA is crucial in this paper.
6. In response to reviewer one’s comments about how the regressions were displayed, and the editor’s request to reduce the number of tables, the regression tables have been combined into two more comprehensive tables. This makes it easier to compare the regression results across the various conditions, and reduces the number of tables.
7. Interpretation of the mediation models is included in the discussion (p. 16). In addition, an erroneous minus sign has been removed from the figures (which addresses some of the confusion expressed by reviewer one).
8. The reason for testing the FEEBS with different crime types is outlined on page 6.
9. Random assignment of participants was added to p. 8.
10. Further detail, in the form of an example, provided for why the forensic evidence in the scenarios was considered to be of weak probative value (addressing comments from reviewers one and three).
11. Added details of t-test statistic and p-value to page 11.
12. As suggested by reviewer 1, reiterated the types of attitudes represented by the FEEBp on page 12. We also did the same on page 13 regarding the FEEBpd subscale, although this wasn’t mentioned by the reviewer.
13. Reviewer three mentioned that we might wish to consider the crime severity as a possible explanation for the correlations between the FEEBS and other attitude measures. However, this is not as explanation in this study (or the previous study) because participants did not know what crime type/severity they were going to read until after they completed the attitude measures. Therefore crime type/severity could not have influenced their responses on the attitude scales.
14. The abstract has been reworded, and is hopefully clearer now.
15. Page two, the sentence beginning “The issue of neutrality...” has been reworded, and is hopefully clearer now.
16. Further detail has been added to pages 8 and 9 to clarify the task carried out by the participants.
Figure 1. Confirmatory factor analysis of the two-factor FEEBS with standardised regression weights (factor loadings)
Figure 2. Partial mediation model expressing the indirect relationship between attitudes measured by the FEEBpp, perceived strength of DNA evidence, and final probability of guilt estimates.

Note. Standardised regression coefficients provided are for robbery followed by sexual assault trials. The standardised regression coefficient between FEEBpp and probability of guilt controlling for perceived strength of evidence is in parentheses. *coefficient significant at $p < .001$. 
Figure 3. Partial mediation model for relationship between FEEBpd, perceived strength of circumstantial evidence, and final probability of guilt in the robbery with no DNA condition.

Note. Standardised regression coefficients provided are for robbery. The standardised regression coefficient between FEEBpd and probability of guilt controlling for perceived strength of circumstantial evidence is in parentheses. *coefficient significant at $p < .001$. 