AARHUS UNIVERSITY

Coversheet

This is the accepted manuscript (post-print version) of the article.

Contentwise, the post-print version is identical to the final published version, but there may be differences in typography and layout.

How to cite this publication

Please cite the final published version:

Anderson, L., Gross, J., Sonne, T., Zajac, R., and Hayne, H. (2016) Where There's Smoke, There's Fire: the Effect of Truncated Testimony on Juror Decision-making. *Behav. Sci. Law*, 34: 200–217.
doi: [10.1002/bsl.2212](https://doi.org/10.1002/bsl.2212).

Publication metadata

Title:	Where there's smoke, there's fire: The effect of truncated testimony on juror decision-making
Author(s):	Anderson, L., Gross, J., Sonne, T., Zajac, R., & Hayne, H.
Journal:	Behavioral Sciences & the Law
DOI/Link:	10.1002/bsl.2212
Document version:	Accepted manuscript (post-print)

General Rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Where there's smoke, there's fire: The effect of truncated testimony on juror decision-making

Lakin Anderson¹, Julien Gross¹, Trine Sonne², Rachel Zajac¹, and Harlene Hayne¹

¹*Department of Psychology, University of Otago, Dunedin, New Zealand*

²*Center on Autobiographical Memory Research, Department of Psychology and Behavioral Sciences, Aarhus University, Aarhus, Denmark*

The research was funded by Marsden grants (09-UOO-074, 15-UOO-061) from the Royal Society of New Zealand to HH. TS was supported by a grant from the Danish National Research Foundation (DNRF93).

Corresponding Author:
Professor Harlene Hayne
Department of Psychology
University of Otago, PO Box 56, Dunedin
New Zealand 9054
Email: hayne@psy.otago.ac.nz

Abstract

In countries that allow child complainants of abuse to present their direct evidence via pre-recorded videotape, the recording is sometimes truncated for relevance or admissibility purposes before it is presented to the jury. In two experiments, we investigated how this practice affects mock jurors' judgements of child credibility and defendant culpability when truncation omitted the child's less plausible allegations. Mock jurors read a transcript of a 6-year-old girl making an abuse allegation against the janitor at her school. Some jurors read this allegation only (truncated version), while others also read either one or two additional—but less plausible—allegations by the same child. Contrary to what we predicted, the presence of these additional allegations did not decrease jurors' belief in the core allegation, nor did it influence their judgments about the child complainant's honesty or cognitive competence. In fact, under at least one condition, reading additional, less plausible allegations made jurors *more* likely to pronounce the defendant guilty of the core allegation—even when jurors did not believe the additional allegations. This finding stands in stark contrast to prior research on jurors' evaluation of adults' testimony that includes implausible details. Future research in this area will help to elucidate the conditions under which the presentation of truncated testimony may or may not influence juror decision-making.

Around the world, experts have long raised concerns about the way that child witnesses are treated in the criminal court (Bala, 1999; Bennett, 2003; Brennan & Brennan, 1988; Cashmore & Bussey, 1996; Cashmore & Trimboli, 2006; Eastwood & Patton, 2002; Gupta, 1994; Landstrom & Granhag, 2010; Pipe & Henaghan, 1996; Powell, Wright, & Hughes-Scholes, 2011; Scott, 1994; Zajac, O'Neill, & Hayne, 2012). Although most witnesses find testifying in court unpleasant, the experience can be particularly stressful for children, who often hold misperceptions about the legal process and may be extremely intimidated by the formality of the proceedings (Bala, 1999; Cashmore, 1992; Davies, 1992; Flin, 1990; Goodman et al., 1992; Hall & Sales, 2008; Hill & Hill, 1987; Murray, 1995; Pipe & Henaghan, 1996; Powell et al., 2011). Researchers have argued that high levels of stress and confusion might limit children's ability to provide complete and accurate accounts of their experiences (Cashmore, 1992; Davies & Noon, 1991; Goodman et al., 1998; Tobey, Goodman, Batterman-Faunce, Orcutt, & Sachsenmaier, 1995)—especially when coupled with the long delays that are commonplace between a criminal allegation and the trial (Connolly & Read, 2006; Hanna, Davies, Henderson, Crothers, & Rotherham, 2010; Klemfuss & Ceci, 2009).

In light of these concerns, legislative changes in many countries permit special provisions for children to provide their evidence in court. New Zealand, the country where we conduct our research, was the first common law country in the world to adopt a suite of legislative changes that were specifically designed to address some of the difficulties that child witnesses face. Among these changes was a provision that allows a witness under the age of 18 who is alleging sexual abuse to provide his or her evidence-in-chief via a pre-recorded evidential interview conducted by a specialist interviewer (Evidence Amendment Act, 1989). As well as reducing the time that children spend answering lawyers' questions, pre-recording children's primary evidence reduces the memory demands that are imposed by long delays because the interview can be conducted shortly after an allegation is made. For these reasons, most child witnesses in New Zealand criminal cases now give their evidence-in-chief in this way (Ministry of Justice, 2010). The legislation also makes provision for children to give evidence from behind a screen so that the defendant is blocked from their line of sight, or to testify from another room outside of the courtroom, via closed-circuit television (CCTV).

These changes were applauded by many and led to similar changes in other countries (Bala, 1999; Gupta, 1994; McDonald & Tinsley, 2011; see Hanna et al., 2010, for a review of relevant legislation from other jurisdictions). Consistent with at least one purpose of the legislative changes, a number of researchers from around the world have now shown that

providing testimony by alternative means reduces children's stress (e.g., Cashmore, 1992; Davies, 1999; Davies & Noon, 1991; Davies, Wilson, Mitchell, & Wilson, 1995; Goodman et al., 1998; Landstrom & Granhag, 2010; Murray, 1995; Nathanson & Saywitz, 2003; Saywitz & Nathanson, 1993). But what impact does this practice have on the jurors who are asked to evaluate children's testimony? Because children are most often called to testify about suspected sexual abuse—in which there is typically no physical evidence or corroborating eyewitnesses—jurors' interpretation of the child's testimony is the most crucial element in the trial (Connolly, Price, Lavoie, & Gordon, 2008; Pezdek et al., 2004).

Research findings concerning the effect of alternative modes of evidence on jurors' evaluations of children's evidence are somewhat mixed. Some studies have shown that mock jurors' ratings of child witnesses' credibility do not differ according to the mode by which the evidence is presented (e.g., Landstrom & Granhag, 2010; Ross, Hopkins, Hanson, Lindsay, Hazen, & Eslinger, 1994; Swim, Borgida, & McCoy, 1993). In contrast, other research has shown that mock jurors rate children who testify via alternative means as less confident, less believable, less honest, less convincing, and less accurate than children who do not (e.g., Goodman et al., 1998, 2006; Orcutt, Goodman, Tobey, Batterman-Faunce, & Thomas, 2001; Tobey et al., 1995; Landstrom, Granhag, & Hartwig, 2007). In fact, in some studies, mock jurors have rated the defendant as less likely to be guilty when a child complainant's testimony is presented via alternative means (Eaton, Ball, & O'Callaghan, 2001; Goodman et al., 1998; but see Davies et al., 1995; Hanna et al., 2010). Given that the primary purpose of legislation enabling children to testify via special measures was to improve the quality and quantity of their testimony, it is concerning that these same measures could decrease the impact of that testimony in the eyes of the jury.

There is another concern about the effect of pre-recorded evidence on decision-making: jurors do not necessarily see the child's evidence in its entirety. In New Zealand, portions of a child's pre-recorded evidential interview—or the transcript of it—can be presented to the jury without showing the entire interview. Decisions to edit testimony are made based on two factors: 1) whether the probative value of the content is outweighed by its potential to be prejudicial; and 2) whether the content is relevant to proceedings (Mahoney, McDonald, Optican, & Tinsley, 2010). Similar provisions for providing only portions of children's evidence also exist in other countries (e.g., England, Wales, Australia; McDonald & Tinsley, 2011).

One concern with presenting only portions of the evidential interview is that it prevents the jury from viewing the presented portions of the interview in their original context. During an evidential interview, children's accounts are not necessarily limited to reporting information

specific to their core allegations; children may also report other information. Sometimes that additional information comprises tangential or unrelated information, but sometimes it also includes additional allegations. Some of these additional allegations might have a bearing on jurors' assessments of a child's credibility. For example, additional bizarre and implausible claims occur at times in children's allegations of sexual abuse as well as in other contexts (Bruck & Ceci, 1995; "Collapse of bizarre abuse case reveals 'mistakes'", 2012; Hood, 2001; Linder, 2003; Schreiber & Parker, 2004). In the highly publicized McMartin Preschool case, children claimed that they had watched a baby being beheaded and were made to drink the blood. They also reported that they had watched animal sacrifices (Linder, 2003). Similarly bizarre allegations were also made against Kelly Michaels in the Wee Care Nursery School case. In that case, children made claims that Kelly raped and assaulted them with knives, swords, spoons, and Lego blocks; they also alleged that Kelly licked peanut butter off their genitals, made them drink her urine and eat her feces, amputated their genitals, and made them lick each other like cats (Bruck & Ceci, 1995; Schreiber & Parker, 2004). Although these additional allegations might not refer to the specific charges in a particular trial, jurors would likely factor this additional information into their assessment of the child's credibility and believability. How might allegations of this kind—or, more importantly, their exclusion—influence jurors' evaluations of children's evidence regarding the core allegation for an offense with which the defendant has been charged?

Unfortunately, there is currently almost no research that helps us to answer this question. What we know so far is that, when it comes to adults' testimony, jurors are sceptical of implausible or bizarre details that are part of a witness's allegation. In a study by Peace, Brower, and Rocchio (2014), mock jurors read an adult's eyewitness statement about a mugging. The statement either contained 5, 10, or 15 bizarre details; these were mildly bizarre (e.g., the perpetrator wore dark red gloves), moderately bizarre (e.g., the perpetrator wore white gloves), or extremely bizarre (e.g., the perpetrator wore a baseball glove). As the bizarreness level increased, the perceived level of credibility of the witness decreased. Similarly, the more bizarre details that the testimony contained, the less credible, believable, and plausible the witness and the evidence were perceived to be.

Do similar factors affect jurors' evaluation of children's evidence? We do know that mock jurors evaluate children's testimony differently when shown the portion of the evidential interview that includes the rapport-building phase and questions regarding children's understanding of truth and lies. When these phases of the interview are included in the transcript, mock jurors rate the child witness as more credible and honest, and are more likely to believe the allegation (Krähenbühl, 2012). But what about additional allegations that a child

might make in the course of recounting a central allegation that has led to a charge against the defendant? To date, there are no published studies examining how jurors interpret segments of a child's testimony in the absence of the context of the full interview, particularly when the child's core allegations are presented against a backdrop of additional—potentially bizarre or implausible—claims.

The aim of the current study was to examine how mock jurors rate the believability of a child's testimony when that testimony was presented either in full, or had been truncated in a way that is currently allowable in many courts around the world. Specifically, we examined the effect of presenting only portions of a child's transcribed testimony, on mock jurors' perceptions of the child's credibility and of the defendant's guilt or innocence. To do this, in two experiments we asked jury-eligible adults to evaluate a transcript that included only a core allegation (physical assault by a janitor) or transcripts that included the core allegation in conjunction with one or more additional allegations that were less plausible. In Experiment 1, the additional, less plausible allegations involved allegations relevant to physical abuse; in Experiment 2, the additional, less plausible allegations involved allegations relevant to sexual abuse. We hypothesised that the inclusion of less plausible allegations would have a negative impact on mock jurors' perception of the child's credibility and on their evaluation of the core allegation.

Experiment 1

Method

Participants. Participants were recruited from an introductory psychology course or via word of mouth. All of the participants were jury eligible. The final sample size was 64 (49 females; M age = 19.73 years; SD = 2.32). Participants identified as New Zealand European (81.3%), Māori (4.7%), or 'other' (14%). An additional two participants were excluded because they failed to demonstrate that they had thoroughly read, comprehended, and retained the information that was contained in their assigned transcript. The research was reviewed and approved by the University of Otago Human Ethics Committee, which is approved by the New Zealand Health Research Council and whose guidelines are consistent with those of the American Psychological Association. Written informed consent was obtained from all participants. Participants received course credit or a movie voucher in return for their participation.

Procedure. Participants were tested individually in the laboratory. They were told that they would be asked to read a transcript of an evidential interview with a child, putting themselves in the role of a juror, and that they would then answer some questions from the experimenter before completing a computer-based questionnaire.

Each participant then read the same background information about the case:

Warren Smith is a 55-year-old cleaner [janitor] at a local primary school. Warren has a mild intellectual impairment. He works most days as part of a community-based employment scheme. He has worked in this particular school for almost a year and there have been no prior complaints about his behaviour. He is charged with assault on a child. The allegation was made by a 6-year-old, female child. The child initially disclosed the alleged assault to her mother and was subsequently interviewed by a specialist interviewer for the police.

Next, participants read a transcript of a mock evidential interview that was conducted with a child complainant, Jasmine. Based on a corpus of transcripts from real evidential interviews with child complainants of abuse, we constructed four different interview transcripts for this experiment. Participants were randomly assigned to one of four transcript conditions ($n = 16$ per condition). Each transcript began with a series of questions designed to establish that the child witness understood the difference between truth and lies, knew what making a promise meant, and promised to tell the truth.

Each transcript then included questions from the interviewer and responses from the child witness about the same core allegation—Jasmine alleged that Warren hit her on the leg in the school bathroom and that, as a result, her leg went purple. Following the core allegation, the transcripts differed according to the number and nature of additional allegations. For participants in the *Core Only* condition, the transcript concluded at the end of the core allegation. For participants in the *Core + 1* condition, the core allegation was followed by the additional allegation that Warren locked Jasmine in a cupboard at school along with two of her friends. In the *Core + 2* condition, the core allegation was followed by the additional allegation that Warren drove Jasmine from school to a cave where some of his friends—dressed as wizards—hit her with swords and sticks and took photos. Finally, for participants in the *Core + 1 and 2* condition, the transcript contained all three allegations.

To ensure that the questioning style used to elicit each of the three allegations was equivalent, we balanced the number of open-ended, specific-open, and multiple-choice questions used to elicit each allegation. We also balanced the number of off-topic responses given by the child, and the number of minimal responses (e.g., *umm* or *uh-huh*), pauses, and instructions to

speak up given by the interviewer. The allegations were also similar in length (Core Allegation: 545 words; Additional Allegation 1: 485 words; Additional Allegation 2: 539 words).

When the participant had finished reading the transcript, the experimenter began the verbal interview phase. During this phase, participants were asked several questions about what they had read and made several judgements about the case. First, participants were asked three questions to ensure that they had thoroughly read, comprehended, and retained the information contained in their assigned transcript (i.e., “*How old is Jasmine?*”, “*How many times was Jasmine allegedly assaulted?*”, and “*Whom did Jasmine first tell about the assault by Warren?*”). Participants who could not answer these questions correctly were excluded from the final sample; two participants were excluded for this reason (see Participants above).

Next, participants were asked to render a decision regarding each of the allegations in their assigned transcript. All participants were asked, *Did Warren Jasmine?* Participants in the *Core + 1* condition were also asked, *Did Warren lock Jasmine in a cupboard?* Participants in the *Core + 2* condition were also asked, *Did Warren take Jasmine somewhere in a car?* Participants in the *Core + 1 and 2* condition were asked all three questions. After providing their decision, participants rated how strongly they believed Jasmine's allegation of physical assault on a 7-point scale (1 = *Not at all*; 4 = *Neutral*; 7 = *Very strongly*), and they were asked to describe why they had made that choice. Finally, participants rated the overall likelihood that Warren committed the assault (1 = *Extremely unlikely*; 4 = *Neutral*; 7 = *Extremely likely*), and gave an overall confidence rating regarding Warren's guilt (1 = *Extremely confident that Warren is innocent*; 4 = *Neutral*; 7 = *Extremely confident that Warren is guilty*).

Following the verbal interview, participants were seated at a computer and were presented with a series of questions specific to the case. The questions were based on those used by Kulisek (2014) and were presented using MediaLab (Jarvis, 2012). First, using a 7-point scale (1 = *very unlikely*; 4 = *neutral*; 7 = *very likely*), participants answered a series of 18 questions regarding their views on the circumstances surrounding Jasmine's allegation(s) (Appendix 1). Next, participants used another 7-point scale (1 = *low*; 4 = *moderate*; 7 = *high*) to rate Jasmine on 11 characteristics: memory ability, language ability, innocence, ability to distinguish imagination from reality, emotional maturity, knowledge of right and wrong, obedience, honesty, confidence, likeability, and intelligence.

Based on a reliability analysis conducted by Kulisek (2014), we used each participant's responses to selected questions to calculate two scores reflecting that participant's perception of Jasmine's credibility: an *honesty* score (combined mean of ratings for questions 1, 15, 16, 18, and the rating of honesty; Cronbach's $\alpha = .82$) and a *cognitive competence* score (combined mean of

ratings for questions 5-9, 11, and ratings of memory ability, language ability, ability to distinguish imagination from reality, knowledge of right and wrong, and intelligence; Cronbach's $\alpha = .68$).

After completing the questionnaire, participants were thanked for their time and debriefed.

Results

Figure 1 shows the number of participants who answered "yes" to the questions about the allegations that were contained in their assigned transcript¹. Across all four conditions, the majority of the participants believed that Warren did hit Jasmine on the leg. In addition, as Figure 1 shows, participants were more likely to believe that the core allegation happened than they were to believe that the additional allegations happened, confirming that participants considered these additional allegations as less plausible.

Recall that the primary research question was whether the inclusion of additional, less plausible allegations influenced mock jurors' evaluation of the core allegation. To answer this question, we used Fisher's Exact Test (FET) with the Freeman-Halton extension² to compare the number of participants who said "yes" to the question about the core allegation, *Did Warren hit Jasmine?* Although this analysis did not reach conventional levels of significance ($p = .063$, FET), given that the pattern of participants' responses was in the opposite direction to what we had expected (Figure 1), we investigated the differences between the transcript conditions further. Contrary to our hypothesis, participants in the *Core + 1 and 2* condition were *more* likely to say "yes" to the question about the core allegation ($n = 15$) than were participants in the *Core only* condition ($n = 9$; $p = .037$, FET). That is, the presence of two additional allegations—both of which contained details that were less plausible than those in the core allegation—increased the likelihood that participants believed the core allegation. The number of participants in the *Core + 1* and the *Core + 2* conditions who believed the core allegation were intermediate between these two extremes ($ns = 13$ and 11 , respectively) and was not statistically different from either (ps ranged from $.252$ to $.716$, FET).

¹ Exploratory analyses indicated that there were no gender differences for any of the findings; the substantial lack of balance in the number of males and female participants in the sample precluded more formal analyses.

² The Freeman-Halton extension of Fisher's exact test is used to analyse a 4×2 contingency table when 20% of the expected cell frequencies are < 5 (Freeman & Halton, 1951).

Next, we took a closer look at the pattern of responses for participants whose assigned transcripts had contained additional, less plausible allegations. We asked the following question: if participants believed that the additional allegations *did not happen*, how did they judge the core allegation? We found that 7 out of 16 participants in the *Core + 1* condition did not believe the allegation that Warren had locked Jasmine in the cupboard; of those 7 participants, 71% still believed the core allegation. Similarly, for participants in the *Core + 2* condition, 10 out of 16 did not believe the allegation that Warren had taken Jasmine somewhere in a car; of those 10 participants, 60% still believed the core allegation.

Perhaps the most compelling data come from those participants whose assigned transcript contained the core allegation as well as both additional allegations. Ten out of 16 participants in the *Core + 1 and 2* condition did not believe that Warren locked Jasmine in the cupboard, and of those 10 participants, 90% still believed the core allegation. Eleven out of the 16 participants did not believe that Warren had taken Jasmine somewhere in a car; of those 11 participants, 91% still believed the core allegation. Finally, 9 out of 16 participants did not believe both of the additional allegations, but of those 9 participants, 89% still believed the core allegation. In summary, across all conditions, participants were highly likely to believe the core allegation, even when they rejected the additional allegations that were part of the same transcript of the same child's testimony.

In the final part of our analysis, we calculated a series of 4 (Transcript Condition) x 2 (Core Allegation Decision: Yes, No) between-subjects analyses of variance (ANOVAs) to examine the effect of these two variables on participants' judgments of Jasmine's believability, honesty, and cognitive competence; the likelihood of the defendant's guilt; and confidence in verdict.

Regardless of the number or nature of the allegations contained in the transcripts, participants' ratings on the questions assessing Jasmine's believability, honesty, and cognitive competence, the likelihood of the defendant's guilt, or participants' confidence in their decisions regarding the allegations were similar. That is, there was no significant main effect of transcript condition on any of these variables. The mean ratings were all in the 4 to 5 range: largest $F(3, 56) = 0.85, p = .47, h_p^2 = .04$.

Not unexpectedly, relative to participants who did not believe the core allegation, participants who believed it also believed Jasmine's allegation more strongly, "Yes" $M = 5.04, SD = 0.80$, "No" $M = 2.66, SD = 0.77; F(1, 56) = 75.31, p < .01, h_p^2 = .57$, thought that it was more likely that Warren was guilty of the allegation, "Yes" $M = 4.64, SD = 1.05$, "No" $M = 3.19$,

$SD = 1.28$; $F(1, 56) = 11.84, p = .001, h_p^2 = .17$, were more confident that Warren was guilty, "Yes" $M = 4.91, SD = 0.85$, "No" $M = 2.94, SD = 1.06$; $F(1, 56) = 41.80, p < .01, h_p^2 = .43$, and rated Jasmine as more honest, "Yes" $M = 3.93, SD = 0.84$, "No" $M = 2.68, SD = 0.62$; $F(1, 56) = 20.57, p < .01, h_p^2 = .27$. In contrast, participants' ratings of Jasmine's cognitive competence did not differ as a function of their decisions, $F(1, 56) = 1.34, p = .25, h_p^2 = .02$. The Transcript $\times$ Core Allegation Decision interaction was not significant for any of the dependent variables, largest $F(3, 56) = .94, p = .43, h_p^2 = .05$.

Experiment 2

Regardless of whether they had received information about a core allegation only or about additional allegations that contained less plausible details, the majority of participants in Experiment 1 indicated that Warren was guilty of the core allegation. In fact, in stark contrast to our original hypothesis, participants were most likely to believe the core allegation of physical assault when they were exposed to the transcript containing the highest number of less plausible allegations. That is, when a child's core allegations were presented against a backdrop of less plausible allegations, mock jurors still appeared to believe that something must have happened. In Experiment 2, we assessed the generality of this finding by creating three new transcripts that contained additional allegations that varied with respect to plausibility. In contrast to Experiment 1, these new allegations were relevant to child sexual abuse.

Method

Participants. Participants were recruited in the same way as in Experiment 1. All of the participants were jury eligible. The final sample comprised 64 participants (40 females; M age = 19.80 years, $SD = 2.89$). Participants identified as New Zealand European (71.9%), Māori (6.3%), or 'other' (21.8%). An additional 10 participants were excluded because they failed to demonstrate that they had thoroughly read, comprehended, and retained the information contained in their assigned transcript.

Procedure. To provide a point of comparison, we retained the original core allegation of physical assault from Experiment 1. We then constructed two new additional allegations that were taken directly from real evidential interviews with child complainants of sexual abuse. With the exception of identifying information, these allegations were presented exactly as they had

appeared in the transcripts of the original evidential interviews. For participants in the *Core + 1* condition, the core allegation was followed by an additional allegation that Warren made Jasmine get in the bath with him, telling her that he would make her do it by getting her undressed. Jasmine also alleged that, while they were in the bath, Warren forced her to eat her "poos." For participants in the *Core + 2* condition, the core allegation was followed by an additional allegation that Warren took Jasmine to another location. During the interview, Jasmine changes her story from saying that Warren took her to a library building to saying that Warren took her to his mother's house or to his grandmother's house, and that Warren's friend's mother and grandmother were present. Jasmine also alleged that Warren locked her beneath a trapdoor and stuck burning paper and a stick up her bottom, and that 20 of his friends were there. Finally, for participants in the *Core + 1 and 2* condition, the transcript contained all three allegations. Participants were randomly assigned to one of four transcript conditions ($n = 16$ per condition).

When the participant had finished reading the transcript, the experimenter began the verbal interview phase. Participants were asked to make a decision regarding whether each of the allegations in their assigned transcript had occurred. All participants were asked, *Do you think that Warren hit Jasmine in the toilets?* Participants in the *Core + 1* condition were also asked, *Do you think that Warren made Jasmine go in the bath with him?* Participants in the *Core + 2* condition were also asked, *Do you think that Warren took Jasmine to the library building?* Participants in the *Core + 1 and 2* condition were asked all three questions. Immediately after providing each decision, participants were asked to rate their confidence in that decision on the same 7-point scale used in Experiment 1. Following the verbal interview, participants completed the same series of computer-based questions that we used in Experiment 1 to calculate perceptions of Jasmine's *honesty* (Cronbach's $\alpha = .77$) and *cognitive competence* (Cronbach's $\alpha = .72$).

Results

Figure 2 shows the number of participants who indicated that they believed the allegations that were contained in their assigned transcript³. As in Experiment 1, irrespective of transcript condition, the majority of the participants believed that Warren had committed the crime in the core allegation. In addition, participants were much more likely to believe the core allegation than they were to believe either of the additional allegations, again confirming that these allegations were viewed as less plausible.

³ Again, a substantial imbalance in the number of male and female participants in the sample precluded analyzing potential gender differences.

As in Experiment 1, we used Fisher's Exact Test (FET) with the Freeman-Halton extension to analyze whether the number of participants who believed the core allegation differed as a function of which additional allegations, if any, were contained in participants' assigned transcript. As shown in Figure 2, there was no significant difference in the number of participants who believed the core allegation as a function of transcript condition ($p = .67$, FET), nor did transcript condition affect participants' confidence in that decision (mean confidence ratings ranged from 3.06 to 3.50), $F(3, 60) = 0.97, p = .41, h_p^2 = .05$, or in the confidence ratings for the decisions that they rendered to the additional allegations, largest $t(30) = 1.48, p = .15$.

As in Experiment 1, we took a closer look at the pattern of responses for participants whose assigned transcripts had contained additional allegations. In the *Core + 1* condition, 13 out of 16 participants did not believe that Warren had forced Jasmine take a bath with him and eat her "poos;" of those 13 participants, 46% still believed the core allegation. In the *Core + 2* condition, 12 out of 16 participants did not believe the allegation that Warren had taken Jasmine to another location, where he locked her beneath a trapdoor and stuck burning paper and a stick up her bottom; of those 12 participants, 67% still believed the core allegation. Eight out of 16 participants in the *Core + 1 and 2* condition did not believe the 'bath' allegation; of those 8 participants, 75% believed the core allegation. Eight out of 16 participants did not believe the 'burning stick' allegation, and of those 8 participants, again, 75% believed the core allegation. Finally, 6 out of 16 participants did not believe either of the additional allegations, but of those 6 participants, 83% still believed the core allegation.

Again, we used 4 (Transcript Condition) x 2 (Core Allegation Decision: Yes, No) between-subject ANOVAs to examine whether participants' ratings of the child's honesty and cognitive competence differed according to transcript condition or to participant's decision regarding the core allegation. There was a main effect of transcript condition, $F(3, 56) = 6.53, p < .01, h_p^2 = .26$. Post hoc Student Newman-Keuls tests indicated that participants in the *Core Only* condition rated Jasmine as being more honest, $M = 4.45, SD = 0.84$, than did participants in the three other transcript conditions (mean ratings ranged from 3.30 to 3.86)., There was also a main effect of core allegation decision, $F(1, 56) = 20.57, p < .01, h_p^2 = .27$. Participants who believed the core allegation rated Jasmine as being more honest, "Yes" $M = 3.94, SD = 0.81$, than did participants who did not believe the core allegation, "No" $M = 3.39, SD = 1.03$. The interaction was not significant, $F(3, 56) = .38, p = .77, h_p^2 = .02$. As in Experiment 1, participants' ratings of Jasmine's cognitive competence did not significantly differ as a function

of transcript condition, decision, or the interaction between these variables, largest $F(1, 56) = 2.81, p = .10, h_p^2 = .05$.

General Discussion

Many countries around the world have altered their legislation to allow child witnesses to present pre-recorded evidence as an alternative to testifying live in the courtroom. With this modification has also come the provision for a child's recorded evidential interview—or a transcript of it—to be truncated before it is presented to the jury. This provision allows legal professionals to edit out parts of the interview(s) that could be seen as inadmissible in court. In turn, this means that the judge does not need to issue jurors with an instruction to ignore parts of the witness's testimony—something that jurors are often asked to do when testimony is presented live. The use of truncated testimony also spares jurors from watching long segments of an evidential interview that may not be relevant to the charge or charges at hand. The distinction between what is relevant and what is not, however, is not a straightforward one. Although omitted segments might not refer to the specific charges in a given trial, the additional information could help jurors to place the relevant details in context, thereby aiding their decisions about the credibility of the witness and the culpability of the defendant.

Unfortunately, the extant database on how excluding portions of an interview might affect jurors' decision-making is sparse. In the present research, we manipulated the content of simulated interview transcripts to explore how truncating evidence influenced mock jurors' judgments of child credibility and defendant culpability. Participants either read a core allegation of physical assault in isolation, or the same core allegation accompanied by either one or two additional—but less plausible—allegations. We hypothesized that the addition of these less plausible allegations would decrease the likelihood that participants would believe that the defendant was guilty of the core allegation. Contrary to our hypothesis, in two experiments, we found that across all transcript conditions, the majority of our participants believed the core allegation—that the defendant, Warren, hit Jasmine, the child complainant. The presence of additional allegations did not significantly decrease participants' belief that the defendant committed the crime in the core allegation, nor did it significantly influence their judgments about the child complainant's honesty or cognitive competence. In fact, in Experiment 1,

exposure to two additional, less plausible allegations actually made participants *more* likely to believe that the defendant was guilty of the core allegation—even when they did not necessarily believe those additional allegations.

Our findings stand in stark contrast to prior research on mock jurors' evaluation of adults' testimony. In prior research with adults, the inclusion of additional implausible or bizarre details has been shown to have a negative effect on the perceived credibility of the witness and the believability of the evidence (Peace et al., 2014). As in the Peace et al. study, the additional allegations that we used in the present research included implausible and bizarre details. For example, in Experiment 1, the additional allegations included Jasmine alleging that Warren had driven her to a remote place where additional men dressed as wizards had hit her with a sword. In Experiment 2, the additional allegations included Jasmine alleging that Warren had forced her to have a bath with him and had forced her to eat feces and/or that Warren and 20 of his friends had taken her to another location and assaulted her by inserting a burning object into her bottom. Although most of the participants did not believe one or both of these allegations, their view about these implausible events had no bearing on their evaluation of the core allegation—that Warren had hit Jasmine at school, leaving a bruise on her leg.

Why might this be the case? One possibility is that our mock jurors placed considerable weight on the nature of the questions that the interviewer asked. In the Peace et al. (2014) study, participants read an eyewitness's statement about an alleged crime in the absence of the questions that elicited that statement. In contrast, participants in our study read a full simulated evidential interview that included not only what the child said, but also what the interviewer said. In preparing the transcripts, we were careful to ensure that the questioning style reflected elements of best practice, and did not differ as a function of allegation. Although the participants were not asked to evaluate the quality of the interviewing, many provided explicit reference to the interviewer's questioning style when providing a justification for their decision. For example, when asked why they believed an allegation, participants' responses included "*the questions were not leading*," "*clear questions—the interviewer didn't suggest anything*," "*she seemed to understand the questions*," "*there was no persuasion from the interviewer*," "*no ideas were put in her head*," or "*no words were put in her mouth*."

Although our mock jurors were correct in assuming that the presence of leading or misleading questions can taint a child's testimony, they were not correct in assuming that the absence of leading or misleading questions means that a child's testimony is accurate. Based on what we know from prior research, the only safe assumption is that the interviewer has not made the child's account *less* accurate during the interview. Over the past two decades, research has

shown that there are plenty of things that can happen before a child is interviewed that can taint his or her testimony even in response to good interviewing, including direct attempts to mislead (Leichtman & Ceci, 1995; Principe, DiPuppo, & Gammel, 2013), overhearing others' conversations about an event (Principe, Kanaya, Ceci, & Singh, 2006; for review, see Principe & Schindewolf, 2012), or stereotype induction (Leichtman & Ceci, 1995). In our study, mock jurors appeared to approach their decision-making with at least some basic understanding that poor interviewing practices can distort children's reports, but their awareness of other potential external influences on children's accounts appeared to be limited.

What other factors may have played a role in the mock jurors' evaluation of the witness and her testimony? Previous research has shown that mock jurors estimate a child's credibility using a number of different factors including the child's age and gender, language ability, demeanor, and perceived honesty, suggestibility, and intelligence (Bottoms, Golding, Stevenson, Wiley, & Yozwiak, 2007). In the present study, many of our mock jurors used the child's age as a justification for their decisions, and her young age worked in her favor. For example, when asked why they believed that Warren was guilty, participants said things like, "*why would a child make it up,*" "*she's only 6,*" "*kids are pretty truthful,*" "*too much detail for a child to make up,*" "*she's young, she wouldn't be able to make up these stories,*" "*she's a young child and they tell the truth,*" "*weird thing for a 6 year old to make up,*" or "*details in there that you wouldn't expect a child to know*". Mock jurors sometimes find testimony provided by younger children to be more believable than testimony provided by older children (Myers, Redlich, Goodman, Prizmich, & Imwinkelried, 1999; Nightingale, 1993), because young children are perceived to lack the cognitive capacity that is required to fabricate an allegation of abuse (for a review, see Bottoms et al., 2007). In this way, a perceived limitation in children's cognitive ability increases the perceived credibility of their evidence (Ross, Dunning, Toglia, & Ceci, 1990). In future research of this kind, it will be important to explore the interaction between the child's age and the plausibility of the allegations on jurors' evaluation of the witness and the testimony.

In addition to the child's age, research has also shown that the outcome of a trial is significantly influenced by the demeanor of the child witness, especially in cases of suspected sexual abuse (Golding, Dunlap, & Hodell, 2009). In the present study, we specifically used transcripts of the child's testimony as a way to encourage our participants to make their judgments on the basis of the information provided in the interview, rather than on the basis of additional features of the witness, like demeanor. Despite the fact that participants only had transcripts to go on, however, some made explicit reference to the child's demeanor when explaining their guilty verdicts (e.g., "*she seemed genuine,*" "*she was quite upset,*" or "*she cried*"). In an

attempt to gain a clearer understanding of the relation between allegation plausibility and witness demeanor, we are currently exploring how mock jurors evaluate plausible and implausible allegations that are presented via pre-recorded video, where participants have the explicit opportunity to assess the child's demeanor.

Although the bulk of our analyses resulted in null findings, the fact that similar outcomes occurred across two experiments involving four different allegations adds considerable weight to our conclusion that the inclusion of additional implausible or bizarre allegations had little or no effect on mock jurors' evaluation of a core allegation of physical assault (see also Nuzzo, 2015; Wells & Windschitl, 1999). On the basis of these findings, we might be tempted to conclude that there is no cause for concern in trials where a jury sees only portions of a child's evidential interview. But before declaring the practice safe, additional research is clearly required. In the present experiments, for example, the core allegation was heard first; the less plausible allegations came later. Under these conditions, it is possible that participants established a positive view of the witness at the beginning of the interview and anything she said later, even if the participants did not believe it, did not influence their subsequent decision about the initial allegation. Given this possibility, we are currently examining the effect of allegation order on mock jurors' verdicts.

In addition, our sample consisted primarily of female participants. Research has shown that females are more likely to exhibit a pro-victim/pro-prosecution bias in their judgements relative to males—a pattern that is particularly strong in child sexual abuse cases. In these cases, women's judgements typically involve a higher number of guilty verdicts and their rating of the believability of the child is higher than those of men (Bottoms et al., 2007; Quas, Bottoms, Haegerich, & Nysse-Carris, 2002; Redlich, Myers, Goodman, & Qin, 2002; Ross et al., 1994; Schutte & Hosch, 1997). In future research, it would be ideal to specifically examine gender differences in jurors' decision-making using our stimuli. At the same time, however, there is no provision to balance the gender composition of juries in actual trials—so the practical implications of the findings could be somewhat limited.

In the case of an actual trial, jurors would not make their decisions after hearing just the child's direct evidence. Instead, the child would be cross-examined, either live in the courtroom or via an alternative method (e.g., CCTV or via pre-recording). If the allegations that we used here had been heard in court, the defense would have undoubtedly challenged the veracity of Jasmine's claims, particularly her claims about the less plausible or bizarre aspects of her evidence. Many in the legal profession would ardently argue that cross-examination provides a mechanism for inaccuracies to be uncovered (e.g., Wigmore, 1974). Despite this widely held view, there is mounting evidence that cross-examination questioning is not an effective way of

arriving at the truth, and that children's responses to this style of questioning may actually hold little diagnostic value when it comes to accuracy (Zajac, Irvine, Ingram, & Jack, in press; see Zajac, O'Neill, & Hayne, 2012, for a review). Given the faith that is often placed in the cross-examination process, there is a dearth of research examining whether cross-examination helps or hinders jurors tasked with determining the veracity of a child's account (but see Fogliati & Bussey, 2014, 2015; Zajac et al., in press). The effect of cross-examination-style questioning on jurors' evaluation of potentially implausible claims is another important avenue for future research.

In conclusion, the results of the present study raised at least as many questions as they answered. In discussing our findings, we have signaled a number of areas for additional research. The overarching take-home message is that, when it comes to children, we often alter our practice in the courtroom before we adequately evaluate the potential impact of that practice on children's accounts or on how they are perceived by a jury. In our view, it is more prudent to understand the impact first—providing children with a firm, evidence-based platform from which to have their voices heard in court.

References

- Bala, N. (1999). Child witnesses in the Canadian criminal courts: Recognizing their capacities and needs. *Psychology, Public Policy, and Law*, 5, 323-354. DOI: 10.1037//1076-8971.5.2.323
- Bennett, K. (2004). Legal and social issues surrounding closed-circuit and television testimony of child victims and witnesses. *Journal of Aggression, Maltreatment & Trauma*, 8, 233-271. DOI: 10.1300/J146v08n03_02
- Bottoms, B. L., Golding, J. M., Stevenson, M. C., Wiley, T. R. A., & Yozwiak, J. A. (2007). A review of factors affecting jurors' decisions in child sexual abuse cases. In J. D. Read, D. Ross, M. Toglia, & R. Lindsay (Eds.), *The psychology of eyewitness memory* (pp. 509-543). Mahwah, NJ: Erlbaum.
- Brennan, M., & Brennan, R. (1988). *Strange language: Child victim witnesses under cross-examination*. Wagga Wagga, Australia: CSU Literary Studies Network.
- Bruck, M., & Ceci, S. J. (1995). *Jeopardy in the courtroom*. Washington, DC: American Psychological Association.
- Cashmore, J. (1992). *The use of closed-circuit television for child witnesses in the ACT*. Sydney, AU: Australian Law Reform Commission.
- Cashmore, J., & Bussey, K. (1996). Judicial perceptions of child witnesses. *Law and Human Behavior*, 20, 313-334. DOI: 10.1007/BF01499026
- Cashmore, J., & Trimboli, L. (2006). *Child sexual assault trials: A survey of juror perceptions*. Sydney: New South Wales Bureau of Crime Statistics and Research. Retrieved from: <http://www.bocsar.nsw.gov.au/Documents/CJB/cjb102.pdf>
- Collapse of bizarre abuse case reveals 'mistakes' (2012, August 17). *The Dominion Post*. Retrieved from <http://www.stuff.co.nz/dominion-post/news/7495365/Collapse-of-bizarre-abuse-case-reveals-mistakes>.
- Connolly, D., & Read, J. (2006). Delayed prosecution of historic child sexual abuse: Analyses of 2064 Canadian criminal complaints. *Law and Human Behavior*, 30, 409-434. DOI: 10.1007/s10979-006-9011-6
- Connolly, D. A., Price, H. L., Lavoie, J. A., & Gordon, H. M. (2008). Perceptions and predictors of children's credibility of a unique event and an instance of a repeated event. *Law and Human Behavior*, 32, 92-112. DOI: 10.1007/s10979-006-9083-3
- Davies, G. M. (1992). Influencing public policy on eyewitnessing: Problems and possibilities. In F. Losel, D. Bender, & T. Bliesener (Eds.), *Psychology and law: International perspectives* (pp. 265-274). Berlin: De Gruyter.

- Davies, G. M. (1999). The impact of television on the presentation and reception of children's testimony. *International Journal of Law and Psychiatry*, 22, 241-256. DOI: 10.1016/S0160-2527(99)00007-2
- Davies, G. M., & Noon, E. (1991). *An evaluation of the live link for child witnesses*. London: UK Home Office.
- Davies, G. M., Wilson, C., Mitchell, R., & Milsom, J. (1995). *Videotaping children's evidence: An evaluation*. London: UK Home Office. Retrieved from <http://webarchive.nationalarchives.gov.uk/20130128103514/http://rds.homeoffice.gov.uk/rds/pdfs2/r20.pdf>
- Eastwood, C., & Patton, W. (2002). *The experiences of child complainants of sexual abuse in the criminal justice system*. Canberra, AU: Report to the Criminology Research Council. Retrieved from: <http://www.secasa.com.au/assets/Statistics/the-experiences-of-child-complainants-of-sexual-abuse-in-the-criminal-justice-system-2002.pdf>
- Eaton, T. E., Ball, P. J., & O'Callaghan, M. G. (2001). Child-witness and defendant credibility: Child evidence presentation mode and judicial instructions. *Journal of Applied Social Psychology*, 31, 1845-1858. DOI: 10.1111/j.1559-1816.2001.tb00207.x
- Evidence Amendment Act. (1989). Retrieved from <http://www.knowledge-basket.co.nz/databases/legislationnz/acts/view/?d1=public/text/1989/an/104.html>
- Flin, R. (1990). The case for change. In J. R. Spencer & R. Flin (Eds.), *The evidence of children* (pp. 282-305). London: Blackstone.
- Fogliati, R., & Bussey, K. (2014). The effects of cross-examination on children's reports of neutral and transgressive events. *Legal and Criminological Psychology*, 19, 296-315. doi:10.1111/lcrp.12010.
- Fogliati, R., & Bussey, K. (2015). The effects of cross-examination on children's coached reports. *Psychology, Public Policy, and Law*, 21, 10-23. doi: 10.1037/law0000036.
- Freeman, G. H., & Halton, J. H. (1951). Note on exact treatment of contingency, goodness of fit and other problems of significance. *Biometrika*, 38, 141-149. DOI: 10.1093/biomet/38.1-2.141
- Golding, J. M., Dunlap, E. E., & Hodell, E. C. (2009). Jurors' perceptions of children's eyewitness testimony. In B. L. Bottoms, C. J. Najdowski, & G. S. Goodman (Eds.), *Children as victims, witnesses, and offenders: Psychological science and the law* (pp. 188-208). New York, NY: Guilford Press.

- Goodman, G. S., Taub, E., Jones, D. P. H., England, P., Port, L., Rudy, L., & Prado, L. (1992). Testifying in criminal court: Emotional effects on child sexual assault victims. *Monographs of the Society for Research in Child Development*, 57, 1-142. DOI: 10.2307/116612
- Goodman, G. S., Tobey, A. E., Batterman-Faunce, J. M., Orcutt, H., Thomas, S., Shapiro, C., & Sachsenmaier, T. (1998). Face-to-face confrontation: Effects of closed-circuit technology on children's eyewitness testimony and jurors' decisions. *Law and Human Behavior*, 22, 165-203. DOI: 10.1023/A:1025742119977
- Goodman, G. S., Myers, J. E. B., Qin, J., Quas, J. A., Castelli, P., Redlich, A. D., & Rogers, L. (2006). Hearsay versus children's testimony: Effects of truthful and deceptive statements on jurors' decisions. *Law and Human Behavior*, 30, 363-401. DOI: 10.1007/s10979-006-9009-0
- Gupta, T. (1994). Child witnesses – An Australian perspective. *Child Abuse Review*, 3, 179–182. DOI: 10.1002/car.2380030306
- Hall, S. R., & Sales, B. D. (2008). *Courtroom modifications for child witnesses: Law and science forensic evaluations*. Washington, DC: American Psychological Association.
- Hanna, K., Davies, E., Henderson, E., Crothers, C., & Rotherham, C. (2010). *Child witnesses in the New Zealand criminal courts: A review of practice and implications for policy*. Auckland, NZ: Institute of Public Policy.
- Hill, P. E., & Hill, S. M. (1987). Videotaping children's testimony: An empirical view. *Michigan Law Review*, 85, 809-833. DOI: 10.1016/0145-2134(93)90005-P
- Hood, L. (2001). *A city possessed: The Christchurch Civic Creche case*. Dunedin, New Zealand: Longacre.
- Jarvis, B. G. (2012). MediaLab (Version 2012) [Computer Software]. New York, NY: Empirisoft Corporation.
- Klemfuss, J., & Ceci, S. J. (2009). Normative memory development and the child witness. In K. Kuehnle & M. Connell (Eds.), *The evaluation of child sexual abuse allegations: A comprehensive guide to assessment and testimony* (pp. 153-180). Hoboken, NJ: Wiley.
- 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
- Kulisek, N. R. (2014). Effects of child age and type of detail reported on credibility of child abuse (Doctoral dissertation). Retrieved from <http://scholar.utc.edu/cgi/viewcontent.cgi?article=1131&context=theses>
- Landström, S., & Granhag, P. A. (2010). In-court versus out-of-court testimonies: Children's

- experiences and adults' assessments. *Applied Cognitive Psychology*, 24, 941-955. DOI: 10.1002/acp.1606
- Landström, S., Granhag, P. A., & Hartwig, M. (2007). Children's live and videotaped testimonies: How presentation mode affects observers' perception, assessment and memory. *Legal and Criminological Psychology*, 12, 333-347. DOI: 10.1348/135532506X133607
- Leichtman, M. D., & Ceci, S. J. (1995). The effects of stereotypes and suggestions on preschoolers' reports. *Developmental Psychology*, 31, 568-578. DOI: 10.1037/0012-1649.31.5.758
- Linder, D. (2003). *The McMartin Preschool abuse trial: A commentary*. Retrieved from: <http://law2.umkc.edu/faculty/projects/ftrials/mcmartin/mcmartinaccount.html>.
- Mahoney, R., McDonald, E., Optican, S. L., & Tinsley, Y. (2010). *The Evidence Act 2006: Act & analysis* (2nd edn). Wellington, NZ: Thomson Reuters.
- McDonald, E., & Tinsley, Y. (2011). Evidence issues. *The Canterbury Law Review*, 17, 123-159.
- Ministry of Justice (2010). *Alternative pre-trial and trial processes for child witnesses in New Zealand's criminal justice system*. Retrieved from <http://www.justice.govt.nz/publications/global-publications/a/alternative-pre-trial-and-trial-processes-for-child-witnesses-in-new-zealands-criminal-justice-system>
- Myers, J. E. B., Redlich, A., Goodman, G. S., Prizmich, L., & Imwinkelried, E. (1999). Juror's perceptions of hearsay in child sexual abuse cases. *Psychology, Public Policy, and Law*, 5, 388-419. DOI: 10.1037//1076-8971.5.2.388
- Murray, K. (1995). *Live television link: An evaluation of its use by child witnesses in Scottish criminal trials*. Central Research Unit, The Scottish Home Office. Retrieved from: <http://www.gov.scot/Publications/1998/12/5e1fb1ac-4e60-4c86-9b89-25f989249d3d>
- Nathanson, R., & Saywitz, K. J. (2003). The effects of the courtroom context on children's memory and anxiety. *Journal of Psychiatry & Law*, 31, 67-98. DOI: 10.1177/009318530303100105
- Nightingale, N. N. (1993). Juror reactions to child victim witnesses: Factors affecting trial outcome. *Law and Human Behavior*, 17, 679-694. DOI: 10.1007/BF01044689
- Nuzzo, R. (2015). Fooling ourselves. *Nature*, 426, 182-185. Retrieved from <http://www.nature.com/news/how-scientists-fool-themselves-and-how-they-can-stop-1.18517>
- Orcutt, H., Goodman, G. S., Tobey, A., Batterman-Faunce, J., & Thomas, S. (2001). Detecting deception in children's testimony: Factfinders' abilities to reach the truth in open court and closed-circuit trials. *Law and Human Behavior*, 25, 339-372. DOI:

10.1023/A:1010603618330

- Peace, K. A., Brower, K. L., & Rocchio, A. (2014). Is truth stranger than fiction? Bizarre details and credibility assessment. *Journal of Police and Criminal Psychology*, 30, 38-49. DOI: 10.1007/s11896-014-9140-7
- Pezdek, K., Morrow, A., Blandon-Gitlin, I., Goodman, G. S., Quas, J. A., Saywitz, K. J., . . . Brodie, L. (2004). Detecting deception in children: Event familiarity affects criterion-based content analysis ratings. *Journal of Applied Psychology*, 89, 119-126. DOI: 10.1037/0021-9010.89.1.119
- Pipe, M., & Henaghan, M. (1996). Accommodating children's testimony: Legal reforms in New Zealand. *Criminal Justice and Behavior*, 23, 145-167. DOI: 10.1177/0093854896023002008
- Powell, M. B., Wright, R., & Hughes-Scholes, C. H. (2011). Contrasting the perceptions of child testimony experts, prosecutors and police officers regarding individual child abuse interviews. *Psychiatry, Psychology and Law*, 18, 33-43. DOI: 10.1080/13218710903566995
- Principe, G. F., DiPuppo, J., & Gammel, J. (2013). Effects of mothers' conversation style and receipt of misinformation on children's event reports. *Cognitive Development*, 28, 260-271. DOI: 10.1016/j.cogdev.2013.01.012
- Principe, G. F., Kanaya, T., Ceci, S. J., & Singh, M. (2006). Believing is seeing: How rumors can engender false memories in preschoolers. *Psychological Science*, 17, 243-248. DOI: 10.1111/j.1467-9280.2006.01692.x
- Principe, G. F., & Schindewolf, E. (2012). Natural conversations as a source of false memories in children: Implications for the testimony of young witnesses. *Developmental Review*, 32, 205-223. DOI: 10.1016/j.dr.2012.06.003
- Quas, J. A., Bottoms, B. L., Haegerich, T. M., & Nysse-Carris, K. L. (2002). Effects of victim, defendant and juror gender on decisions in child sexual assault cases. *Journal of Applied Social Psychology*, 32, 1993-2021. DOI: 10.1111/j.1559-1816.2002.tb02061.x
- Redlich, A. D., Myers, J. E., Goodman, G. S., & Qin, J. (2002). A comparison of two forms of hearsay in child sexual abuse cases. *Child Maltreatment*, 7, 312-328. DOI: 10.1177/107755902237262
- Ross, D., Dunning, D., Toglia, M., & Ceci, S. J. (1989). Age stereotypes, communications modality, and mock jurors' perceptions of the child witness. In S. J. Ceci, D. Ross, & M. Toglia, (Eds.), *Perspectives of children's testimony* (pp. 23-37). New York, NY: Springer-Verlag
- Ross, D., Hopkins, S., Hanson, E., Lindsay, R., Hazen, K., & Eslinger, T. (1994). The impact of protective shields and videotape testimony on conviction rates in a simulated trial of child sexual abuse. *Law and Human Behavior*, 18, 553-566. DOI: 10.1007/BF01499174

- Saywitz, K. J., & Nathanson, R. (1993). Children's testimony and their perceptions of stress in and out of the courtroom. *Child Abuse & Neglect*, 17, 613-622. DOI: 10.1016/0145-2134(93)90083-H
- Schreiber, N., & Parker, J. F. (2004). Inviting witnesses to speculate: Effects of age and interaction on children's recall. *Journal of Experimental Child Psychology*, 89, 31-52. DOI: 10.1016/j.jecp.2004.03.008
- Schutte, J., & Hosch, H. (1997). Gender differences in sexual assault verdicts: A meta-analysis. *Journal of Social Behavior & Personality*, 12, 759-772.
- Scott, J. (1994). Preparation of the child witness for court – A Scottish view. *Child Abuse Review*, 3, 177-179. DOI: 10.1002/car.2380030305
- Swim, J., Borgida, E., & McCoy, K. (1993). Videotaped versus in-court witness testimony: Is protecting the child witness jeopardizing due process? *Journal of Applied Social Psychology*, 23, 603-631. DOI: 10.1111/j.1559-1816.1993.tb01105.x
- Tobey, A., Goodman, G. S., Batterman-Faunce, J. M., Orcutt, H., & Sachsenmaier, T. (1995). Balancing the rights of children and defendants: Effects of closed-circuit television on children's accuracy and jurors' perception. In M. Zaragoza, J. Graham, G. Hall, R. Hirschman, & Y. Ben-Porath (Eds.), *Memory and testimony in the child witness* (pp. 214-239). Thousand Oaks, CA: Sage.
- Wells, G. L., & Windschitl, P. D. (1999). Stimulus sampling and social psychological experimentation. *Personality and Social Psychology Bulletin*, 25, 1115-1125. DOI: 10.1177/01461672992512005
- Wigmore, J. H. (1974). *Evidence in trials at common law*. Boston, MA: Little, Brown & Company.
- Zajac, R., Irvine, B., Ingram, J., & Jack, F. (in press). *The diagnostic value of children's responses to cross-examination questioning*. *Behavioral Sciences & the Law*.
- Zajac, R., O'Neill, S., & Hayne, H. (2012). Disorder in the courtroom? Child witnesses under cross-examination. *Developmental Review*, 21, 181-204. DOI: 10.1016/j.dr.2012.06.006

Figure 1. The number of participants who believed each allegation as a function of transcript condition in Experiment 1. $N = 16$ in each condition.

Figure 2. The number of participants who believed each allegation as a function of transcript condition in Experiment 2. $N = 16$ in each condition.

Appendix 1

Questions regarding the circumstances of Jasmine's allegation(s)

1. How likely is it that Jasmine made up the allegation(s) against Warren?*
2. How likely is it that Jasmine is telling the truth about the allegation(s) against Warren?
3. How likely is it that Jasmine honestly believes she was assaulted by Warren when really she was not?
4. How likely is it that Jasmine's statements were influenced by the specialist interviewer?
5. How likely is it that Jasmine reported all the major details about the event(s)?#
6. How likely is it that Jasmine remembers all the major details about the event(s) with Warren?#
7. How likely is it that Jasmine reported all the minor details about the event(s) with Warren?#
8. How likely is it that Jasmine remembers all the minor details about the event(s) with Warren?#
9. How likely is it that Jasmine completely understood what happened between her and Warren?#
10. How likely is it that Jasmine is afraid to tell everything that happened with Warren?
11. How likely is it that Jasmine completely understood all of the questions asked by the specialist interviewer?#
12. How likely is it that Jasmine's memory of the event(s) is affected by stress?
13. How likely is it that Jasmine confused an earlier abusive experience with what happened with Warren?
14. How likely is it that Jasmine forgot to tell the specialist interviewer about some things that happened because she was nervous?
15. How likely is it that Jasmine had a motive to make a report against Warren?*
16. How likely is it that some of the details that Jasmine reported are not accurate?*
17. How likely is it that an adult convinced Jasmine to make a false report against Warren?
18. How likely is it that if Jasmine were questioned again, her story would change?*

* Responses used to compile honesty score

Responses used to compile cognitive competence score