

JUDICIAL DECISION MAKING IN TIMELY AND DELAYED PROSECUTIONS OF CHILD SEXUAL ABUSE IN CANADA

A Study of Honesty and Cognitive Ability in Assessments of Credibility

Deborah A. Connolly
Simon Fraser University

Heather L. Price
University of Regina

Heidi M. Gordon
University of Toronto

Child sexual abuse is an alarmingly common criminal offense. Whether prosecutions occur shortly after the alleged offense or after a lengthy delay, complainant credibility is often the central issue at trial. In both law and in psychology, credibility is said to be a function of two relatively distinct factors: honesty and cognitive ability. Complainant age informs evaluations of both such that younger children are seen as more honest but less cognitively competent than older children and adults. When a complainant describes a recent event, current age may be used to assess honesty and cognitive ability. However, when a complainant describes an event that occurred in the distant past, we argue that current age is most informative in evaluations of honesty, whereas age at the time of the alleged offense should inform evaluations of cognitive ability. In this research, we analyzed judicial assessments of complainants' credibility in 52 timely (child complainant) and 49 delayed (adult complainant) criminal prosecutions of child sexual abuse. Judicial comments concerning cognitive ability suggest that adults were viewed more positively than children, despite the fact that all complainants were children when the alleged offense occurred. As expected, comments related to honesty suggested that children were seen to be more honest than adults unless they had been exposed to suggestive influences.

Keywords: historic child sexual abuse, criminal prosecutions, credibility

Imagine two criminal trials involving allegations that a child had been sexually abused by her stepfather on several occasions between the ages of 6 and 8 years. The complainants remained silent about the abuse throughout its duration.

Deborah A. Connolly, Department of Psychology, Simon Fraser University, Burnaby, British Columbia, Canada; Heather L. Price, Department of Psychology, University of Regina, Regina, Saskatchewan, Canada; Heidi M. Gordon, Department of Human Development and Applied Psychology, University of Toronto, Toronto, Ontario, Canada.

This research was supported by a Social Sciences and Humanities Research Council (SSHRC) operating grant to Deborah A. Connolly, a Natural Sciences and Engineering Research Council graduate scholarship to Heather L. Price, and a SSHRC graduate scholarship to Heidi M. Gordon. We thank Jocelyn Conway for her invaluable assistance with data coding.

Correspondence should be sent to Deborah A. Connolly, Department of Psychology, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia V5A 1S6 Canada. E-mail: debc@sfu.ca

Now, imagine that, in one case, the complainant was 9 years old at trial and in the other case she was 45 years old at trial. Imagine further that both cases turn on the perceived credibility of the complainants and complaints. The complainants, accused persons, and allegations are very similar; however, the ages of the complainants at trial and the delays to prosecution are vastly different. How are such assessments made in cases that are, at the same time, strikingly similar and markedly dissimilar?

Although we use the term *imagine* liberally, these scenarios are not remarkable. Child abuse is distressingly common. According to the World Health Organization (2006), approximately 40 million children are abused every year. In Canada, children and youth account for 61% of all reported sexual assaults and 21% of all physical assaults (AuCoin, 2003). Moreover, delayed disclosure/prosecution is common. According to Lewis (2006), the incidence of prosecutions of child sexual abuse (CSA) that is alleged to have occurred in the distant past is substantial in each of the jurisdictions she studied (i.e., Canada, United States, Australia, New Zealand, England, Scotland, Wales, and Ireland). In many cases involving timely and delayed prosecutions of CSA, trial outcome turns on evaluations of credibility (e.g., Castelli, Goodman, & Ghetti, 2005; Goodman, Batterman-Faunce, & Schaaf, 2002; McCarron, Ridgway, & Williams, 2004; Pezdek et al., 2004). Assessments of credibility have been described as “notoriously difficult” (*R. v. Marquard*, 1993, para. 49) when the prosecution is commenced shortly after the alleged offense was committed—in all likelihood, this complexity is compounded when there is a very long delay to trial.

It was not long ago that neither of the scenarios described above would have proceeded to trial. Laws in most common-law jurisdictions had the effect of frustrating the prosecution for timely and delayed allegations of CSA. Most complaints of CSA made around the time of the offense could not proceed to trial because the complainants’ evidence was the subject of a judicial warning: Triers of fact were told that it would be dangerous to convict on a child’s evidence alone (Van Tongeren Harvey & Dauns, 2001). Most complaints of CSA made some time *after* the offense occurred would not have proceeded to trial because the complainants’ evidence was the subject of a judicial warning: Triers of fact were told that failure to complain at the time of the offense was evidence of falsity of the allegation (Lewis, 2006). In short, until relatively recently, a criminal remedy for CSA was out of reach for most child victims. During the mid- to late-1980s, several of these laws were abrogated in Canada, and they were abrogated or significantly modified in most other common-law jurisdictions (see Connolly & Read, 2003, for a more detailed discussion of the laws and their abrogation or erosion). Coincident with these legal changes was a substantial increase in the incidence of criminal prosecutions of both timely and delayed CSA allegations (Connolly, Price, & Gordon, 2009; Connolly & Read, 2003; Lewis, 2006), and the challenge of assessing credibility of children and adults reporting childhood events became much more prominent in criminal courts. Throughout this article, we refer to timely prosecutions as CSA and delayed prosecutions as historic CSA (HCSA).

In this article, we examine credibility assessments in criminal prosecutions involving similar complaints but different delays to prosecution (immediate

versus delayed)¹ and complainants of different ages. First, we review the psychological literature on credibility. In particular, we focus on the age of the speaker as a central feature that informs evaluations of honesty and cognitive ability. We use psychological theory to argue that evaluations of overall credibility should be different in CSA and HCSA prosecutions. This is followed by a discussion of overall credibility evaluations in Canadian case law and relevant legal scholarship. We highlight the striking similarities between legal and psychological factors allied with overall credibility. As in our discussion of psychological theory, we propose that age of the speaker should be viewed differently in CSA and HCSA cases when it is used to inform assessments of honesty and cognitive ability. In this section, support for our assertion comes from the Supreme Court of Canada's direction on how to assess overall credibility in adults who allege having been sexually abused as children. This review of psychological and legal scholarship is followed by an examination of the actual relative salience given by judges to honesty and cognitive ability in a sample of CSA and HCSA criminal prosecutions in Canada.

Psychological Scholarship on Perceived Credibility

An overall evaluation of credibility has been conceptualized as involving two constructs: perceived honesty and perceived cognitive ability (Bottoms, 1993; Goodman, Bottoms, Herscovici, & Shaver, 1989; Ross, Dunning, Toggia, & Ceci, 1990; Ross et al., 2003). Generally, children are seen as more honest but less cognitively competent than adults. In CSA cases involving child complainants, honesty is said to be more salient than cognitive ability and so children are often judged to be more credible than adults (e.g., Bottoms, 1993). This is due, in large part, to (a) children's perceived impoverished cognitive sophistication and consequent inability to fabricate a convincing allegation concerning events about which they have little knowledge (i.e., sex; Bottoms, 1993) and (b) children's perceived innate innocence. Indeed, lying is cognitively complex and the ability to do so convincingly develops throughout the childhood years (e.g., Peterson, Peterson, & Seeto, 1983; Wimmer, Gruber, & Perner, 1984). Furthermore, there is compelling evidence that child victims are seen as innocent and vulnerable (e.g., Meyer, 2007). It is our position that this inability/unwillingness to lie concerns the person at the time he or she describes a past event, not his or her age at the time the alleged event occurred. Accordingly, in both CSA and HCSA prosecutions, the complainant's honesty should be evaluated on the basis of his or her age at trial.

We now turn to a discussion of cognitive ability. To begin, two clarifications are important. First, this discussion assumes an ability to encode and retrieve

¹ To the best of our knowledge, the term *delay* has not been defined in the legal or psychological literature. For the purposes of our research, we defined *delay* as 2 or more years from the end of the alleged abuse to trial, because this is the modal limitations period on personal injury cases in civil courts in Canada and the United States (see <http://www.canlii.org> for access to each statute of limitations in each province and territory in Canada and Nolo, 2005, for statutes of limitations in the United States).

verbal memories. Although memories may be observed with implicit measures (e.g., Cordon, Pipe, Sayfan, Melinder, & Goodman, 2004), our interest is in verbal recall of past events because in criminal prosecutions verbal recall is assessed. Accordingly, the following discussion does not apply if the alleged offense occurred when the child was younger than about 3 years old when, because of infantile amnesia, verbal memories are likely to be inaccessible (e.g., Howe, Courage, & Peterson, 1994). Second, in most trials there is a delay from the time a complaint is made to trial. Martone, Jaudes, and Cavins (1996) reported that, on average, 12 to 16 months expired between the time a charge was made and the trial was complete. Accordingly, delayed recall (and its effect on memory reports, as discussed below) is present in both CSA and HCSA cases.

Both the complainant's age at the time of the offense and her age at trial are relevant considerations in an assessment of cognitive ability, but, we argue, the complainant's age at the time of the offense should carry more weight. Ornstein, Ceci, and Loftus (1998) reasoned that, to understand what adults may remember about childhood events, one must consider three stages of remembering: encoding, retention, and retrieval. First, we submit that what a person understands and encodes about an event does not vary as a function of her age at the time of reporting. That is, what a witness actually encoded about an event that occurred when she was, say, 5 years old does not vary as a function of her age at trial (e.g., 8, 25, or 45 years old). Only the complainant's age when the event occurred is relevant to understanding what he or she *may* be able to remember at trial. Second, although the preservation of information is expected to vary as a function of the retention interval (for a review, see Read & Connolly, 2007), the forgetting curve for long-term autobiographical memory is not linear: It has been described as a power function wherein most forgetting occurs within a few years of the event and the absolute amount of subsequent forgetting is relatively small (Rubin & Wenzel, 1996). The slope of the forgetting curve may vary as a function of the strength of the initial memory trace such that a stronger trace has a slower rate of forgetting than a weaker trace (Wixted, 2004). Furthermore, the forgetting curve for peripheral details is steeper than the forgetting curve for central details (for a review, see Read & Connolly, 2007) and for the gist of an event (Brainerd & Reyna, 1990, 1995). In fact, memory for the most central details and for the gist of an emotionally salient event can be quite resistant to forgetting, even over a lengthy retention interval (e.g., Brewin, Andrews, & Gotlib, 1993; Hardt & Rutter, 2004; Peterson, 1999; Peterson & Whalen, 2001). Given this, and given that there is often a substantial delay to trial even in timely complaints, we submit that the differences in absolute amount of forgetting of central or gist details in CSA and HCSA complaints may be relatively small. Thus, in terms of retention, a longer delay in HCSA versus CSA cases may not be weighty enough to shift the emphasis from complainant age at the time of the alleged abuse to her age at trial. Third, retrieval is a present concern, and the age of the complainant at trial is most relevant to understanding the influence of retrieval attempts on accuracy of memory reports. In summary, the age of the complainant when the alleged offense occurred and her age at trial are relevant when evaluating cognitive ability, although her age at the time of the offense may have a greater influence on what she is able to recall and should carry more weight in a legal analysis.

Jurisprudence

The notion that overall credibility is a function of two distinct constructs described in the psychological literature as honesty and cognitive ability is also found in law. The most obvious connection can be found in assessments of the very ability, or competence, of a witness to provide viva voce evidence. In most common-law jurisdictions, judges may or must inquire into a child's capacity to testify (see Connolly, Gagnon, & Lavoie, 2008, for a brief description of the relevant law in different common-law jurisdictions). Typically, there are two branches to the competence inquiry: the child's comprehension of and commitment to truth telling and the child's cognitive capacity to communicate the evidence (interpreted as the ability to perceive, recollect, and communicate the evidence; Bala, Lee, Lindsay, & Talwar, 2000; Lyon, 2000).

When a witness is found competent to testify, his or her overall credibility is in issue. We argue that this assessment is based on the two factors that psychologists call honesty and cognitive ability. In case law, a distinction is made between credibility and reliability. We argue that credibility is analogous to psychologists' notion of honesty, as should be clear from its description in *R. v. Pinnock* (2004):

[T]o assess M.T.'s credibility . . . I have considered the following evidence. M.T. has a criminal record that includes offences of dishonesty. He has, in giving his evidence tried to repudiate prior statements he made to people in authority. Detective M. Young testified prior to this application that he had concerns from the outset with respect to M.T.'s candour. There was evidence presented that M.T. was trying to hold the police at ransom . . . Clearly M.T. was an individual prepared to manipulate or try and manipulate the process to his advantage. (para. 10)

Although we have been unable to locate a definition of reliability in case law, its analysis places it in a distinct category and, we argue, aligns it with psychologists' notion of cognitive ability. That is, it is concerned with the evidence itself and includes issues such as the apparent completeness of the report (e.g., *R. v. Robichaud*, 2007; *R. v. Stea*, 2007), clarity of the report (*R. v. Stea*), consistency (e.g., *R. v. Stewart*, 1994), corroboration (e.g., *R. v. Stea*), apparent accuracy of the report (e.g., *R. v. Joudrie*, 1997; *R. v. Stea*), and ability to observe (e.g., *R. v. Ross*, 2006). These measures are analogous to the measures Ross et al. (2003) used to assess cognitive ability: memory for details, consistency, accuracy, event comprehension, suggestibility, and intelligence. Arguably, the two constructs referred to as *reliability* by legal professionals and *cognitive ability* by psychologists are comparable.

Canada's highest court, the Supreme Court of Canada, distinguished between credibility and reliability in several recent cases (e.g., *R. v. Gagnon*, 2006; *R. v. L. M.*, 2008). In *R. v. Joudrie* (1997) the Ontario Court of Appeal ordered a new trial, in part, because the trial judge failed to distinguish credibility and reliability for the jury. The court held that the jury may have put too much weight on the appearance of honesty and not enough weight on inaccuracies with respect to dates of the alleged offenses.

The theoretical distinction between credibility (honesty) and reliability (cognitive ability) has important explanatory power. For instance, as discussed earlier,

it explains why, in some circumstances, overall credibility ratings are higher for children than adults. The practical significance of the distinction was articulated in *R. v. Stewart* (1994), where the Ontario Court of Appeal stated:

We all know from our personal experiences as trial lawyers and judges that honest witnesses, whether they are adults or children, may convince themselves that inaccurate versions of a given event are correct and they can be very persuasive. The issue, however, is not the sincerity of the witness but the reliability of the witness's testimony. Demeanor alone should not suffice to found a conviction where there are significant inconsistencies and conflicting evidence on the record. (para. 19)

Notwithstanding different epistemologies, the literature in psychology and law converge on the understanding that overall credibility consists of two distinct elements, honesty (or credibility) and cognitive ability (or reliability). In this context, we introduce the Supreme Court of Canada direction vis à vis assessing overall credibility in HCSA cases:

In general, where an adult is testifying as to events which occurred when she was a child, her credibility should be assessed according to criteria applicable to her as an adult witness. Yet with regard to her evidence pertaining to events which occurred in childhood, the presence of inconsistencies, particularly as to peripheral matters such as time and location, should be considered in the context of the age of the witness at the time of the events to which she is testifying. (*R. v. W.(R.)*, 1992, p. 134)

We argue that the Supreme Court of Canada directed Canadian criminal courts to consider the age of the witness at trial when assessing honesty (or credibility) and to consider the age of the witness at the time of the alleged offense when considering cognitive ability (or reliability).

In summary, in CSA cases, perceptions of credibility are often determinative of case outcome (e.g., Castelli et al., 2005; Goodman et al., 2002; McCarron et al., 2004; Pezdek et al., 2004). In all likelihood, the role of perceived credibility is at least as central, if not more so, in HCSA cases as the passage of time is likely to have diminished or eliminated any corroborative evidence that might have been available at the time of the alleged offense. However, we know very little about how judges come to their decisions. In the present study, we begin to investigate this issue by looking at the relative salience of perceptions of honesty and cognitive ability in CSA and HCSA cases separately.

The Present Study

That overall credibility is a function of two components—honesty and cognitive ability—is consistent with psychological literature, case law, and legal scholarship. When a listener assesses honesty and cognitive ability, the age of the speaker is often considered. When the speaker is reporting a recent autobiographical event, his or her current age is the appropriate age to consider in evaluations of both honesty and cognitive ability. However, when the speaker is reporting a long-past autobiographical event, two ages are relevant: the speaker's current age and his or her age at the time of the reported event. Psychological literature, as

well as case law and legal scholarship, support this assertion. In the present study, we ask the question, “Do trial judges make such a distinction?”

We randomly selected 52 CSA and 49 HCSA complaints adjudicated by judges alone.² Judicial comments related to overall credibility concerning each complaint were coded as related to memory for the alleged offense, credibility, or reliability (using the system developed by Connolly et al., 2009). In most cases, judges did not provide an evaluation for each comment. A judge may say something like “[t]hat statement was written for the Sexual Assault Centre. In it she did not mention J.L.S.A. kissing her” (*R. v. J.L.S.A.*, 2001, para. 32) or “the complainant has limited education” (*R. v. Audet*, 1999, para. 29) without stating whether the detail was considered neutral, positive, or negative vis à vis the complainant’s overall credibility. We suspect that this is not an oversight on the part of the judges. Rather, we speculate that it is purposeful and intended to avoid having each comment become a potential ground of appeal. Whatever the reason, we could not code the valence of the vast majority of comments. Although imperfect, we attempted to gather some data concerning valence by including verdict as an independent variable in the analyses. In convict cases, the judges held that the essence of the complainants’ allegations were true, whereas in acquit cases, the complainants’ evidence was insufficient to meet the standard of proof. Accordingly, our premise is that comments made in convict cases are more commonly positive than negative vis à vis the complainant, whereas the reverse is true in acquit cases. It is almost certainly true that verdict does not always accurately reflect the valence of judicial comments. For instance, a judge may comment on the demeanor of the complainant at trial believing, without stating, it to be inappropriate. However, other evidence is powerful enough to support a conviction. We see no way to address this problem directly. We suggest, however, that across almost 8,000 judicial comments, on average, our assumption is reasonable.

Following from the literature review of psychological scholarship and jurisprudence, we expect that there will be more concerns about honesty in HCSA than in CSA cases. However, given that age at the time of the alleged offense should carry more weight when assessing cognitive ability, we do not expect differences as a function of type of case (CSA vs. HCSA) in concerns about cognitive ability.³

² Only bench trials were included for two reasons. First, jurors do not provide reasons, and so it is not possible to evaluate their decisions. Second, in some jurisdictions it is possible to interview jurors after a case to better understand the reasons for their verdict. This is not possible in Canada, as it is an offense for jurors to discuss the content of their deliberations outside of the jury room (*Criminal Code*, RSC 1985, s. 649).

³ Some of the case law we report and rely on was released after the cases analyzed in the present data. The Supreme Court of Canada case that directed trial judges to consider the age of the complainant at trial in evaluations of credibility and his or her age at the time of the offense in evaluations of reliability was released in 1992 (*R. v. W. (R.)* 1992) and should have informed evaluations in the cases included in these analyses. More recent case law further developed the distinction between credibility and reliability and, as such, may have aided the application of this direction.

Method

We obtained full-text reports of trial decisions involving CSA from 1998 to 2002 using Quicklaw, a Canadian database that contains all Supreme Court of Canada decisions, decisions from provincial Courts of Appeal, written decisions from the provincial Superior Courts, and written decisions from provincial Courts that were forwarded to Quicklaw.⁴ A search strategy was used that included the key words *child* (and variations thereof) and the following offenses: *sexual offence(s)*, *sexual assault*, *sexual interference*, *sexual intercourse*, *gross indecency*, *indecent assault*, *incest*, *rape*, *bestiality*, and *buggery*. All cases that included the search term *child*, as well as at least one of the offense terms and that involved a complainant under the age of 19 years when the alleged offense began, were included. A total of 468 cases were located: 208 included CSA complaints and 284 cases included HCSA complaints.⁵ There were 892 complainants in the full data set: 321 CSA complainants and 571 HCSA complainants. From these data, we selected 60 cases involving 101 complainants: 32 cases included at least one CSA complaint (52 complainants) and 32 cases involved at least one HCSA complaint (49 complainants).⁶ Sample cases were selected so that within each of the CSA and HCSA sample files, there was an approximately equal number of convict and acquit cases (23 convict cases in the CSA and HCSA data sets) and the frequency of alleged abuse was similar (in 26 CSA and 23 HCSA cases, the complainants alleged more than one instance of abuse). Cases were also matched for approximate complainant age at the time the alleged offense began. The sample is described in Table 1.

Each complaint was independently coded for presence of complainant-related elements that were discussed by the judge in rendering his or her decision. First, we investigated how judges described the alleged offense. In particular, we looked at whether the judge reported specific details about particular instances of the allegation or whether he or she described the offense(s) in general terms or made comments about memory failures. Second, we investigated how judges described their evaluations of credibility. The coding system was developed to study these cases as well as the cases described in Connolly et al. (2009; see the Appendix for a description of the coding). Coding was continuous in that the coder coded the data reported here immediately after she coded the data described in Connolly et al. (2009). Six major categories and 12 minor categories were coded; a definition for each minor category is provided in the Appendix. Reliability data were obtained at the beginning of the combined project. Overall kappa was 0.77 for assignment of comments to the 18 categories. Once acceptable intercoder reliability was achieved, one person coded all of the cases.

⁴ Decisions that judges believe should be available to the legal community were forwarded to Quicklaw. The reasons that judges forward decisions are likely varied and may be idiosyncratic. It is important to note that this is not an exhaustive set of criminal cases heard between 1998 and 2002. The data set, however, contains most decisions that are available to the political, legal, and research communities and upon which policy and practice are informed.

⁵ In 23 cases, there were both CSA complainants and HCSA complainants, so the total number of CSA ($n = 208$) and HCSA ($n = 284$) cases exceeds the total number of cases accessed ($n = 468$).

⁶ In four cases, there were both CSA and HCSA complainants, and so the total number of CSA ($n = 32$) and HCSA ($n = 32$) cases exceeds the total number of cases accessed ($n = 60$).

Table 1
Profiles of CSA and HCSA Cases

Variable	CSA	HCSA	Test of differences
Complainant			
Age began	10.92 (3.63)	9.38 (3.81)	$t(83) = 1.89, p = .06$
Age ended	12.03 (3.70)	11.54 (3.40)	$t(81) = 0.62, p = .54$
Age trial	13.40 (3.56)	25.52 (12.88)	$t(79) = 5.41, p < .01$
Gender (% male)	15.5	39.1	$\chi^2(1, N = 91) = 6.34, p = .01$
Accused			
Age began	37.73 (17.06)	34.41 (12.89)	$t(56) = 0.84, p = .40$
Age ended	38.65 (16.91)	37.06 (12.49)	$t(56) = 0.41, p = .68$
Age trial	39.36 (17.27)	50.16 (14.51)	$t(55) = 2.56, p = .01$
Gender (% male)	100	97.92	$\chi^2(1, N = 93) = 0.95, p = .33$
Duration (in months)	10.72 (16.71)	25.99 (31.16)	$t(82) = 2.71, p < .01$
Delay to trial (in months)	15.13 (5.53) (range = 6–24)	163.19 (119.48) (range = 36–408)	$t(84) = 7.72, p < .01$
Frequency			$\chi^2(1, N = 90) = 0.32, p = .57$
% 1–3 times	68.29	51.15	
% 4 or more times	42.86	48.78	
Relationship			$\chi^2(3, N = 84) = 1.31, p = .72$
% Parent	15.79	19.56	
% Other relative	15.79	23.91	
% Nonfamily connection	31.58	26.09	
% Community connection	36.84	30.43	
Allegation			$\chi^2(2, N = 88) = 8.14, p = .017$
% Fondling	59.10	29.54	
% Nonpenile penetration	22.73	45.45	
% Penile penetration	18.18	25.00	
Verdict			$\chi^2(1, N = 101) = 0.08, p = .78$
% Acquit	55.77	53.06	
% Convict	44.23	46.94	

Note. CSA = childhood sexual abuse; HCSA = historic child sexual abuse.

Comments were organized into three major categories: memory for the offense, credibility, and reliability. General and specific details of the allegation, as well as comments about memory failures, were classified as memory details. As discussed earlier in the Jurisprudence section, case law provided some direction on how to define credibility and reliability for the present study. Reliability concerned the evidence itself and included comment concerning inconsistencies, corroboration, and overall quality. Credibility concerned comments related to the conduct of the complainant himself or herself and included comments about his or her conduct around the time of abuse, conduct around the time of disclosure, and conduct at trial. A fourth major category, judicial inferences, concerned assumptions made by the judge without an apparent evidentiary basis.

Only units of information that were relevant to complainants' credibility were coded; information pertaining to the defendants' or other witness's evidence was not coded unless it related directly to the complainants' evidence. Each unit of information was coded only once, regardless of how often it was mentioned by the judge. The coder did not read the outcome of the case until she had finished coding the case. When there were two or more complainants in a case, comments related to each complainant were coded separately. If a comment related to more than one complainant, one code was applied to each complainant.

Results

A total of 7,776 comments were coded, and frequencies of comments are reported in Table 2. Twenty-seven comments were unclassified and are not discussed further.⁷ The most common classification concerned memory for the offense (45% of all comments), followed by comments classified as credibility (32% of all comments), and then comments classified as reliability (18.3% of all comments). The percentage of comments in each classification was similar in CSA and HCSA cases; although there was a large numeric difference in reliability comments, the difference was not significant, $\chi^2(1, N = 1427) = 1.08, p > .05$. This distribution of comments is consistent with an earlier report using a different set of HCSA cases (Connolly et al., 2009), in which judicial comments related to credibility were far more common than those related to reliability.

The numbers of comments expressed as a percentage of all judicial comments (excluding unclassified comments) in each minor category were the dependent variables. Using Wilks's criterion, we found a significant multivariate main effect of type of case, $F(17, 81) = 2.52, p < .01, \eta_p^2 = .35$; and the multivariate interaction between type of case and verdict was significant, $F(17, 81) = 2.53, p < .01, \eta_p^2 = .35$. We conducted a series of 2 (type of case: HCSA or CSA) \times 2 (verdict: convict or acquit) analyses of variance, one for each dependent variable.⁸ The means are in Table 3. As can be seen in Table 1, CSA and HCSA cases varied on the following dimensions: duration of offense, the nature of the allegation, age of the complainant when the alleged offense began, and gender of the complainant. We reran the analyses and concurrently covaried these variables. With the following exception, the results were identical; The main effect of case

⁷ Examples of comments that were not classified are as follows: The judge found this case to be "troubling"; the judge did not find the defense questions confusing, but he acknowledged that the complainant may have a different perception; and the judge acknowledged a lack of parental involvement with the complainant.

⁸ In 16 cases, two or more complainants were adjudicated together. This may compromise the assumption of independence of ANOVA, notwithstanding that judges are directed to adjudicate each complaint independent of the others. To test whether our conclusions were compromised, we omitted all cases that included two or more complaints. The remaining data included only 44 complaints, and cell sizes were somewhat unequal (ranging from 9 to 13 complaints), resulting in a substantial loss of power. When this analysis was run, the effect of case type on reasons for delay and the assumption that someone would have known were no longer significant. Given the small effect sizes in the present data, it is not surprising that some effects were lost when our sample size was cut in half. It is our conclusion that a loss of power explains the effect losses. Of course, some readers may conclude that the effects we describe in this report are attributable to a possible violation of the independence assumption of ANOVA.

Table 2
No. of Judicial Comments in Each Major and Minor Category Within Each Class of Comments

Judicial comment	Frequency		% of total		Absent	
	CSA	HCSA	CSA	HCSA	CSA	HCSA
Memory						
Overall	1,418	2,088	41.96	47.49	1	0
Specific incidents	1,040	1,381			9	6
General details	307	533			27	18
Memory failures	71	174			29	22
Reliability						
Overall	742	685	21.96	15.58	0	0
Inconsistencies	202	207			7	10
Corroboration	186	173			16	19
Quality of evidence	354	305			4	4
Credibility						
Overall	1,098	1,431	32.49	32.54	0	1
Conduct associated with abuse	523	708			8	2
Conduct after alleged assault	351	530			10	4
Resistance to assault	172	178			18	14
Disclosure	306	418			6	3
Reasons for delay	86	188			26	12
Possibility of Collusion	60	52			30	25
Motive to Fabricate	89	91			28	20
Conduct after disclosure	41	31			39	38
Reason for disclosure	30	56			32	16
Conduct at trial	269	305			15	12
Demeanor	206	204			16	13
Personality	63	101			31	25
Inferences						
Overall	106	181	3.14	4.12	21	9
Someone should have known	17	57			41	19
General inferences	89	124			26	19
Other comments	15	12			40	40
Overall total	3,379	4,397	100.00	100.00		

Note. The major categories are presented in bold font and the minor categories are presented in standard font.

type on general details was lost. Here we report the main analyses, rather than the analyses of covariance. The results are organized in three paragraphs, one concerning each of the following: memory for the offense, credibility, and reliability, respectively.

The way judges described complainants' memory for the alleged offense varied as a function of type of case as shown in the main effect of general details, $F(1, 97) = 4.99, p < .05, \eta_p^2 = .05$. Specifically, there were more general descriptions in cases involving HCSA than in cases involving CSA. However, given that this effect disappeared in the analysis of covariance, it is not clear if the effect is attributable to delay or another variable.

Table 3
Mean Percentages (and Standard Deviations) of Judicial Comments as a Function of Case Type and Outcome

Judicial comment	Acquit		Convict		Marginal means		Marginal means	
	CSA	HCSA	CSA	HCSA	CSA	HCSA	CSA	HCSA
Memory								
Specific details	22.75 (18.88)	22.81 (15.50)	22.78 (17.21)	33.37 (20.02)	36.31 (17.47)	34.84 (18.63)	27.45 (19.92)	29.15 (17.64)
General details ^a	6.21 (11.02)	14.69 (15.84)	10.22 (14.05)	7.78 (9.50)	10.57 (13.05)	9.17 (11.37)	6.91 (10.31)	12.75 (14.60)
Memory failure	2.01 (3.40)	4.00 (4.83)	2.95 (4.21)	2.08 (3.49)	1.47 (2.35)	1.77 (2.96)	2.04 (3.40)	2.81 (4.03)
Reliability								
Inconsistencies	6.92 (5.07)	6.79 (7.25)	6.86 (6.14)	4.49 (3.25)	3.31 (3.39)	3.90 (3.34)	5.84 (4.49)	5.16 (5.98)
Corroboration	3.44 (4.63)	3.47 (4.10)	3.46 (4.35)	6.69 (5.88)	5.21 (8.97)	5.95 (7.54)	4.88 (5.41)	4.29 (6.81)
Quality evidence ^b	19.00 (13.87)	8.37 (5.90)	13.97 (12.03)	7.84 (7.90)	5.75 (3.56)	6.79 (6.15)	14.06 (12.08)	7.14 (5.07)
Credibility								
Conduct/assault	8.59 (10.93)	12.46 (8.79)	10.42 (10.07)	9.95 (6.43)	10.25 (8.76)	10.10 (7.60)	9.19 (9.16)	11.43 (8.75)
Resistance	2.78 (3.53)	2.61 (2.95)	2.70 (3.24)	5.10 (4.52)	4.50 (3.52)	4.80 (4.02)	3.81 (4.13)	3.50 (3.33)
Reasons/delay ^a	1.34 (1.63)	3.22 (2.95)	2.23 (2.51)	2.40 (2.85)	4.17 (3.11)	3.28 (3.08)	1.81 (2.29)	3.66 (3.03)
Collusion	2.84 (5.21)	1.34 (2.79)	2.13 (4.27)	3.41 (4.66)	2.38 (3.24)	2.89 (4.01)	3.09 (4.93)	1.83 (3.02)
Fabrication ^b	4.87 (6.53)	1.02 (1.57)	3.05 (5.20)	1.29 (2.08)	3.16 (3.02)	2.23 (2.73)	3.29 (5.34)	2.03 (2.57)
Conduct/disclose	1.65 (2.65)	0.81 (1.55)	1.25 (2.22)	0.19 (0.72)	0.37 (1.12)	0.28 (0.94)	1.00 (2.15)	0.60 (1.37)
Reasons/disclose	1.71 (2.29)	1.64 (2.14)	1.68 (2.20)	0.64 (1.26)	1.81 (1.58)	1.22 (1.53)	1.24 (1.97)	1.72 (1.88)
Demeanor/ trial	5.28 (7.00)	5.38 (4.14)	5.33 (5.78)	7.84 (5.56)	4.00 (4.62)	5.92 (5.42)	6.41 (6.47)	4.73 (4.38)
Personality	1.95 (3.54)	3.21 (4.44)	2.54 (4.01)	2.31 (2.84)	1.63 (3.92)	1.97 (3.40)	2.11 (3.22)	2.47 (4.24)
Inferences								
Someone would have known ^b	0.42 (1.04)	2.39 (3.35)	1.35 (2.59)	0.50 (0.94)	0.37 (1.00)	0.43 (0.96)	0.45 (0.99)	1.44 (2.71)
General	3.99 (4.41)	2.72 (2.74)	3.39 (3.74)	1.74 (2.59)	2.31 (2.65)	2.03 (2.61)	3.00 (3.85)	2.53 (2.68)

Note. CSA = childhood sexual abuse; HCSA = historic child sexual abuse.

^aEntered into a main effect of type of case. ^bEntered into an interaction between type of case and verdict.

There were several differences in judicial comments about credibility of the complainant. The main effect of reasons for delayed disclosure, $F(1, 97) = 11.78$, $p < .01$, $\eta_p^2 = .11$, was significant because comments concerning delayed disclosure were more likely to be discussed in HCSA cases than in CSA cases. There was also an interaction between type of case and verdict in the percentage of comments concerning motive to fabricate, $F(1, 97) = 12.74$, $p < .01$, $\eta_p^2 = .12$. In acquit cases, judicial comments about motive to fabricate were more likely to be discussed in CSA cases than in HCSA cases, $t(53) = 2.92$, $p < .01$. Conversely, in convict cases, motive to fabricate was more likely to be discussed in HCSA cases than in CSA cases, $t(44) = 2.44$, $p < .05$. There was also an interaction between type of case and verdict in the percentage of judicial comments about the possibility that someone would have seen the offense, $F(1, 97) = 7.56$, $p < .01$, $\eta_p^2 = .07$. Judges were more likely to comment on their assumption that someone should have seen the offense in HCSA cases than in CSA cases but only when the case ended in an acquittal, $t(53) = 3.01$, $p < .01$; not when it ended in a conviction, $t(44) = 0.45$, $p = .65$.

In terms of reliability, judicial comments about the quality of the evidence varied as a function of case type and verdict $F(1, 97) = 5.59$, $p < .05$, $\eta_p^2 = .05$. Quality of the evidence comments were more common in CSA cases than in HCSA cases when the case ended in an acquittal, $t(53) = 3.62$, $p < .01$; but not when the case ended in a conviction, $t(53) = 1.16$, $p = .25$.

Discussion

Do judges discuss cases involving delayed versus timely prosecutions of CSA differently? In the introduction, we discussed the two components of overall credibility, honesty and cognitive ability, and argued that they are analogous to the legal concepts of credibility and reliability, respectively. We used this analysis to argue that the Supreme Court of Canada directed Canadian courts to consider complainant age at trial when evaluating honesty but to consider complainant age at the time of the offense when evaluating cognitive ability. Because children are seen as more honest but less cognitively competent than adults, this led us to the somewhat counterintuitive hypothesis that comments related to honesty would differ as a function of case type (more such comments in HCSA cases than in CSA cases) but that comments about cognitive ability would not.

Honesty/Credibility

Three minor categories provided significant results and inform this discussion: the assumption that someone should have seen, reasons for delaying disclosure, and motive to fabricate. Each effect is discussed in turn. First, in acquit cases, judges were more likely in HCSA complaints than in CSA complaints to discuss their assumption that someone should have seen the abuse. In the other three cells associated with this interaction, the percentages of such comments were trivial. Why would judges be more likely to assume that someone would have seen in HCSA acquit cases? One possibility is that in delayed allegations the accused was more commonly a close relative and so the alleged abuse occurred in a crowded environment, such as in the home. We looked at the percentage of cases that involved an accused parent in each of the four cells associated with the

interaction. In the HCSA acquit cell, only 8% of the accused persons were parents, whereas 17.4% of the accused persons were parents in the CSA acquit cell, 13.3% were parents in the CSA convict cell, and 33.3% were parents in the HCSA convict cell. It is possible, therefore, that the circumstances of the alleged offense were such that it was *less* likely to have been observed in HCSA acquit cases than in other cases. An alternative explanation is that this type of comment reflects judicial concern about honesty—if the allegation was true, surely someone would have noticed. These types of comments were more common in HCSA cases than in CSA cases, when the case ended in an acquittal. This is consistent with the notion that children are seen as more honest than adults.

Second, comments about reasons for delayed disclosure were more likely to have been discussed in HCSA cases than in CSA cases. The temporal distance between the end of the alleged abuse and trial was less than 2 years in CSA cases, whereas in HCSA cases it was 2 or more years. Accordingly, if there was a delay in CSA cases, it was relatively short and the reasons in some cases may have been systemic—having to do with delays in the justice system rather than the complainants' delayed disclosure. Accordingly, delayed disclosure simply may not have occurred in most CSA cases. However, there are two reasons that this explanation is unsatisfactory. First, in a very thorough literature review, London, Bruck, Ceci, and Shuman (2005) reported that up to two-thirds of CSA victims delay disclosing their abuse for some period of time. Also, the Supreme Court of Canada in *R. v. D.D.* (2000) held that delayed disclosure is the norm in CSA cases. Given this, we expect that some, if not most, of the child complainants in the present CSA sample delayed disclosure for some period of time. Second, in the present data, we defined delay as the temporal distance between the *end* of abuse and trial. About half of the child complainants alleged repeated abuse, and the average duration of abuse in the entire CSA sample was about 10 months. Accordingly, in most CSA cases, there was a delay from the *beginning* of the abuse to disclosure. Despite the fact that delayed disclosure was probably present in many CSA cases, comments about reasons for delaying reporting were present in just under half of the CSA complainants, but it was present in about 75% of the HCSA complaints. Historically in law, delayed disclosure was taken as an indication of dishonesty (see Connolly et al., 2009, for a more comprehensive discussion of this point). To the extent that this attitude continues to exist in courts of law, the heightened interest in reasons for delay in HCSA cases versus CSA cases may reflect increased concern about honesty in HCSA versus CSA cases. This provides further evidence for our hypothesis that children would be seen as more honest than adults.

A third effect is inconsistent with our hypothesis concerning age and perceptions of honesty. In convict cases, comments about motive to fabricate were more common in HCSA than in CSA cases. That the effect was only present in convict cases suggests that the comments related to the absence of motive to fabricate. Conversely, in acquit cases, comments about motive to fabricate were more common in CSA than in HCSA cases. Given the verdict in the cases we suggest that comments expressing concerns that there may have been a motive to fabricate were more common in CSA than HCSA cases. This is inconsistent with our expectation that children would be seen as more honest than adults. To better understand this unexpected effect, we looked more closely at the category of comments. Motive to fabricate

comments included anger, revenge, financial compensation, and external or suggestive influences. In CSA cases, a common concern involved external or suggestive influences: For instance, in *R. v. P. G.* (2002), the judge wrote, “The child resisted this initially, but, in the court’s opinion, succumbed eventually to the repeated insistence from her mother that others were responsible for the child’s injuries” (para. 52). Also, in *R. v. Westcott* (1999), the judge stated, “All of the mothers were anxious, agitated, and in an emotional state when they talked to their daughters” (para. 55). In *R. v. R. E. M.* (2001), the judge wrote, “Before providing a statement to the police L.G. spoke with three Grade six friends, . . . who agreed that Mr. M.’s behavior of touching “butts” was inappropriate” (para. 73). It is interesting that few nefarious motives were inferred by the judge. That is, children were not depicted as being evil or bad. In fact, one might characterize the descriptions of the children as innocent and naïve yet vulnerable to suggestive external forces to fabricate an allegation. In one case, the judge stated, in reference to the child complainants, “a very real possibility of collusion, collaboration and/or concoction (even if inadvertent) exists” (*R. v. Westcott*, 1999, para. 58). It is interesting that the judge allowed for the apparent concoction to be inadvertent, suggesting that the child was not to blame for having succumbed to external forces to fabricate an allegation. The child’s innate innocence and naïvete may be the very qualities that, independent of external forces, sustain perceptions of honesty. However, if external suggestive forces were present, children may be seen as less able than adults to resist the forces and continue to report honestly.

In the context of a CSA vignette, Castelli et al. (2005) reported that a child who was questioned in a suggestive way was judged by mock jurors to be less credible than a child who was questioned in a neutral manner (see also Tubb, Wood, & Hosch, 1999). Our data suggest that the basis of the effect is in judgments of honesty rather than judgments of cognitive ability. Moreover, we provide the first demonstration of this effect in a sample of bench trials.

Cognitive Ability/Reliability

With respect to cognitive ability/reliability, as expected, comments concerning inconsistencies and corroboration did not vary as a function of type of case. Furthermore, when the allegation was held to be true (i.e., convict cases), judicial comments about the quality of the evidence did not vary as a function of type of case. However, comments concerning quality of the evidence did vary when there was cause to distrust the allegation. In acquit cases, it was more common for judges to discuss quality of the evidence in CSA than in HCSA complaints. This category contained general remarks about the evidence, without reference to specific or general details. For instance, comments about the quality of the evidence included, “her evidence was quite cogent,” “her recollection of crucial pieces of evidence made the judge cautious,” and “a number of matters give the court pause as to the child’s identification of the [accused] as the assailant.” Given that, generally, children are less competent communicators than adults (Saywitz, Nathanson, & Snyder, 1993; Saywitz, Snyder, & Nathanson, 1999), it is reasonable to speculate that the “quality” of their communication would be inferior to that of adults, and this could translate into differences in the perceived quality of evidence. Children’s communicative competence, however, is not the only factor influencing comments about quality of the evidence. Concerns

regarding general inconsistencies, report completeness, and plausibility of the allegation were quite common. In summary, when the allegation was held to be true, comments that reflect complainants' cognitive ability did not vary as a function of type of case. However, when the allegation did not meet the standard of proof, judicial comments about cognitive ability, expressed as quality of the evidence comments, were more common in CSA cases than in HCSA cases.

In *R. v. W. (R.)* (1992), the Supreme Court of Canada directed lower courts to consider the complainant's age at the time of the alleged offense when considering questions of reliability. What a person understands and encodes about an event is critical to knowing what the person may later be able to recall. The Supreme Court of Canada further stated in *R. v. B. (G.)* (1990), "[w]hile children may not be able to recount precise details and communicate the when and where of an event with exactitude . . . this does not mean that they have misconceived what happened to them and who did it" (p. 55). In the same judgment, the Court held that "a flaw, such as a contradiction, in a child's testimony should not be given the same effect as a similar flaw in the testimony of an adult" (p. 54). Children and adults may not encode the same information and, whether the information is retrieved shortly after the event or many years later, the memory report can only contain information that was encoded. In the present research, we did not find differences as a function of type of case in the analyses of the more concrete categories of inconsistencies or corroboration. However, "quality of the evidence," arguably a much more elusive category, varied such that when there was reason for the judge to mistrust the allegation, children's evidence was seen as lower in quality than adults' evidence. Although this is consistent with the theory that children are perceived as less cognitively competent than adults, it is inconsistent with the Supreme Court of Canada's direction that the age of the complainant at the time of the offense is the most relevant age to consider when assessing reliability.

In summary, in the absence of reasons to be concerned that external suggestive influences were present, our data support the notion that complainants in CSA cases were seen as more honest than complainants in HCSA cases. However, if there was evidence that the complainant was exposed to external suggestive influences, children's perceived innocence and naïvete may have heightened their perceived vulnerability to such environments and made them appear less honest than adults. We speculate that the inherent perceived naïvete and innocence of children may backfire when circumstances are such that there is an apparent reason to lie about the alleged abuse. We found some support for our hypotheses that comments expressing concerns about cognitive ability would be similar in CSA and HCSA cases. This was observed in measures that are clearly related to reliability, inconsistencies, and corroboration. In a more elusive measure (quality of the evidence), our data suggest that children are viewed as less cognitively competent than adults reporting childhood events.

Limitations

The complainants' age at trial and delay to trial are confounded. We did not statistically control this confound for two reasons. First, age at trial and length of delay are naturally occurring confounds in this context. That is, the longer the

delay to trial, the older the complainant is likely to be. To statistically control the co-occurrence of these two variables would create artificial data that rarely occur in the courtroom. Second, in the present data, these two variables are very highly correlated ($r = .93$): Statistical control is not possible even if advisable. As a result, we cannot know whether the complainants' age at trial or the delay to trial accounts for differences in how judges describe overall credibility in CSA and HCSA cases. On the basis of psychological research and our analysis of case law, described in the introduction, we submit that age at trial is a more powerful explanation. However, we invite readers to draw their own conclusions.

We studied written decisions rather than trial transcripts. There are both practical and methodical reasons for this. By using written decisions rather than trial transcripts, we were able to investigate a much larger sample of cases (the cost and time involved in evaluating 101 complaints based on transcripts would have been prohibitive). More important, in our view, the information of interest is more likely to be present in decisions than trial transcripts. We were interested in studying the information that judges use to decide overall credibility CSA and HCSA cases. The mere presence of evidence at trial does not mean that it contributed to a judge's decision: Trial evidence may be accepted in whole, in part, or not at all. We reasoned that judges would report the evidence they considered in an explanation of their verdict.

In the present study, we matched cases on several characteristics that we felt were most relevant to the present research (i.e., number of complainants per case, frequency of the alleged offense, age of the complainant when the offense began [we attempted to match within 2 years, but despite this, HCSA complainants were slightly younger when the alleged offense began], and verdict). Even with a large data set from which to select cases, it was not possible to match on other variables and obtain a reasonable sample size. A few unmatched but measured variables varied as a function of type of case (duration, nature of the allegation, age of the complainant when the alleged offense began, and gender of the complainant), although an analysis of covariance suggests that they contributed little to the effects reported. However, there may be other unmeasured variables that contributed to the differences observed.

The effect sizes associated with the individual variables are very small. This is not surprising, given the complexity of judicial decision making. It is important to note that, in the context of criminal trials, a variable that provides small but significant explanatory power can have great practical significance if it raises a reasonable doubt or masks the presence of reasonable doubt. Also, although the individual variables did not explain a lot of the variance in type of case, a linear combination of dependent variables explained a modest amount of variance, as demonstrated in the multivariate effect size.

Intercoder reliability was not up to standard in the present study. Although our initial kappa value was good, once intercoder reliability was achieved, only one person coded the cases. It is possible that the criteria shifted over time. In retrospect, more rigorous and ongoing reliability should have been conducted. We are somewhat reassured that coding criteria did not shift considerably for the following reason: Coding of the present data and the data reported in Connolly et al. (2009) was conducted by the same person. Although the purpose of the present

study is quite different from that of Connolly et al., the two data sets were compared and found to be remarkably similar.

The sample places clear limits on generalizability of these data. We studied only criminal complaints involving sexual assaults against children. We have no data that would allow us to generalize our conclusion to other criminal offences or to civil trials. In civil trials, the standard of proof is “on a balance of probabilities” rather the criminal standard of “beyond a reasonable doubt.” It is possible that this difference in the standard of proof may affect how judges evaluate evidence. The selection criteria also limit the generalizability of these data. In Provincial Courts, where many cases are heard, only decisions that were forwarded to Quicklaw between 1998 and 2002 are included in these analyses. As noted in Footnote 4, we are not aware of a predictable set of criteria that inform judges’ decisions about which cases should be forwarded. Only bench trials were reported. We have no basis on which to speculate that similar effects would be found in cases heard by juries.

The research is archival and so causal conclusions cannot be drawn from these data. However, as noted throughout this article, there are several findings that could be the subject of controlled research.

Conclusion

Imagine two criminal trials involving similar sexual offenses against similarly aged children. In both cases, there is no evidence other than the allegation of the complainants and the denial of the accused. The outcomes rest on perceptions of credibility. Now, imagine that in one case the complainant is still a child, and in the other case the complainant is an adult. Should the credibility evaluations be similar, or are there reasons to weight credibility factors differently? Our evaluation of the psychological literature and case law suggests that there should be differences. Specifically, evaluations of cognitive ability should be similar because the most relevant age for such an assessment is when the offense occurred. We found support for this in our analyses of inconsistencies and corroboration. However, in analysis of the more ephemeral category “quality of the evidence,” the age of the complainant at trial may have carried more weight than is warranted. Evaluations of honesty, we argued, should vary as a function of the age of the complainant at trial. For the most part, our data support this and are consistent with the theory that children are seen to be more honest than adults. This finding must be qualified. Although children may be viewed as innately more honest, they may also be seen to be more vulnerable to external pressure to fabricate allegations. Researchers have shown that children who are questioned suggestively are seen as less credible than children who are interviewed in a neutral way. The present study is the first to demonstrate a similar effect with judges, and we suggest that the locus of the effect is in perceptions of honesty rather than perceptions of cognitive ability.

References

- AuCoin, K. (2003). Children and youth as victims of violent crime. *Juristat: Canadian Centre for Justice Statistics*, 25(1, Catalogue No. 85-002-XIE). Retrieved August 28, 2009, from <http://www.statcan.gc.ca/pub/85-002-x/85-002-x2005001-eng.pdf>

- Bala, N., Lee, K., Lindsay, R., & Talwar, V. (2000). A legal and psychological critique of the present approach to the assessment of the competence of child witnesses. *Osgoode Hall Law Journal*, 38, 409–451.
- Bottoms, B. L. (1993). Individual differences in perceptions of child sexual assault victims. In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims, child witnesses: Understanding and improving testimony* (pp. 229–261). New York: Guilford Press.
- Brainerd, C. J., & Reyna, V. F. (1990). Can the Age \times Learnability interactions explain the development of forgetting? *Developmental Psychology*, 26, 194–204.
- Brainerd, C. J., & Reyna, V. F. (1995). Learning rate, learning opportunities, and the development of forgetting. *Developmental Psychology*, 31, 251–262.
- Brewin, C. R., Andrews, B., & Gotlib, I. H. (1993). Psychopathology and early experience: A reappraisal of retrospective reports. *Psychological Bulletin*, 113, 82–98.
- Castelli, P., Goodman, G. S., & Ghetti, S. (2005). Effects of interview style and witness age on perceptions of children's credibility in sexual abuse cases. *Journal of Applied Social Psychology*, 35, 297–319.
- Connolly, D. A., Gagnon, N. C., & Lavoie, J. A. (2008). The effect of a judicial declaration of competence on the perceived credibility of children and defendants. *Legal and Criminological Psychology*, 13, 257–277.
- Connolly, D. A., Price, H. L., & Gordon, H. M. (2009). Judging the credibility of historic child sexual abuse complaints: How judges describe their decisions. *Psychology, Public Policy, and Law*, 15, 102–123.
- Connolly, D. A., & Read, J. D. (2003). Remembering historical child sexual abuse. *Criminal Law Quarterly*, 47, 438–480.
- Cordon, I. M., Pipe, M.-E., Sayfan, L., Melinder, A., & Goodman, G. S. (2004). Memory for traumatic experiences in childhood. *Developmental Review*, 24, 101–132.
- Criminal Code, R. S. C. (1985). c. C-46.
- Goodman, G. S., Batterman-Faunce, J. M., & Schaaf, J. M. (2002). Nearly 4 years after an event: Children's eyewitness memory and adults' perceptions of children's accuracy. *Child Abuse & Neglect*, 26, 849–884.
- Goodman, G. S., Bottoms, B. L., Herscovici, B. B., & Shaver, P. R. (1989). Determinants of the child victim's perceived credibility. In S. J. Ceci, D. F. Ross, & M. P. Toglia (Eds.), *Perspectives on children's testimony* (pp. 1–22). New York: Springer-Verlag.
- Hardt, J., & Rutter, M. (2004). Validity of adult retrospective reports of adverse childhood experiences: Review of the evidence. *Journal of Child Psychology and Psychiatry*, 45, 260–273.
- Howe, M. L., Courage, M. L., & Peterson, C. (1994). The emergence and early development of autobiographical memory. *Psychological Review*, 104, 499–523.
- Lewis, P. (2006). *Delayed prosecution for childhood sexual abuse*. New York: Oxford University Press.
- London, K., Bruck, M., Ceci, S. J., & Shuman, D. W. (2005). Disclosure of child sexual abuse: What does the research tell us about the ways that children tell? *Psychology, Public Policy, and Law*, 11, 194–226.
- Lyon, T. D. (2000). Child witnesses and the oath: Empirical evidence. *Southern California Law Review*, 73, 1017–1074.
- Martone, M., Jaudes, P. K., & Cavins, M. K. (1996). Criminal prosecution of child sexual abuse cases. *Child Abuse & Neglect*, 20, 457–464.
- McCarron, A. L., Ridgway, S., & Williams, A. (2004). The truth and lie story: Developing a tool for assessing child witnesses' ability to differentiate between truth and lies. *Child Abuse, Review*, 13, 42–50.
- Meyer, A. (2007). The moral rhetoric of childhood. *Childhood*, 14, 85–104.
- Nolo. (2005). *Statutes of limitations: Is your lawsuit timely?* Retrieved September 9, 2009, from <http://www.nolo.com/legal-encyclopedia/article-29941.html>

- Ornstein, P. A., Ceci, S. J., & Loftus, E. F. (1998). Adults' recollections of childhood abuse: Cognitive and developmental perspectives. *Psychology, Public Policy, and Law*, 4, 1025–1051.
- Peterson, C. (1999). Children's memory for medical emergencies: 2 years later. *Developmental Psychology*, 35, 1493–1506.
- Peterson, C., & Whalen, N. (2001). Five years later: Children's memory for medical emergencies. *Applied Cognitive Psychology*, 15, S7–S24.
- Peterson, C. C., Peterson, J. L., & Seeto, D. (1983). Developmental changes in ideas about lying. *Child Development*, 54, 1529–1535.
- Pezdek, K., Morrow, A., Blandon-Gitlin, I., Goodman, G. S., Quas, J. A., Saywitz, K. J., . . . & Brodie, L. (2004). Deception detection in children: Event familiarity affects criterion-based content analysis ratings. *Journal of Applied Psychology*, 89, 119–126.
- R. v. Audet. (1999). OJ No. 5306. (Sup. Ct.)
- R. v. B. (G.) (1990). 2 S. C. R. 30.
- R. v. D. D. (2000). S. C. R. 275.
- R. v. Gagnon. (2006). SCJ No. 17.
- R. v. J. L. S. A. (2001). MJ No. 439.
- R. v. Joudrie. (1997). OJ. No. 1619.
- R. v. L. M. (2008). SCJ No. 31.
- R. v. Marquard. (1993). 4 S. C. R. 223.
- R. v. P. G. (2002). O. J. No. 2208.
- R. v. Pinnock. (2004). O. J. No. 3692.
- R. v. R. E. M. (2001). B. C. J. No. 2492.
- R. v. Robichaud. (2007). AJ No. 1169.
- R. v. Ross. (2006). N. S. J. No. 233.
- R. v. Stea. (2007). O. J., No. 2219.
- R. v. Stewart. (1994). O. J. No. 811.
- R. v. W. (R.) (1992). 2 S. C. R. 122.
- R. v. Westcott. (1999). O. J. No. 916.
- Read, J. D., & Connolly, D. A. (2007). The effects of delay on long-term memory for witnessed events. In M. P. Toglia, J. D. Read, D. F. Ross, & R. C. L. Lindsay (Eds.), *Handbook of eyewitness psychology: Vol. 1: Memory for events* (pp. 117–155). Mahwah, NJ: Erlbaum.
- Ross, D. F., Dunning, D., Toglia, M. P., & Ceci, S. J. (1990). The child in the eyes of the jury. *Law and Human Behavior*, 14, 5–23.
- Ross, D. F., Jurden, F. H., Lindsay, R. C. L., & Keeney, J. M. (2003). Replication and limitations of a two-factor model of child witness credibility. *Journal of Applied Social Psychology*, 33, 418–431.
- Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review*, 103, 734–760.
- Saywitz, K. J., Nathanson, R., & Snyder, L. S. (1993). Credibility of child witnesses: The role of communicative competence. *Topics in Language Disorders*, 13, 59–78.
- Saywitz, K. J., Snyder, L., & Nathanson, R. (1999). Facilitating the communicative competence of the child witness. *Applied Developmental Science*, 3, 58–68.
- Tubb, V. A., Wood, J. M., & Hosch, H. M. (1999). Effects of suggestive interviewing and indirect evidence on child credibility in a sexual abuse case. *Journal of Applied Social Psychology*, 29, 1111–1127.
- Van Tongeren Harvey, W., & Dauns, P. (2001). *Sexual offences against children and the criminal process*. Toronto, Ontario, Canada: Butterworths.
- Wimmer, H., Gruber, S., & Perner, J. (1984). Young children's conception of lying: Lexical realism–moral subjectivism. *Journal of Experimental Child Psychology*, 37, 1–30.

- Wixted, J. T. (2004). On common ground: Law of retrograde amnesia. *Psychological Review*, *111*, 864–879.
- World Health Organization. (2006). *Preventing child maltreatment: A guide to taking action and generating evidence*. Geneva, Switzerland: Author. Retrieved September 20, 2009, from http://whqlibdoc.who.int/publications/2006/9241594365_eng.pdf

Appendix

Coding Categories and Descriptions

Category	Description	Example
Memory		
Specific details	At least one detail from a specific or single incident of alleged abuse. Multiple incidents were considered independent when the complainant was physical free from the accused between episodes.	“One time the accused unbuttoned his pants”
General details	General information concerning how and when the alleged abuse normally occurred. Often involved past tense such as “he would” and “it was always.”	“He would try to kiss me”
Memory failure	Complainant either admitted memory failure of a detail(s) or the judge discussed memory failures or alluded to impoverished memory.	“She was missing memories”
Reliability		
Inconsistencies	Details reported differently on different occasions, omitted or included details across interviews and/or at trial, and differences between complainant’s testimony and other witnesses regarding specific details.	“Said it happened every time she wore a skirt, but it only happened once”
Corroboration	Evidence that could be used to support the allegation.	“Witness saw the girls on the accused’s lap”
Quality of the evidence	General comments about characteristics of the evidence.	“No complainant gave straightforward evidence”

(Appendix continues)

Appendix (*continued*)

Category	Description	Example
Credibility		
Conduct after assault	Complainant's emotional and/or behavioral state attributable to the alleged abuse or that was considered relevant when considering the alleged abuse. Reported as occurring during the alleged abuse, shortly after it ended, or long after the alleged abuse ended.	"Visited the accused and told the accused's mother that she was going to miss them"
Resistance to assault	Physical, verbal, or emotional resistance or non-resistance during or after the alleged abuse.	"Did not attempt to resist because she did not care anymore"
Reasons for delay	Complainant's reason for not disclosing the abuse immediately.	"He told her that what he was doing was not wrong"
Possibility of collusion	Judge discussed the issue of collusion, whether allowing for its possibility or stating that s/he believed none was present.	"The complainants' accounts have come to merge"
Motive to fabricate	Reasons for or against the possibility of fabrication.	"Not sophisticated enough to fabricate a story"
Conduct after disclosure	Complainant's behavior or emotions displayed to an authority figure around the time of disclosure. An authority figure was defined as one who would be required by law to report suspected child abuse.	"She felt more comfortable in 1999 talking to the Police"
Reason for disclosure	Complainant's reason(s) for disclosing the alleged abuse.	"She saw the accused picking up her 4-year-old boy"
Demeanor	Complainant's attitudes, emotions, and behaviors when testifying.	"Mode of expression was awkward"
Personality traits	Enduring constructs of the complainant such as intelligence, memory, propensity to lie, or details of education and/or social class.	"Does not see herself as worthwhile"
Inferences		
Someone would have known	Statements or evidence relating to the presence of, or knowledge of presence of, other people during the abuse.	"Would not have been able to see from where he was"

Appendix (*continued*)

Category	Description	Example
Other inferences	Inferences made by the judge for which external support is not provided.	“Every memory fades with time”

Received June 3, 2009
Revision received January 12, 2010
Accepted January 13, 2010 ■

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at <http://notify.apa.org/> and you will be notified by e-mail when issues of interest to you become available!