Mock-Juror Reactions to Multiple Interview Presentation and Rapport-Building

Genevieve F. Waterhouse[1]

Anne M. Ridley[2]

Ray Bull[3]

Rachel Wilcock[1]

Author Note

[1]Department of Psychology, University of Winchester; [2] Department of Psychology, London South Bank University; [3]Department of Criminology and Law, University of Derby.

Correspondence should be addressed to Genevieve F. Waterhouse, Department of Psychology, Faculty of Humanities and Social Sciences, University of Winchester, Sparkford Road, Winchester, UK. SO22 4NR. Email: genevieve.waterhouse@winchester.ac.uk

This is an Accepted Manuscript of an article published by Taylor & Francis in PSYCHOLOGY, CRIME & LAW on 21 August 2019, available online: https://www.tandfonline.com/doi/full/10.1080/1068316X.2019.1652745.

Mock-Juror Reactions to Multiple Interview Presentation and Rapport-Building

Abstract

In the UK and some US states, video-recorded investigative interviews of child victims/witnesses can be presented in court as the child's evidence-in-chief. However, there is scarce advice or research on the effect that presenting different sections of the interviews may have on juror perceptions of the child's testimony. Two aspects of testimony presentation are examined here: first, whether to show the rapport-building phase of the interview, and second, the presentation of multiple interviews (i.e., more than one interview with the same child). Participants (n = 103) informed they were watching two interviews of the same child separated by a week had more positive perceptions of the child's testimony than those informed they were watching just one extended interview with a ten-minute break. Also, those watching the rapport-building phase had less positive perceptions of the child's testimony than those who did not watch this phase. Participants' perceptions of the interviewer and their case progression decisions were mainly not related to the above presentational differences. Thus, (i) mock-jurors were not inherently biased against multiple interviews and (ii) decisions regarding whether or not to show the rapport-building phase in court may have significant effects on jurors' perceptions of the child and their testimony.

Keywords: jury decision-making; repeated interviews; child witness; rapport-building; eyewitness testimony.

Introduction

In a number of jurisdictions worldwide, including some US states, England, Wales, and New Zealand, children's investigative interviews can be used in court to replace their live evidence-in-chief (Anderson, Gross, Sonne, Zajac, & Haynes, 2016). The use of video-recorded interviews in this context has been shown to reduce children's stress (Davies, Wilson, Mitchell, & Milsom, 1995). However, the courts have some power to edit these interviews for presentation to jurors and there has been little research to determine the effect this may have on jurors' perceptions of testimony. In particular, the Crown Prosecution Service for England and Wales suggests that interviews are edited to ensure they are an acceptable length by removing inadmissible and irrelevant evidence. Concerns regarding length have also been raised in New Zealand (Burrows & Powell, 2014). An aspect of investigative interviewing that may affect the length and could be interpreted as irrelevant and thus, possibly be edited out, is the rapport-building phase (highly recommended in relevant guidance). Additionally, in cases where children have been interviewed more than once (or 'multiple interviews'), editing or solely showing one interview may be considered to reduce the length of the interviews. However, there is a distinct lack of research examining how mock-jurors' perceptions are affected by the rapport-building phase or presentation of single or multiple interviews with a child.

In England and Wales, child interviews are conducted according to the 'Achieving Best Evidence in Criminal Proceedings: Guidance on interviewing victims and witnesses, and guidance on using special measures' (henceforth ABE; Ministry of Justice, 2011). These guidelines advise interviewers to initiate their interviews with a brief but meaningful rapport-building phase. This partly entails the interviewer asking the child neutral questions which are unrelated to the alleged offence in question and can be answered positively. The aims of including this phase in the interview are to create a positive atmosphere, reduce the

interviewee's anxiety, and free cognitive resources for recalling the key event (Ministry of Justice, 2011). The rapport-building phase also often includes a discussion of 'ground rules' for the interview, although arguably this component does not necessarily result in rapport being built between the interviewer and interviewee. 'Ground rules' include a description of how the interview will be conducted, a brief discussion ascertaining the child's understanding of truth and lies, and 'rules' such as empowering the child to say when they do not know the answer to a question. The rapport-building phase, therefore, includes no discussion of the alleged event itself and could be removed from the video recording in order that the jury is not subjected to an 'overly-long' interview.

Although multiple interviews do involve additional, but arguably 'overly-long', discussion of the alleged event itself, there is little guidance as to how or whether to present multiple video-recorded interviews. This is despite multiple interviews being conducted in many cases in the UK (Plotnikoff & Woolfson, 2001) and elsewhere (such as the USA, Norway, and Portugal; Carnes, Nelson-Gardell, Wilson, & Orgassa, 2001, Korkman, Pakkanen, & Laajasalo, 2017, Peixoto, Ribeiro, Fernandes, & Almeida, 2015). Additionally, reminiscence (i.e., new information provided in a later recall session) often occurs in such interviews (Hershkowitz & Terner, 2007; Katz & Hershkowitz, 2012; Waterhouse, Ridley, Bull, La Rooy, & Wilcock, 2016), and so counsel may wish to show more than one of the child's interviews in court.

There is scarce research on the effect of presenting the rapport-building section or multiple interviews with the same child to mock-jurors, and no research has examined the impact of both these factors together. The only study to have examined how rapport may affect jurors' perceptions of a child's testimony found that when mock-jurors were presented with the full written rapport-building transcript as well as the substantive section of an interview with either a four or six year old (compared to the substantive section alone), mock-

jurors judged the child as more accurate, honest, credible and confident (Krähenbühl, 2012). They also judged the interview to be of better quality, the child to have understood more about the interview concept, and to have given more complete and clear information. Thus, mock-jurors' ratings of the child witness and the interview were more positive when they read the full transcript including the rapport-building phase than when they just read the substantive section. Krähenbühl offered two possible explanations for this. First, the full rapport-building transcript gave jurors a greater insight into why and how the interview was conducted and confirmed that the child also understood this, possibly due to the discussion of the 'ground rules'. Viewing this section, therefore, allowed the mock-juror a greater awareness of the interviewer and interviewee's shared understanding of the purpose of the interview; a crucial feature of effective communication (Krähenbühl). This greater insight therefore may have affected the way that jurors made their credibility judgements. Second, Krähenbühl suggested that the findings may have been caused by the increased length of the transcript that included the full rapport-building section. However, she suggested this was unlikely because other studies have shown interview length to have no effect on trial convictions in the field (e.g., Wilson & Davies, 1999).

An alternative explanation is that reading the rapport-building gave mock-jurors a greater understanding of the emotional context of the interview. Rapport-building is, in theory, meant to ease children's anxieties regarding the investigative interview and to create camaraderie between the interviewer and the interviewee. Thus, Krähenbühl's (2012) participants who read the full rapport-building transcript may have perceived there to be a better interviewer-interviewee relationship than those who were not exposed to the rapport-building section and because of this viewed the child and the interview more positively. However, Bottoms, Rudnicki, and Nysse-Carris (2004, as cited in Bottoms, Quas, & Davis, 2007), found that watching supportive interviews including rapport-building and other forms

of non-verbal interviewer-provided social support, such as eye contact with the child, smiling, a relaxed body posture and warm vocal intonation, decreased adults' perceptions of a seven or eight year old child's believability and credibility. In their study, the group of adults who viewed non-supportive interviews with child witnesses (i.e., interviews with no rapport-building or non-verbal support) judged the children as more credible, less nervous, and they believed the child more than did the group who watched supportive interviews. Bottoms *et al.* (2007) suggested the reasons for more positive perceptions of non-supportive interviews may be that these interviews encouraged more sympathy for the child, or that adults may believe supportive interviews to be coercive. These two studies' contradictory results may be due to the differing formats of the interviews (video vs. transcript) or the aspects of rapport-building shown. For instance, it is possible Krähenbühl's (2012) sample's perceptions were affected by the presentation of the 'ground rules' section rather than the inclusion of social support, or that Bottoms *et al.*'s (2004) findings were a result of exposure to non-verbal supportive behaviours which were not visible in Krähenbühl's transcript study rather than rapport-building. At present, the effects of exposure to social support in the form of either rapport-building or generally supportive interviews on mock-jurors' perceptions of child witnesses are unclear.

Research regarding mock-juror perceptions of multiple interviews is even scarcer. The only study, to the authors' knowledge, that has examined mock-jurors' opinions of testimony provided in multiple interviews is that of Yozwiak, Golding, and Marsil (2004). Mock-jurors were given a written account of a (fictional) child sexual abuse case, including descriptions of the alleged victim's testimony, the defendant's testimony, the testimony of the detective and of a friend of the defendant, and the judge's instructions. There were two different versions of the summary of the six year old victim's interviews. In one version, the child victim gave complete disclosure in two interviews, and in the other, the child victim

partially disclosed in the first interview and then fully disclosed in the second. They found that mock-jurors perceived the latter child as less believable and gave significantly fewer guilty verdicts than for the former. This finding is troubling, as Yozwiak *et al.* highlighted, because children (and adults) often do not fully disclose in their first investigative interview, which may in fact be the cause and benefit of a second interview in some cases (e.g., Waterhouse *et al.*, 2016). Therefore, presenting interviews in which children partially disclose across multiple interviews may be a relatively common occurrence in courts, which in turn may affect jurors' decision-making. However, there are a number of limitations to the ecological validity of this study. First, participants were only provided with summaries of the interviews, which is unlikely to occur in courts (e.g., in England and Wales). Second, as in Krähenbühl (2012), the mock-jurors read the information about the trial. Brief written summaries of a child's recall in two interviews make it very easy to determine how consistent the child was. Watching two full video-recorded interviews may affect participants' perceptions of consistency differently particularly given that there is more information to keep track of. In addition, mock-jurors could use non-verbal and tonal cues when assessing a child's testimony which cannot be accurately portrayed when mock-jurors read transcripts.

Related to time of disclosure is the issue of perceived inconsistency of testimony: studies do support the notion that mock-jurors perceive inconsistent testimony as less reliable than consistent testimony. In Leippe, Manion, and Romanczyk's (1992) study, mock-jurors' ratings of a five to six year old child's consistency within an interview were strongly and positively associated with their ratings of the child's believability and memory accuracy. Relatedly, Quas, Thompson, and Clarke-Stewart (2005) asked participants to what extent they agreed with statements about consistency and found that 29% agreed that 'Inconsistencies in a child's report of sexual abuse indicate that the report is false' (p. 439). Fisher, Brewer, and Mitchell (2009) described this as part of the 'courtroom theory' of

memory (p.126) which, based on an understanding of forgetting, results in people viewing new information provided at later dates (i.e., reminiscence) with suspicion. Fisher *et al.* argue that it is presumed that this additional information comes from a non-crime source and so may be inaccurate. At the other end of the scale, Leippe *et al.* (1992) suggest that consistency may be an indicator of expertise and that if a message is internally consistent, the listener may infer that the communicator is an expert in the subject. Thus, perceived inconsistency may well affect some mock-jurors' perceptions of a child's testimony, despite research suggesting that consistency is an unreliable indicator of children's accuracy, and that instead the quality of the interview (i.e., the inclusion of best practice question types) is a better indicator of accuracy (Fisher, *et al.*, 2009).

The present innovative study therefore acts as a crucial study examining the effects of being exposed to rapport-building and multiple interviews on adults' perceptions of a child interview. It aimed to increase ecological validity by providing participants with video-recorded interviews rather than transcripts. Regarding rapport-building, it also aimed to disentangle the effects of exposure to the neutral rapport-building discussion from those of viewing the 'ground rules' of the interview. Thus, all mock-jurors were only provided with a brief written account of the 'ground rules', and those that watched the rapport-building section only viewed the neutral discussion part (and not the 'ground rules'). This was in order to examine if removing this non-substantive section of the interview would affect mock-juror perceptions. Based on conflicting previous research, it was hypothesised that participants who viewed the rapport-building recording would perceive the child significantly differently from those who did not watch the rapport-building phase, but this hypothesis was non-directional.

To examine these two variables, the same interviews were presented with and without the rapport-building phase, and written descriptions stated that the interviews shown were

either one interview with a ten-minute refreshment break within it, or two interviews with a week delay between them. Due to the lack of prior research examining mock-jurors' perceptions of single versus multiple interviews, no directional hypotheses were made as to how perceptions would be affected by whether participants believed they were watching a single interview or two separate interviews, but perceptions were expected to differ.

Finally, the present study also examined juror perceptions of consistency within single and multiple interviews by asking participants to rate the likelihood that specific details provided within the child's testimony actually occurred in what the child witnessed. These details differed according to the consistency with which they were provided (once, twice, or contradictory) to determine if the consistency of recall of individual details affects mock-jurors' perceptions of the child's credibility regarding those specific details. It was predicted that mock-jurors would rate repeated details as more likely to have happened than those that were mentioned only once or that were contradicted by the child. These perceptions may also have been affected by how long a delay the participants believed there to have been between the two interviews (i.e., a matter of minutes or a week). Therefore, the effect of such delay on these perceptions was examined.

Method

Recruitment

A referral sampling design was utilised. The survey was put online using Qualtrics survey software. The video recorded interviews were uploaded to YouTube, using the strictest privacy settings, and embedded in the Qualtrics survey. Participants were recruited via social media. Friends and family were asked to complete the survey themselves via a Facebook page and status updates. They were also asked to invite their own friends and

family to the page, and to circulate the link to the survey to people they knew. An email link was sent to those who requested it.

Sample

One hundred and twenty-five participants completed the survey online. Completers who spent too much (> 24 hours) or too little time (less time than the videos took to watch) on the survey were removed from the sample (n = 22). This left 103 participants in the final sample, with 25 or 26 participants per condition.

The final sample ranged in age from 19 to 69 years (M = 39.47, SD = 15.73), the majority were female (64.1%), and described themselves as *English, Welsh, Scottish, Northern Irish, or British* (87% of the 100 participants who provided this information). All of the final sample stated they were jury eligible, but only 14 participants had ever been part of a real jury (13.6%). There were 13 students in the sample.

Design

The study had a 2 (rapport-building phase: shown vs. not shown) x 2 (description of number of interviews: two vs. one) between-participants design. For the within-subjects aspect of the design, the consistency of the detail (provided in clip one only, provided in clip two only, provided in both clips, contradicted in second clip) was added as a further independent variable.

Materials and Procedure

The Interview Videos

The two interviews used were selected from interviews conducted for a previous experimental study (Waterhouse, 2016). In this previous study, children had viewed a video-

recording of a non-violent theft and then experienced three interviews about the theft, conducted according to the UK's ABE guidelines (Ministry of Justice, 2011). All the interviews were conducted in a supportive manner. During the substantive section of each of these interviews, children were first asked to give their full free recall of the video-recording. This was elicited by asking 'Tell me everything you remember about the film you watched last week' and then using minimal prompters (such as echoing what the child had said, or nodding) to encourage further recall. When children had completed their free recall and stopped responding to prompters, the interviewer asked probing questions based on their free recall. So, for example, if a child had recalled a lady with a blue top, the interviewer would ask the child to tell them everything they could about the lady with the blue top. If necessary, the interviewer also asked wh- questions (who, what, when and where) to obtain further details. Option-posing questions were occasionally used for clarification and leading questions were avoided. Second and third interviews of the children were conducted in the same way.

Only the first two interviews were used for the present study. The first interview was conducted a week after the child viewed the event and the second a week later. A seven-year-old girl's video-recorded interviews were selected because she experienced rapport-building before each interview, looked similar in her first and second interviews (i.e., no change in hairstyle or school uniform), the interviews were conducted in the same room, and neither the child nor the interviewer referred to the delay between the two interviews. Additionally, the child provided mainly accurate information in both interviews but also some inaccurate details. In her second interview, she also gave some details she had previously given in her first interview (repeated details), some new accurate details, some new inaccurate details, and contradicted some of the details she gave in the first interview (i.e., said the victim wore a dress and tights in the first interview but leggings and a skirt in

the second interview). Parental consent was obtained both to conduct and video-record the interviews and to use the interviews in the present study.

The video-recorded interviews were anonymised by cutting sections in which the child's name or personal details (such as discussion of the child's family in the rapport-building section) were mentioned, and pixelating the school badge on the child's jumper. The rapport-building section was separated from the substantive section.

The Online Survey

Participants were first directed to a page where they were asked to give fully informed consent, followed by a page on which they were asked to confirm their UK jury-eligibility. Prior to watching the interviews, the participants were asked to provide demographic details, their profession, and state whether they had ever sat on a real jury.

Before viewing the interviews, participants were asked to treat the child's testimony as if she had actually witnessed the incident and was giving testimony in a criminal trial, about which they were given the following information:

> This testimony is part of a criminal trial for the alleged theft of the victim, Jade Richards', handbag by the defendant, Jon Ellis. It is alleged that Jon Ellis stole Jade Richards' handbag in Kingston, London, in the afternoon of the 15th November, 2014. The state is charging Jon Ellis with theft. The trial started after the defendant entered a plea of 'not guilty.' The videos you are about to watch consist of the investigative interview of Mary Lakes[1], a seven year old witness for the prosecution.

The next instructions differed according to the rapport-building condition.

[1] This is a pseudonym.

Participants in the Rapport Condition.

Participants who viewed the rapport-building part(s) of the interview(s) were informed that the interview(s) 'includes a rapport-building session prior to the discussion of the alleged theft, which you will be shown'. They then watched the rapport-building part. Prior to watching the substantive phase, they were given the following details:

> At this point, the interviewer checked Mary's understanding of truths and lies, which she understood, and explained to Mary a number of ground rules about the interview. The interview then continued with the interviewer asking her to tell everything she remembered about the event.

Participants in the No Rapport Condition.

Participants who did not view the rapport-building part(s) were given the following information:

> You will be shown the interview from the point at which Mary and the interviewer began discussing the theft. Just prior to this, the interviewer discussed Mary's recent trip to the aquarium with her in order for Mary to feel more relaxed and for her to get to know the interviewer and build a relationship. The interviewer also checked Mary's understanding of truth and lies, which she understood, and explained to Mary a number of ground rules about the interview. In total, this took five minutes. The interview then continued with the interviewer asking her to tell everything she remembered about the event.

These participants then watched the substantive part of the interview. Neither group, therefore, viewed the ground rules section of the interviews.

One Versus Two Interview Condition.

After having watched the first substantive phase, the participants were presented with a page, which either informed them that 'the following interview was conducted one week later' or that 'the interview was then paused for a 10 minute refreshment break. The interview then continued with the interviewer asking Mary to tell everything she remembered again.'

All participants viewed both substantive sections of the interview and those in the no rapport conditions only viewed these sections. However, for those in the rapport conditions, both groups viewed the first interview rapport-building, but only the two interview condition viewed both first and second rapport-building videos.

After the appropriate group had viewed the second rapport-building video, they were informed that 'Mary Lakes was then reminded of the ground rules, including the importance of telling the truth. The interview then continued, with the interviewer asking Mary to tell everything she remembered again.' They then watched the substantive phase of the interview.

Participants in the two interview x no rapport group were given the following information:

> The following interview was conducted one week later. After a discussion of Mary's recent walk in the park, to make her feel at ease with the interviewer, the interviewer asked Mary to tell everything she remembered again.

They then watched the substantive phase of the interview.

After the participants had watched the relevant interviews, they were presented with the perceptions questionnaire (see below). Following this, participants viewed a page that provided debriefing information and thanked them for their participation.

[Table 1 placed here]

Mock-Juror Perceptions Questionnaire

Participants were first asked questions about their general perceptions of the child witness (e.g., credible, accurate, anxious) and the interviewer (see Table 1). Responses were given on a ten-point Likert scale with characteristics provided at each end of the scale. Participants were also asked about their memories for specific comments the child made in the interview (which varied according to whether the statement was accurate, and when the child made it, see Table 2). The contradictions were details for which the child had provided conflicting information in an interview. For example, the child said the victim was wearing leggings and a skirt in the second interview, but in the first interview had said the victim was wearing a dress and tights. For each detail, the participants were first asked if they remembered the child mentioning the detail in her interview. If they did remember her saying it, they were then asked how likely (on a ten-point Likert scale) they felt it was that that detail had actually happened in the witnessed event. Finally, the participants were asked two questions about the case. First, 'If you were in charge of the prosecution's case (e.g., the side trying to persuade the jury that the defendant is guilty), how likely would you be to show these interview clips in court?' Participants responded on a ten-point Likert scale from Very Likely to Very Unlikely. The second question was 'If other evidence in the case was equally balanced, what would your verdict be based on the child's evidence?' to which participants could answer Guilty or Not Guilty.

After filling out these questions, the participants were asked to identify in which interview clip they remembered the child recalling each of the details in Table 2 (i.e., interview clip 1, 2, or both). They were then asked if the child had contradicted herself on any of these points. If they answered yes, they were asked to indicate how she had

contradicted herself. This was designed to determine whether participants remembered these contradictions without it necessarily affecting their initial perceptions of the child and her testimony.

[Table 2 placed here]

Pilot Study

A pilot study was conducted with six participants who were asked to complete the survey and comment on the clarity of the questions and information provided. Apart from clarifying that the participants could take notes (and including a question regarding whether they had or not), no changes to the questions themselves needed to be made. Therefore, these pilot participants' responses were included in the final sample.

Results

The first analysis examined how rapport and number of interviews might affect mock-juror perceptions. Additionally, a within-subject comparison of mock-juror perceptions of the reliability of specific recall statements according to their consistency within the interviews was conducted.

Data Adjustments.

Prior to running any analyses, participants' ratings of the child's testimony, the interviewer, their likelihood of using the video recordings in court, and the likelihood of specific details having occurred were reverse coded. This was for ease of interpretation and meant that higher scores now indicated positive responses (e.g., very believable, very likely) and lower responses indicated negative responses (e.g., very anxious, very inaccurate).

When analysed for the relevant groups, many of the child witness and interviewer perception dependent variables had non-normal distributions. This continued to be the case after three child witness ratings and two interviewer ratings were removed as outliers (z-scores greater than ±3.29). Non-normality was generally due to a large negative skew of the scores. Thus, the results were unlikely to meet the assumption of multivariate normality required for MANOVAs. However, Tabachnick and Fidell (2013a) state that MANOVA is robust to non-normality as long as the sample size in the smallest cell is at least 20. All cells continued to have a count of over 20, and therefore the non-transformed scores were entered into factorial MANOVAs and ANOVAs, irrespective of the lack of normality in their distributions.

Multicollinearity and Singularity.

Multicollinearity describes the situation in which dependent variables are highly correlated, and singularity describes a situation where two variables are so correlated that they may be measuring the same thing, making one of the variables redundant. According to Tabachnick and Fidell (2013b), statistical problems related to multicollinearity and singularity occur when the correlation between two variables (*r*) is greater or equal to .90. Interpretational problems occur when $r > .70$. In the current study, the majority of rated child witness variables' correlation scores are below .70. However, for the relationship between credibility and believability ($r = .767$) and truthfulness and believability ($r = .760$), the correlations are higher than .70. Despite this, all three variables were still included in the MANOVAs described below. This is because although truthfulness, believability, and credibility are similar concepts, the authors believe them to be different perceptions of the child, and their correlations are only high enough to cause interpretational issues and thus findings for these variables are interpreted with care. The correlations between the

interviewer perception scores were below the .70 threshold (*r*s < .574) and so all of these dependent variables were also included in the interviewer perception MANOVAs.

Group Comparisons

Perceptions of the Child Witness.

Mean non-transformed scores for each group are provided in Table 3. A factorial MANOVA was conducted with rapport condition (watched rapport-building phase x did not watch it) and interview number condition (presented as one interview x presented as two) as the independent variables, and participants' responses to the eight child witness questions (see Table 1) as the dependent variables.[2] Using Pillai's trace, there was a significant effect of rapport condition, $V = 0.18$, $F(8, 86) = 2.36$, $p = .024$, and a significant effect of the number of interviews, $V = 0.16$, $F(8, 86) = 2.08$, $p = .046$. The interaction between rapport condition and number of interviews was not significant, $V = 0.06$, $F(8, 86) = 0.64$, $p = .742$. Separate two-way ANOVAs revealed differences between groups for three of the eight dependent variables. For ratings of believability, both the number of interviews, $F(1, 98) = 14.99$, $p = .006$, $\omega = 0.25$, and rapport condition, $F(1, 98) = 5.19$, $p = .025$, $\omega = 0.20$, had an effect, although the effect size was small to medium for both. Participants who believed they were viewing two separate interviews rated the child as more believable ($M = 8.76$, $SD = 1.28$) than those who believed they were viewing one ($M = 8.00$, $SD = 1.52$), and participants who did not view the rapport-building rated the child as more believable ($M = 8.67$, $SD = 1.23$) than those who did watch the rapport-building ($M = 8.06$, $SD = 1.61$). Participants' credibility ratings were also affected by whether they saw the rapport-building phase(s) or not, $F(1, 99) = 6.35$, $p = .013$, $\omega = 0.23$; although there was only a small to medium effect

[2] The MANOVAs and ANOVAs for child and interviewer perceptions and likelihood of using the video clips in court were also run with bootstrapping to overcome the normality issues described above. Bootstrapping provided the same results and so, for simplicity, the results without bootstrapping were reported here.

size, those who did not watch the rapport-building phase(s) rated the child as more credible (M = 8.27, SD = 1.65) than those who did (M = 7.37, SD = 1.95). Additionally, participants' ratings of the child's truthfulness were affected by whether they believed they were watching one interview with a ten minute break, or two separated by a week, $F(1, 98) = 6.08$, $p = .015$, $\omega = 0.22$ (indicating a small to medium effect size); those in the latter condition rated children as more truthful (M = 8.82, SD = 1.10) than those in the former (M = 8.13, SD = 1.65). Thus, number of interviews and whether the participants saw the rapport-building or not affected a number of mock-jurors' perceptions of the child witness.

[Table 3 placed here]

Perceptions of the Interviewer.

Mean scores by group are provided in Table 3. A further factorial MANOVA was conducted with the same independent variables and participants' responses to the three interviewer perception questions (see Table 1) as the dependent variables. Pillai's trace results were non-significant for rapport, $V = 0.02$, $F(3, 95) = 0.63$, $p = .598$, number of interviews, $V = 0.01$, $F(3, 95) = 0.17$, $p = .915$, and the interaction between these variables, $V = 0.06$, $F(3, 95) = 1.99$, $p = .121$. Thus, perceptions of the interviewer were not affected by whether the participant saw the rapport-building phase or not, or whether they thought the video clips were from one interview with a short break or two interviews separated by a week.

Case Progression Perceptions.

Table 3 shows the average scores for each group's ratings of how likely they would be to use the clips in court if they were in charge of prosecution, and the percentage of

participants who stated that if all other evidence was equal, based on the child's videos they would vote guilty in a trial. A factorial ANOVA with the same independent variables was conducted with the dependent variable of mock-jurors' responses regarding how likely they would be to use the video clips in court. There was no significant effect of watching the rapport-building phase, $F(1, 97) = 1.60$, $p = .209$, or believing the interviews to be one continued interview or two separate ones, $F(1, 97) = 0.69$, $p = .407$. However, there was a significant interaction between the two, $F(1, 97) = 5.24$, $p = .024$. The interaction graph (see Figure 1) shows that participants who believed they were watching one interview with a ten minute break were more likely to use the video clips in court if they had not seen the rapport-building video clips. For those in the two interview condition, viewing the rapport-building made less difference to their perceived likeliness to use the video clips in court.

[Figure 1 placed here]

A three-way loglinear analysis was conducted to determine the relationship between rapport, number of interviews, and type of verdict. The expected counts for this analysis reduced the power of the results (four of the eight cells had an expected count of approximately four). However, the final model retained only type of verdict, suggesting that none of the higher-order 3-way or 2-way interactions, nor rapport or number of interviews condition were significant, $\chi^2(1)$s < 2.90, ps $> .089$. Verdict was the only significant predictor of the number of participants in each cell, $\chi^2(1) = 47.60$, $p < .001$. An odds ratio indicated that the odds were 5.06 higher that participants would give a guilty verdict than a non-guilty one. Thus, the model of best fit was to assume all participants had voted guilty. This finding was confirmed by separate chi-square tests conducted to overcome power issues.

Neither number of interviews ($\chi^2(1) = 0.002$, $p = .964$) nor rapport-building condition ($\chi^2(1) = 2.85$, $p = .092$) had significant associations with verdicts.

Summary.

Watching the rapport-building section of interview(s) or believing the video clips came from one interview or two had no effect on perceptions of the interviewer or verdict. The presentation format of the interviews, however, did have an effect on perceptions of the child witness and participants' perceived likelihood of using the video clips in court. Those that thought they were watching two interviews of the child had some more positive perceptions of the child than those that thought they were watching one interview with a short break, and those who did not watch the rapport-building phase of the interview(s) also perceived her more positively than those who did. Those who watched one interview and no rapport-building were most likely to use the video clips in court, whereas those who watched one interview with rapport-building were the least likely to.

Consistency and Believability

On average, participants remembered the child recalling 5.95 of the seven specific details they were questioned about (see Table 2). For the majority of details, at least half of the participants remembered when these details were given correctly (see Table 4). However, a large number of participants did not remember the child saying the 'nice' man had short hair ($n = 44$). When asked whether the child had contradicted herself, 65.7% of participants (correctly) stated she did. Fifty-two participants (50.5%) correctly stated she was contradictory about the victim's clothing, 44 participants (42.7%) incorrectly stated she was contradictory about the thief stealing credit cards, and 32 (31.07%) correctly stated she was contradictory about whether the victim went to get the police. Fewer than ten participants

incorrectly stated the child was contradictory about each of the other specific details asked about (for which she had, in fact, not been contradictory).

In order to examine whether details that were repeated were thought to be more likely to have occurred than those mentioned only once, within-subject comparisons were conducted for participants' responses to the questions about the likelihood of specific details occurring. Thirty-two participants remembered the child recalling all of the seven details and responded to all of the follow-up questions regarding how likely they thought the detail was to have happened. Mean scores for these participants for each detail are provided in Table 4. Within the interview condition groups, only one of the details' ratings were non-normally distributed, and no outliers were identified.

[Table 4 placed here]

To determine whether the perceived likelihood of the details significantly differed from each other or were affected by whether the participants had been informed they had watched two interviews or one, the likelihood scores were entered into a mixed design ANOVA. The 'number of interviews' condition was entered as the between-subject independent variable and the details were entered as the within-subject independent variable.[3] Participants' likelihood scores for each detail were entered as the dependent variable. The mixed design ANOVA found a significant main effect of detail, $F(6, 180) = 6.12$, $p < .001$, but no main effect of number of interviews, $F(1, 30) = .124$, $p = .728$, nor any significant interaction, $F(6, 180) = 1.70$, $p = .123$. Pairwise comparisons with Bonferroni-corrected *p*-values showed that there were four pairs of significantly different scores. One detail

[3] The same results were obtained in a separate mixed design ANOVA with reflected square root transformed scores which normalised the distribution for all details.

provided in interview one only (that the 'nice' man had short hair) was perceived as significantly more likely to have happened than either of the details that had been contradicted (that the 'victim' had gone to get the police, estimated mean difference = 2.14, *SE* = 0.43, 95% CI [0.72, 3.55], $p < .001$, or that the lady was wearing leggings and a skirt, estimated mean difference = 1.37, *SE* = 0.38, 95% CI [0.10, 2.63], $p = .025$). This same detail (the 'nice' man had short hair) was also rated as more likely to have happened than the detail only provided in interview two (that the thief took the lady's credit cards, estimated mean difference = 1.54, *SE* = 0.37, 95% CI [0.32, 2.77], $p = .005$). The final significant difference was between the other piece of information provided in interview one only (that the bag stolen was beige) and the contradiction mentioned in interview one (that the 'victim' had gone to get the police). The information provided in interview one that was not later contradicted was scored as more likely to have happened than the detail that was later contradicted (estimated mean difference = 1.63, *SE* = 0.44, 95% CI [0.18, 3.09], $p = .017$).

Summary.

Participants generally remembered the majority of details that they were asked about. They also mostly correctly remembered when they had heard these details (i.e., in which video clip), but their memory for contradictions was less reliable. Additionally, details provided in the first interview, in particular that the 'nice' man had short hair, were viewed as more reliable than either of the contradicted details and the non-contradicted detail provided for the first time in interview two.

Discussion

In the present study, mock-juror perceptions of a child's testimony and interview were very positive and the majority of participants would have given a 'guilty' verdict in response

to the child's testimony. Whether the mock-juror viewed the rapport-building section(s) of the interview and whether they were told the interview was one session with a ten-minute break or two sessions with a week long delay had no effect on their opinions of the interviewer, or their verdicts. However, some small to medium effects were found. Participants who were told there were two interviews separated by a week had more positive views of the child (more believable and truthful) than those who were told there was one interview with a ten-minute break. Those who did not watch the rapport-building phase also had more positive views of the child (more believable and credible) than those who did watch this phase. Viewing the rapport-building phase also seemed to affect whether participants in the one interview condition would show the video-recorded interviews in court; those who watched the rapport-building video reported they were less likely to show the video clips in court than those who had not watched the rapport-building video. Finally, participants did believe some specific details provided by the child more than others, and in particular those provided only once in the first interview clip, but this was not affected by the number of interviews condition.

These findings are somewhat different from Krähenbühl's (2012) study, which found that reading the rapport-building section of a child interview transcript affected mock-juror perceptions of the child's testimony and understanding positively, as well as their perceptions of the interview itself. The present study, on the other hand, similar to Bottoms *et al.* (2004, cited in Bottoms *et al.*, 2007), found no effect of watching the rapport-building section of an interview on most of the participants' perceptions except to reduce their perceptions of the child's believability and credibility (two highly related concepts). An important difference in the methodologies used is that Krähenbühl's rapport-building included the verbatim transcript of the 'ground rules' and 'truth and lies' sections of the interview along with a neutral rapport-building discussion. The participants in the present study, on the other hand, only

read brief referrals to 'ground rules' plus 'truth and lies' and instead just viewed the neutral discussion section of the rapport-building phase. Krähenbühl's suggested explanation of why rapport-building affected perceptions was that mock-jurors were exposed to more signs of effective communication, such as a shared understanding of the purpose of the interview and how it was meant to proceed. If this is the cause of the positive effects Krähenbühl found, then it could be due to her participants' exposure to the 'ground rules' section specifically, in which the interview and the child's role within it is explained. The present study's findings suggest that it is unlikely to have been exposure to the neutral rapport-building discussion that improved perceptions. Instead, exposure to supportive interviewing in the form of the neutral rapport-building discussion (as in the present study) or a generally supportive interviewing style (Bottoms *et al.*, 2004, cited in Bottoms *et al.*, 2007), seems to lead to less positive perceptions. In order to determine whether these differences are caused by exposure to differing rapport-building sections, further research comparing mock-jurors' perceptions of child testimony presented with or without the relevant sections (i.e., the neutral discussion, 'ground rules', truth and lies) should be conducted. This is important given the view of judges that jurors should not watch 'over-long' child interviews (Burrows & Powell, 2014).

Krähenbühl's alternative explanation of the effect of rapport-building (that the increased length of the transcript affected perceptions rather than the content) is also not supported by the present study's findings. The condition in which the participants spent most time watching the interview(s) (two interviews x rapport-building shown) did not result in significantly different perceptions from the other conditions. This is supported by research which has found interview length to have no effect on jury decision-making in real cases (Wilson & Davies, 1999). Alternatively, the change in medium (i.e., from reading a transcript to watching an interview) may explain the difference in results. Although previous studies have found no significant differences between mock-juror perceptions of written

mock-trials and video-recorded ones (Goodman, Golding, Helgeson, Haith, & Michelli, 1987), Krähenbühl's (2012) participants may have been influenced by the addition of the verbatim rapport-building section because they had no visual cues as to how comfortable the child was within the interview. In the present study, involving video recordings, participants had more cues (including non-verbal ones) to indicate how well the child understood what was going on during the substantive part of the interview, and so the rapport-building itself may not have been so integral to participants' perceptions of the child. The child may also have provided visual cues in the rapport-building which led the participants to believe she was less trustworthy, or, as Bottoms *et al.* (2007) suggest, support (in the form of rapport-building) may appear coercive or reduce sympathy for the child interviewee. This highlights a methodological issue with the present study that limits generalisability. Namely, that all participants viewed the same interviews. Although this methodological decision allows standardisation of extra-legal factors which may affect mock-jurors' perceptions (such as the child's age, gender, or non-verbal communication), it also means these findings may be unique to this set of interviews and/or children of this age (seven years old). Research by Crane, *et al.* (2018), for example, has found perceptions of child witnesses with autism to vary from child to child. Thus, a useful next step would be to compare responses to different children's interviews in which they recall an event they directly experienced and to additionally measure mock-jurors' sympathy for the child.

Further limitations of the present study include the homogeneity of the sample which may not reflect the variety of people who take part in juries; very few participants were non-British or had experience of sitting on a real jury. On the other hand, participants were a community sample, with only a few students. Additionally, the specific wording in the no-rapport condition may have implied that rapport had successfully been built and thus given a better impression of the rapport between the child and the interviewer than watching the

rapport-building session did. Therefore, this wording could explain the rapport-building results. The interviewer had, however, been trained in ABE interviewing techniques by police trainers and so the rapport-building was conducted in an ecologically valid way. The participants were also not informed that the rapport-building phase was included in the child witness interview following governmental guidelines. Therefore, incorrect assumptions may have been made regarding why the rapport-building was conducted which may have affected perceptions. Future participants should be told that this follows best practice to avoid such assumptions being formed.

The choice of crime reduces ecological validity; children's testimony is less likely to be as vital in theft cases as in child sexual abuse cases, which come with their own additional difficulties (including emotional responses and lack of corroborating evidence). Perceptions of children's testimony may also be more cautious in these kinds of cases, whereas in the present study, perceptions of the child and interviewer were very positive, with some possible ceiling effects. Despite this risk, a theft case was chosen to ensure children were protected from harm that could have been caused by a more stressful crime event. However, in future research it would be beneficial to use a crime or testimony that may be more ambiguous and allow more variety in perceptions, thus avoiding ceiling effects.

A novel and important finding of the present study was that participants who believed they were viewing multiple interviews (i.e., two interviews separated by a week) rated the child as more believable and truthful than those that thought they were watching one interview with a ten-minute break. Although these two variables are highly related, this suggests that information provided across two interviews is not immediately perceived as less reliable, and that instead, a child's inconsistencies may be viewed as more understandable and judged more positively when they are provided across two interviews with a delay than when they are perceived as occurring within the same interview. Therefore, research that has

found that mock-jurors perceive inconsistencies as indications of inaccuracy (Leippe *et al.*, 1992; Quas *et al.*, 2005) and the 'courtroom theory' of memory (Fisher *et al.*, 2009) where all inconsistent information is presumed to be obtained via non-crime information may be over-simplifying the situation; mock-jurors may take into account aspects of the interviewing conditions that could affect likely consistency (such as delay between interviews).

Alongside examining how video presentation affected mock-jurors' perceptions, this study also looked at how believable participants perceived specific details within the child's testimony to be. Of particular interest was whether the timing of the detail (i.e., provided in clip one, clip two, or both) and how consistently it was provided (i.e., once, repeated, or contradicted) affected perceptions. The sole significant differences were that two details provided only in interview one (one correct and one incorrect) were perceived as being more likely to be true than the contradicted details and a detail provided only in interview two. Although this suggests that details provided in interview one were perceived as more reliable than other details (which could reflect a memory primacy effect, Ashcraft, 2006), three of the four significant differences related to one detail. This detail was that the 'nice' man had short hair. It may be that the participants did not necessarily believe this more because it was provided only once in the first interview (which many of them did not correctly remember), but because this detail fitted their own schema of how a 'nice' man would look, and so was quite a 'safe bet' in comparison to the more specific other details.

Further interesting findings are that repeated details were perceived as no more believable than other details and that participants did not rate contradictions as less reliable than other details (except the two details given in interview one only, discussed above). This result opposes previous findings that suggest that inconsistency is perceived as an indicator of unreliability (Leippe *et al.*, 1992; Quas *et al.*, 2005), the 'courtroom theory' of memory (Fisher *et al.*, 2009), and that interviews with repeated details are perceived more positively

(Yozwiak *et al.*, 2004). This could be explained by methodological differences; the present study was the only one to expose participants to full interviews and to ask them to make decisions on single details, or, as suggested by Leippe *et al.* (1992), mock-jurors' perceptions of inconsistency may not actually relate to what the child says, but to a more general perception of the child. Alternately, the findings regarding contradictions may be due to varying definitions of 'contradiction'. When participants described the contradictions that they remembered, some described omission errors (i.e., recalling a detail in one interview and not the other). This suggests they may interpret these inconsistencies as contradictions and therefore equally unreliable (as suggested in Fisher *et al.*, 2009). However, as participants did not rate repeated details as any more reliable, it seems mock-jurors do not judge reliability of a specific detail solely on when it was stated or how often it was repeated, irrespective of whether there were one or two interviews.

Finally, in the present study, length of the substantive section on mock-jurors' perceptions was not manipulated; both single and multiple interview conditions had equally long substantive sections. This is unlikely to reflect real cases, where some second interviews are conducted in addition to a full first investigative interview (e.g., as recommended by the ABE, Ministry of Justice, 2011). Thus, real cases involving multiple interviews of children are likely to involve longer interviews for the jurors to watch and may also involve additional inconsistencies than single interviews. Therefore, studies that compare multiple to single interview perceptions without controlling for length or number of contradictions would be useful additions to the literature.

Conclusions

This study provides an examination of mock-juror perceptions in response to a single interview of a child in comparison to two interviews. The findings indicate that mock-jurors

do not disbelieve details provided by a child over two interviews more than those provided in one, and instead have slightly more positive views of the child's testimony when they were told it was given across multiple interviews. Additionally, the present study found that viewing the rapport-building sections of these interviews can actually degrade mock-jurors' pre-deliberation perceptions of the child. Thus, showing multiple interviews in court as the child's evidence-in-chief may not negatively affect jurors' perceptions, but including the rapport-building phase may.

References

Anderson, L., Gross, J., Sonne, T., Zajac, R., & Hayne, H. (2016). Where there's smoke, there's fire: The effect of truncated testimony on juror decision-making. *Behavioral Sciences and the Law, 34,* 200-217. doi: 10.1002/bsl.2212

Ashcraft, M. H. (2006). Short-term working memory. In M. H. Ashcraft (Ed.), *Cognition* (4th ed.). Pearson Education: New Jersey, USA.

Bottoms, B. L., Quas, J. A., & Davis, S. L. (2007). The influence of the interviewer-provided social support on children's suggestibility, memory, and disclosures. In M.-E. Pipe, M. E. Lamb, Y. Orbach & A.-C. Cederborg (Eds.), *Child sexual abuse: Disclosure, delay, and denial* (pp. 135-157). Oxon, UK: Routledge.

Burrows, K. S., & Powell, M. (2014). Prosecutors' recommendations for improving child witness statements about sexual abuse. *Policing and Society: An International Journal of Research and Policy, 24,* 189-207. doi: 10.1080/10439463.2013.784305

Carnes, C. N., Nelson-Gardell, D., Wilson, C., & Orgassa, U. C. (2001). Extended Forensic Evaluation when sexual abuse is suspected: A multisite field study. *Child Maltreatment, 6,* 230-242. doi: 10.1177/1077559501006003004

Crane, L., Wilcock, R., Maras, K. L., Chui, W., Marti-Sanchez, C., & Henry, L. A. (2018). Mock juror perceptions of child witnesses on the autism spectrum: The impact of providing diagnostic labels and information about autism. *Journal of Autism and Developmental Disorders.* Advance online publication. doi: 10.1007/s10803-018-3700-0

Crown Prosecution Service. *Legal guidance. Rape and sexual offences: Chapter 6: Special measures and video evidence*. https://www.cps.gov.uk/legal-guidance/rape-and-sexual-offences-chapter-6-special-measures-video-evidence

Davies, G., Wilson, C., Mitchell, R., & Milsom, J. (1995). *Videotaping children's evidence: An evaluation* (Research Findings No. 20). London, UK: Home Office Research and Statistics Department.

Fisher, R. P., Brewer, N., & Mitchell, G. (2009). The relation between consistency and accuracy of eyewitness testimony: Legal versus cognitive explanations. In R. Bull, T. Valentine, & T. Williamson (Eds.), *Handbook of psychology of investigative interviewing: Current developments and future directions* (pp. 121-136). Chichester, UK: John Wiley & Sons, Ltd.

Goodman, G. S., Golding, J. M., Helgeson, V. S., Haith, M. M., & Michelli, J. (1987). When a child takes the stand: Jurors' perceptions of children's eyewitness testimony. *Law and Human Behavior, 11,* 27-40. doi: 10.1007/BF01044837

Hershkowitz, I., & Terner, A. (2007). The effects of repeated interviewing on children's forensic statements of sexual abuse. *Applied Cognitive Psychology, 21,* 1131-1143. doi: 10.1002/acp.1319

Katz, C., & Hershkowitz, I. (2012). Repeated interviews with children who are the alleged victims of sexual abuse. *Research on Social Work Practice, 23,* 210-218. doi: 10.1177/1049731512467511

Korkman, J., Pakkanen, T., & Laajasalo, T. (2017). Child Forensic Interviewing in Finland: Investigating Suspected Child Abuse at the Forensic Psychology Unit for Children and Adolescents. In S. Johansson, K. Stefansen, E. Bakketeig, & A. Kaldal (Eds.), *Collaborating against child abuse: Exploring the Nordic Barnahus model*. London, UK: Palgrave Macmillan.

Krähenbühl, S. (2012). 'Does the jury really need to hear it all?': The effect of evidence presentation practice on jury assessment of children's eyewitness testimony. *Psychology, Crime & Law, 18,* 847-858. doi: 10.1080/1068316X.2011.579904

Leippe, M. R., Manion, A. P., & Romanczyk, A. (1992). Eyewitness persuasion: How and how well do fact finders judge the accuracy of adults' and children's memory reports? *Journal of Personality and Social Psychology, 63,* 181-197. doi: 10.1037/0022-3514.63.2.181

Ministry of Justice (2011). *Achieving Best Evidence in Criminal Proceedings: Guidance on Interviewing Victims and Witnesses, and Guidance on Using Special Measures.* Retrieved from: http://www.justice.gov.uk/downloads/victims-and-witnesses/vulnerablewitnesses/achieving-best-evidence-criminal-proceedings.pdf.

Peixoto, C. E., Ribeiro, C., Fernandes, R. V., & Almeida, T. S. (2015). Forensic interviewing of witnesses in Portugal. In D. Walsh, G. E. Oxburgh, A. D. Redlich, & T. Myklebust (Eds.), *International developments and practices in investigative interviewing and interrogation. Volume 1: Victims and witnesses* (pp. 188-198). Abingdon, UK: Routledge.

Plotnikoff, J., & Woolfson, R. (2001). *An evaluation of child witness support.* The Scottish Executive Central Research Unit. Retrieved from http://www.scotland.gov.uk/Publications/2001/

Quas, J. A., Thompson, W. C., & Clarke-Stewart, K. A. (2005). Do jurors "know" what isn't so about child witnesses? *Law and Human Behavior, 29,* 425-455. doi: 10.1007/s10979-005-5523-8

Tabachnick, B. G., & Fidell, L. S. (2013a). Multivariate analysis of variance and covariance. In B. G. Tabachnick & L. S. Fidell (Eds.), *Using multivariate statistics* (6th ed., pp. 245-313). New Jersey, USA: Pearson Education, Inc.

Tabachnick, B. G., & Fidell, L. S. (2013b). Cleaning up your act: Screening data prior to analysis. In B. G. Tabachnick & L. S. Fidell (Eds.), *Using multivariate statistics* (pp. 60-116). New Jersey, USA: Pearson Education, Inc.

Waterhouse, G. F. (2016). *Investigating the Forensic Interviewing of Children: Multiple Interviews and Social Support.* (Unpublished doctoral dissertation). London South Bank University, England.

Waterhouse, G. F., Ridley, A., Bull, R., La Rooy, D. J., & Wilcock, R. (2016). Dynamics of repeated interviews with children. *Applied Cognitive Psychology, 30,* 713-721. doi: 10.1002/acp.3246

Wilson, J. C., & Davies, G. M. (1999). An evaluation of the use of videotaped evidence for juvenile witnesses in criminal courts in England and Wales. *European Journal on Criminal Policy and Research, 7,* 81-96. doi: 10.1023/A:1008740231642

Yozwiak, J. A., Golding, J. M., & Marsil, D. F. (2004). The impact of type of out-of-court disclosure in a child sexual assault trial. *Child Maltreatment, 9,* 325-334. doi: 10.1177/1077559504266518

Table 1.

Mock-Juror Perceptions Questions

Questions	Likert Scales	
	1	10
Child Witness		
How believable was the child witness?	Very Believable	Unbelievable
How credible do you think the child witness was?	Very Credible	Very uncredible
How accurate do you think the child was?	Very accurate	Very inaccurate
How truthful do you think the child was?	Very truthful	Very untruthful
How clear was the child's testimony?	Very clear	Very unclear
How anxious do you think the child was when they were interviewed first?	Very calm	Very anxious
How anxious do you think the child was when they were interviewed after the break?	Very calm	Very anxious
How well do you think the child understood the questions asked to them?	Understood completely	Did not understand at all
Interviewer		
How fair do you think the interviewer was being to the child?	Very fair	Very unfair
How friendly do you think the child's interviewer was?	Very friendly	Very unfriendly
How clear do you think the questions the child was asked were?	Very clear	Very unclear

Note. Responses were all reverse scored in the results section for ease of interpretation.

Table 2.

Accuracy and Consistency for Specific Detail Questions.

Detail in Question	**Accuracy**	**Consistency**
Lady said ‘Oh no’ when she got her bag back.	Correct	Repeated
Bag that was stolen was beige.	Incorrect	Interview One only
The ‘nice’ man that gave the bag back had short hair.	Correct	Interview One only
The victim (whose bag was stolen) went to get the police.*	Contradiction	Interview One only
The thief and the ‘nice’ man were by a factory.	Incorrect	Repeated
The thief took credit cards out of the lady’s bag.	Correct	Interview Two only
The victim was wearing leggings and a skirt.*	Contradiction	Interview Two only

* Contradictions are details that were given in one interview clip and directly contradicted in the other interview clip. For the first contradiction in this table, the contradicting detail provided in the second interview clip was that the victim did not get the police. For the second contradiction in this table, the contradicting detail provided in the first interview clip was that the victim was wearing a dress and tights.

Table 3.

Non-transformed Mean Scores (and Standard Deviations) for Child Witness and Interview Perceptions.

	Rapport-building video		**No rapport-building video**	
	One interview	**Two interviews**	**One interview**	**Two interviews**
Child Witness Perceptions				
Believable	7.73 (1.66)	8.40 (1.50)	8.26 (1.35)	8.88 (1.51)
Credible	7.04 (1.99)	7.72 (1.88)	8.15 (1.68)	8.40 (1.63)
Accurate	6.85 (1.57)	7.16 (1.68)	7.11 (1.99)	7.44 (1.61)
Truthful	7.81 (2.00)	8.96 (1.10)	8.26 (1.53)	8.68 (1.11)
Clear	7.15 (2.09)	7.96 (1.21)	7.31 (2.11)	7.68 (1.68)
Anxious in First Clip	5.04 (2.09)	5.58 (2.02)	5.33 (1.78)	5.92 (1.89)
Anxious in Second Clip	7.42 (1.55)	6.84 (1.86)	7.04 (1.37)	7.20 (1.50)
Understood	8.73 (1.12)	9.24 (0.66)	8.78 (1.31)	8.96 (1.02)
Interviewer Perceptions				
Fair	9.12 (0.82)	9.40 (0.96)	9.52 (0.64)	9.36 (0.70)
Friendly	9.23 (1.07)	9.40 (0.87)	9.19 (1.42)	9.28 (0.79)
Clear Questions	8.58 (1.24)	9.04 (1.02)	9.07 (1.44)	8.60 (1.19)
Case Progression Perceptions				
Likelihood of Using Clips in Court	7.00 (2.28)	7.96 (2.44)	8.22 (1.78)	7.84 (2.19)
Percentage Guilty Verdict	79.2%	75.0%	88.0%	91.7%

Note. All scores were on a scale of one to ten and reverse scored so one indicated the least positive response (e.g., Not friendly at all, Very unlikely, or Very anxious) and ten indicated the most positive (e.g., Very clear, or Very friendly).

Table 4.

Mean Scores (and Standard Deviations) of Likelihood of Specific Events to have Occurred.

Detail	Type of Detail	Correctly Recalled n (%)	Likelihood of Having Occurred *M* (*SD*)
Lady said 'Oh no' when she got her bag back.	Correct / Repeated	81 (78.6)	7.09 (2.41)*
Bag that was stolen was beige.	Incorrect / Clip One only	71 (68.9)	7.63 (1.43)*
The 'nice' man that gave the bag back had short hair.	Correct / Clip One only	24 (23.3)	8.13 (1.41)*
The victim (whose bag was stolen) went to get the police.	Contradiction / Clip One only	63 (61.2)	6.03 (2.25)*
The thief and the 'nice' man were by a factory.	Incorrect / Repeated	67 (65.0)	6.81 (2.16)*
The thief took credit cards out of the lady's bag.	Correct / Clip Two only	91 (88.3)	6.59 (1.81)*
The victim was wearing leggings and a skirt.	Contradiction / Clip Two only	65 (63.1)	6.78 (1.31)*

Note. Likelihood scores provided on a ten-point Likert scale, with ten signifying 'Very likely' and one 'Very unlikely'.

* $n = 32$

Figure 1.

Graph showing interaction between rapport and number of interview conditions on average Likert score for likelihood of using video clips in court.