

## Perceptions and Predictors of Children's Credibility of a Unique Event and an Instance of a Repeated Event

Deborah A. Connolly · Heather L. Price ·  
Jennifer A. A. Lavoie · Heidi M. Gordon

Published online: 26 January 2007

© American Psychology-Law Society/Division 41 of the American Psychological Association 2007

**Abstract** Perceptions of children's credibility were studied in two experiments wherein participants watched a videotape of a 4- to 5- or a 6- to 7-year old child report details of a play session that had been experienced once (single-event) or was the last in a series of four similar play sessions (repeat-event). The child's report was classified as high or low accurate. In Experiments 1 and 2, reports of repeat-event children were judged to be less believable on several measures. In Experiment 1, younger children were viewed as less credible than older children. In both experiments, neither undergraduates nor community members correctly discriminated between high- and low-accurate reports. Content analysis in Study 3 revealed the relationship between age and event frequency and children's credibility ratings was mediated by the internal consistency of children's reports. Recent research on children's reports of instances of repeated events has identified several challenges facing children who report repeated abuse. These data bring to light another potential difficulty for these children.

**Keywords** Children · Law · Perceived credibility · Repeat events

In child sexual abuse (CSA) cases, there is rarely any corroborating evidence (Bruck, Ceci, & Hembrooke, 2001; Lamb, 1999; Pezdek et al., 2004) and so decisions about children's credibility made by child protection workers, police, prosecutors, and triers of fact are often determinative of the investigative and/or adjudicative outcome (Goodman, Batterman-Faunce, & Schaaf, 2002; McCarron, Ridgway, & Williams, 2004; Pezdek et al., 2004). Indeed, Lamb (1999) found that research concerning child witnesses is expanding to include communicative competence. And, Castelli, Goodman, and Ghetti (2005) reasoned that perceived credibility may be at least as important as actual accuracy. After all, it is the child's ability to communicate, combined with the ability of those working in the forensic system to understand that communication, which is often determinative of outcome. An important factor to consider when studying perceived credibility is the nature of the event that children may be asked to describe. Often, when children

---

D. A. Connolly (✉) · H. L. Price · J. A. A. Lavoie · H. M. Gordon  
Department of Psychology, Simon Fraser University, 8888 University Dr.,  
Burnaby, British Columbia, Canada V5A 1S6  
e-mail: debc@sfu.ca

 Springer

enter the criminal justice system, they do so as alleged victims of repeated abuse (e.g., CSA). In law, the specificity principle (e.g., *R. v. B. (G.)*, 1990) states an indictment [criminal charge] must describe the offence so as to “lift it from the general to the particular” (p. 209). This principle may require children (and other witnesses alleging a repeated offence) to report details of instances of the alleged offence. Thus, an important question is how children report instances of repeated events and how credible the reports are judged to be.

### Perceptions of credibility

Children’s credibility is often conceptualized along two dimensions: honesty and cognitive competence (Bottoms, 1993; Goodman, Bottoms, Herscovici, & Shaver, 1989; Ross, Dunning, Toglia, & Ceci, 1990; Ross, Jurden, Lindsay, & Keeney, 2003). The influence of these two dimensions on perceptions of credibility may vary with children’s age. That is, generally, compared to older children and adults, younger children are seen to be more honest but less cognitively competent. Accordingly, in cases where honesty is relatively more salient (e.g., CSA perpetrated by a familiar person), younger children are viewed as more credible than older children and adults, whereas in cases where cognitive competence is relatively more salient (e.g., witness to a motor vehicle accident), older children and adults are viewed as more credible than younger children (e.g., Bottoms, 1993; Goodman et al., 1989).

Report consistency is also influential in determining credibility. Brewer, Potter, Fisher, Bond, and Luszcz (1999) found that, among 10 possible indicators of inaccurate testimony, undergraduate students rated inconsistencies with previous statements as the strongest indicator of inaccuracy (see also Conte, Sorenson, Fogarty, & Rosa, 1991). Similarly, Berman, Narby, and Cutler (1995) and Berman and Cutler (1996) found that, compared to a consistent witness, an inconsistent witness was judged to be less credible and such a case was less likely to end with a conviction (but see Leippe & Romanczyk, 1989, Exp. 3 and 4 for a contrary result). Myers, Redlich, Goodman, Prizmich, and Imwinkelreid (1999) reported that actual jurors rated the child’s honesty and consistency as the most important factors in deciding the outcome of a CSA trial they had adjudicated. Other predictors of credibility include confidence, powerful speech, and willingness to admit memory failures (Leippe, Manion, & Romanczyk, 1992) as well as the suggestive nature of the interview used to elicit children’s reports (Tubb, Wood, & Hosch, 1999).

### Perceptions of credibility and event frequency

Most research on children’s credibility has focused on reports of a unique event. There are, though, a few important exceptions. Using Criterion Based Content Analysis (CBCA), Pezdek et al. (2004) rated 3- to 13-year-old children’s reports of an invasive medical procedure (VCUG) that had occurred once (unfamiliar) or more than once (familiar). CBCA scores were higher for reports of familiar than unfamiliar events. In Blandon-Gitlin, Pezdek, Rogers, and Brodie (2005), some children received instruction on how to sew a button on a shirt (familiar), while other children did not receive such training (unfamiliar). One day later, all children engaged in a three-activity session during which half of the children in each familiarity condition sewed a button on a shirt. When later describing the three-activity session, children in the familiar condition received higher CBCA scores than children in the unfamiliar condition. In Strömwall, Bengtsson, Leander, and Granhag (2004), 10- to 13-year-olds experienced a mock medical examination once or four times. One week after the final occasion, children were interviewed about the event and their reports were evaluated using CBCA and reality monitoring (RM)

criteria. Children who engaged in the event repeatedly scored higher than single-event children on several of the CBCA and RM criteria.

A potentially important qualification to these studies has to do with the nature of the target event. In each of the above-cited studies, the events experienced by the repeat-event children were likely highly similar. In the VCUG study, children experienced a highly invasive and painful medical examination. It is probable that the procedure is highly structured and so each experience is very similar to the others. In Blandon-Gitlin et al. (2005), familiarity was manipulated by providing, or not providing, children with training on how to perform the target task. The training focused on details common to the activity of sewing on a button (K. Pezdek, personal communication, July 18, 2006). In Strömwall et al. (2004), all events were identical (with minor variations, such as what the “doctor” and the child talked about). In summary, in the above-cited studies, children in the familiar conditions reported an instance of a repeated event that had been experienced in much the same way during each encounter.

When an event recurs, some event details remain the same (i.e., fixed details) and some details change across instances (i.e., variable details). Importantly, the content and presentation of fixed and variable details differ. Fixed details are described with greater accuracy and more confidence (Connolly & Lindsay, 2001; Roberts & Powell, 2005). Indeed, Roberts and Powell (2005) found that 5- to 6-year-old children reported higher confidence in their reports of fixed, compared to variable details. It is also expected that children describe fixed details with more specificity than variable details (e.g., report the particular “we made a puzzle” rather than the general “we played”) because the particular detail was presented repeatedly and there is no need to move to a higher level of abstraction. Logically, compared to reports of fixed details, reports of variable details will also contain more optionals (e.g., you might do this, or this) and conditionals (e.g., if . . . then). Given these differences in content and presentation style, reports of instances of a repeated event that contain primarily fixed details may be judged differently than reports of instances of repeated events that contain many variable details.

Reports of instances of repeated events that contain more variable details may be judged to be less credible than memory reports of unique events because the latter may contain the very characteristics, as described above, that tend to lower perceptions of credibility. Compared to reports of unique events, reports of instances of repeated events that contain variable details are less consistent and less distinctly associated with particular instances (Farrar & Goodman, 1990; Fivush, 1984, 1997; Fivush, Hudson, & Nelson, 1984; Hudson, 1990; Hudson & Nelson, 1986; Powell & Roberts, 2002; Powell & Thomson, 1996, 2003). They may also be less specifically situated in time because children’s reports of instances of repeated events tend to contain relatively more present tense, impersonal pronouns (“you”), and temporal markers (e.g., Fivush, 1984, 1997; Hudson & Nelson, 1986). In certain circumstances, children who experience several similar instances of an event are also more suggestible than children who experience the same event once (Connolly & Lindsay, 2001; Connolly & Price, 2006; Powell & Roberts, 2002; Price & Connolly, 2004). Based on these data, we expect that children’s reports of a unique event will be judged to be more credible than reports of an instance of a repeated event when the repeated event contains many details that vary across instances.

#### Perceptions of credibility and accuracy

In many studies of children’s credibility, the child’s accuracy is not in issue because actual accuracy is unknown. Among studies where actual accuracy of the child’s report is known, the results are somewhat mixed. A few studies have found participants have limited ability to discriminate accurate from inaccurate reports. For example, Tye, Amato, Honts, Devitt, and

Peters (1999) found discriminability to be just slightly better than chance when undergraduates evaluated the credibility of a lie that a child told to protect a parent from possible punishment. In other studies, actual accuracy had no effect on ratings of credibility. For instance, Leippe et al. (1992, Exp. 2) found that actual accuracy had no effect on perceived credibility ratings of 5- to 6- or 9- to 10-year-olds by undergraduates. Similarly, Goodman et al. (2002) found that social workers and undergraduates were unable to discriminate true from false reports of children who reported details of an interaction that had occurred four years earlier. They found that participants were unable to estimate accuracy and “at times the respondents’ estimates were even in the wrong direction” (p. 865). Similarly, Ceci, Crotteau-Huffman, Smith, and Loftus (1994) found that professionals in psychology, law enforcement, social work, and psychiatry performed at chance when discriminating between reports of events children had experienced from reports of entire events children had not experienced (see also Ceci, Loftus, Leichtman, & Bruck, 1994; Talwar, Lee, Bala, & Lindsay, 2006). In other studies, actual accuracy has predicted ratings of credibility, but in the wrong direction. Leichtman and Ceci’s (1995) now famous “Sam Stone” study is an excellent illustration. Three reports of Sam Stone’s visit to their day care (one highly accurate, one highly inaccurate, and one intermediately accurate) were shown to psychology and legal professionals who were asked to rate the accuracy of the reports. The judges were highly inaccurate: The child who was inaccurate was judged to be highly credible whereas the child who was accurate was judged to be less credible (see also Ceci, Crotteau-Huffman et al., 1994). Goodman et al. (1989) reported a similar result with undergraduates and Talwar and Lee (2002) reported the same results when undergraduates attempted to discriminate lie-tellers from truth-tellers.

With only a few exceptions, the extant research supports the conclusion that undergraduates, community members, and professionals are no better than chance (and sometimes worse) at discriminating children’s accurate from inaccurate reports. Importantly, all of these studies involved children reporting an event that ostensibly occurred once. This result may not generalize to discriminability of children’s reports of an instance of a repeated event. As discussed above, both the content and the presentation of repeat-event children’s reports are different on many characteristics that have been shown to influence ratings of credibility. We are unaware of any literature that could be used to predict the influence of actual accuracy on ratings of credibility when the event being described is an instance of a repeated event.

### The present study

In the present study, each participant viewed one of 31 videotapes, each with a different child describing details of the same past play session. The child was either 4- to 5- or 6- to 7-years-old, classified as high- or low-accurate, and reported an event they had experienced once (unique event) or was the last in a sequence of four similar events. For children who experienced the event four times, 16 critical details changed across experiences and children’s attention was drawn to the details that changed. Participants then completed a questionnaire that measured their perceptions of the child’s accuracy, honesty, overall credibility, consistency, confidence, likeability, and suggestibility. Because judgments were based on memory for particular details of an unfamiliar event, we hypothesized that cognitive competence would be relatively more salient than honesty and so younger children would be seen as less credible than older children. For repeat-event children, the event sequence contained many variable details. As such, we hypothesized that repeat-event children would be viewed as less credible than single-event children. We expected to replicate the extant research findings that participants would be insensitive to actual accuracy in ratings of single-event children. However, given the paucity of comparable research on reports

of instances of repeated events, we had no hypothesis concerning the effect of actual accuracy on credibility ratings for repeat-event children.

## Experiment 1

### Method

#### *Participants*

Participants were 127 undergraduate students (55 males,  $M_{age} = 20.74$  years,  $SD = 2.08$  years). Most students ( $n = 83$ ) reported English as their first language and the remaining 44 students had spoken English for an average of 11.78 years ( $SD = 3.82$  years).

#### *Materials*

*Videotapes.* Sixty-six parents from an earlier study (Connolly & Price, 2006) granted permission for the use of the videotape of their child in the present study. Thirty-one videotapes were selected to include four videotapes in each cell of a  $2 \times 2 \times 2$  matrix. Children were 4- to 5-years-old or 6- to 7-years-old; they described, in free and cued recall, an event that had occurred once or was the last instance in a series of four similar, but not identical, events; and within each event frequency condition children were either high or low on a measure of accuracy. In studies of credibility that include videotaped reports, individual witness characteristics that are unrelated to the independent variables can influence credibility ratings. Luus and Wells (1992) and Leippe et al. (1992) recommended that multiple children be included to minimize these effects. Leippe et al. (1992) initially included three videotapes in each condition. In the current study, there were four videotapes in all conditions except the young child, single-event, high-accuracy condition, wherein there were three videotapes. Inadvertently, one of the videotapes was replaced with a different child and all ratings of that child had to be dropped from the study.

The target event that children described was a play session. The play sessions involved eight activities: play a pretend game, construct a puzzle, colour a sticker, make up a story, draw a picture, search for a hidden treasure, make a play-dough model, and find a prize underneath one of three cups. There were two critical details in each activity for a total of 16 critical details.

Half of the children engaged in the play session once (single-event condition) and half engaged in four similar but not identical play sessions (repeat-event condition). In the repeat-event condition, the 16 critical details changed across experiences (e.g., children constructed a puzzle of a cow, tiger, pig, and mouse during play sessions 1, 2, 3, and 4, respectively). During the only play session for single-event children and the last play session for repeat-event children, the play-session leader wore a special cape to tag it as “Cape Day.” In subsequent interviews, children were asked to think only about “Cape Day” when answering questions. Two weeks after the final, or only, play session, all children participated in a biasing interview about “Cape Day” in which eight experienced details were misrepresented. The day after the biasing interview, each child completed a final memory test in which rapport was first established (i.e., the interviewer discussed family, school, or other interests with the child for a few minutes before beginning the interview). Then, free recall began with an open request for the child to report all that he or she could remember about “Cape Day.” When the child stopped talking for approximately 10 seconds, he or she was prompted with the names of each activity, presented one at a time. When the child appeared to have exhausted his or her free recall of the event, the cued recall test was administered. Children were asked about each critical detail, in the order the details were

presented (e.g., “On Cape Day, you put together a puzzle, what was the picture on the puzzle?”). All cued recall questions were asked, regardless of whether the child reported the information in free recall.

Accuracy was assigned based on information provided about the 16 critical details in free and cued recall: Those were the only details that could be verified as the play sessions were not recorded. Correct responses were details that had been presented during *any* of the play sessions or activities. Incorrect responses were suggested details and details that were neither suggested nor experienced during any of the play sessions. Off-topic comments and “don’t know” responses were not used to compute proportion correct. In free recall, children described critical and non-critical details whereas in cued recall, children described critical details only. About 60% of the interview time involved free recall ( $M = 7.34$  min,  $SD = 2.48$  min) and included descriptions of an average of 6.74 ( $SD = 2.82$ ) critical details. Approximately 40% of the interview time involved children’s responses to cued recall questions. If 40% of the free recall time involved children describing critical details, then approximately 65% of the total interview focused on descriptions of critical details. We expect that there was a positive, but not perfect, correlation between accuracy on critical details and accuracy on non-critical details. Therefore, we reason that accuracy on critical details is a reasonable approximation of overall accuracy.

A proportion correct score was computed (total correct responses/total correct + total incorrect responses). For older and younger children separately, and within each frequency condition, children with the highest and lowest proportion correct scores and whose videotapes were clear enough for use in this study were selected (e.g., in one videotape the child continuously walked out of the range of the camera and could not be seen for considerable periods of time, another child said “I don’t know” to most questions). Videotapes averaged 11.81 min ( $SD = 3.17$  min). A 2 (age: 4- to 5-years, 6- to 7-years)  $\times$  2 (event frequency: 1, 4 sessions)  $\times$  2 (accuracy: high, low) ANOVA was used to investigate differences in length of video across conditions. There were no differences ( $F_s < 1$ ).

*Credibility questionnaire.* Undergraduate participants rated, on 6-point scales, the child’s honesty (questions about honesty, truthfulness, and likelihood of fabrication), cognitive competence (questions about intelligence, how well the child understood the event, and two questions about the child’s accuracy), credibility (credibility and believability), consistency,<sup>1</sup> confidence, likeability, and suggestibility. Questions used to assess honesty and cognitive competence were taken from Ross et al. (2003). Although Ross et al. (2003) included suggestibility and consistency in their measure of cognitive competence, as reported below, others have found these two report characteristics may be predictors of credibility rather than inherent characteristics of credibility. Variables that have been shown to influence ratings of credibility were measured, including consistency (Brewer et al., 1999; Conte et al., 1991; Berman & Cutler, 1996), confidence (Cutler, Penrod, & Stuve, 1988; Lindsay, Wells, & O’Connor, 1989; Luus & Wells, 1994), likeability (Leippe et al., 1992), and suggestibility (Castelli et al., 2005 and Tubb et al., 1999 found that children interviewed in a suggestive way were viewed as less credible than children interviewed in a neutral way). Two random orders of credibility questions were created and half of the participants received each version. After answering the credibility questions, participants provided demographic information. Assignment of videotapes to participants was random with the constraint that each videotape be viewed by an equal number of participants.

<sup>1</sup>An anonymous reviewer correctly stated that because the participant listened to only one report, “coherence” is a more appropriate description of this report characteristic. However, given that it was labeled “consistency” in the questionnaire, we retain that term throughout the paper.

### Design and procedure

This was a  $2 \times 2 \times 2$  between-subjects design. The independent variables were age of the child (4- to 5-, 6- to 7-years-old), event frequency (1, 4 times) and actual accuracy (high, low). Participants were told that they would view a 15–20 min videotape of a child reporting details of a past event. They were further instructed that the videotape was randomly selected from a group of videotapes and that, within this group, some of the children were very accurate in their description of the event and some were not very accurate. The accuracy instruction was included to encourage careful attention to the videotape as participants might be motivated to correctly identify the child as high- or low-accurate. Participants were not told the child's age or whether the child was reporting a unique event or an instance of a repeated event. Finally, participants were informed that they would view the videotape once and then would be asked a series of questions about the credibility of the child. After affirming that the child depicted in the still photograph was not known to them *in any capacity*, participants viewed the videotape on a computer monitor, completed the questionnaire, and were debriefed. Participants received course credit or \$5.00 for participation.

### Results

#### Analyses of actual accuracy

Means and observed values of proportion correct are in Table 1. Because there were only four observations in each cell (three in one cell), we did not run an ANOVA on the full model. T-tests were used to compare proportion correct as a function of the independent variables. Proportion correct did not vary as a function of age ( $M = .80$ ,  $SD = .17$  and  $M = .75$ ,  $SD = .17$ , for younger and older children, respectively),  $t(29) = .80$ ,  $p = .43$ , or event frequency ( $M = .82$ ,  $SD = .13$  and  $M = .74$ ,  $SD = .20$ , for single- and repeat-event children, respectively),  $t(29) = 1.36$ ,  $p = .19$ . Children assigned to the high-accurate group were indeed more accurate ( $M = .90$ ,  $SD = .06$ ) than children assigned to the low-accurate group ( $M = .66$ ,  $SD = .15$ ),  $t(29) = 5.89$ ,  $p < .001$ .

To control for the effects of non-independence of error attributable to the fact that four participants viewed and rated each videotape, a fully nested model with videotape nested within subjects was used to analyze the data. Results are reported in two sections. First, we report ratings of credibility including cognitive competence, honesty, and overall credibility. Second we report ratings of report characteristics that could be used to explain credibility ratings: consistency, confidence, likeability, and suggestibility. All effects are significant at  $p < .05$  unless otherwise stated. The means for all effects reported in this experiment are in Table 2.

#### Credibility ratings

Event frequency affected each measure of credibility in the predicted direction. Single-event children were judged to be more cognitively competent [ $F(1, 23) = 8.58$ ,  $\eta^2 = .27$ ], more honest [ $F(1,$

**Table 1** Proportion correct (observed values in parentheses) as a function of event frequency, accuracy classification, and age

	Single-event		Repeat-event	
	Low accuracy	High accuracy	Low accuracy	High accuracy
4 to 5 years	.73 (.71, .72, .73, .76)	.93 (.86, .94, 1.00)	.65 (.29, .72, .73, .85)	.93 (.90, .91, .94, .95)
6 to 7 years	.71 (.60, .60, .81, .84)	.92 (.86, .94, .94, .95)	.53 (.43, .50, .57, .63)	.84 (.80, .83, .85, .87)

**Table 2** Mean ratings (*SD* in parentheses) of cognitive competence, honesty, credibility, consistency, confidence, likeability, and suggestibility as a function of age, event frequency, and accuracy in Experiment 1

	Age		Event frequency		Accuracy	
	4–5	6–7	Single	Repeat	Low	High
Cog. comp.	3.55 (0.39)	3.97 (0.43)	3.98 (0.36)	3.56 (0.45)	3.76 (0.50)	3.77 (0.42)
Honesty	4.11 (0.69)	4.36 (0.68)	4.61 (0.55)	3.90 (0.59)	4.18 (0.69)	4.31 (0.66)
Credibility	3.55 (0.65)	4.10 (0.68)	4.19 (.52)	3.50 (0.72)	3.70 (0.77)	3.97 (0.64)
Consistency	3.73 (0.95)	4.16 (0.78)	4.38 (0.69)	3.56 (0.87)	3.98 (0.92)	3.93 (0.87)
Confidence	3.27 (0.71)	3.84 (0.89)	3.94 (0.77)	3.21 (0.77)	3.63 (0.72)	3.49 (0.99)
Likeability	4.38 (0.81)	4.90 (0.45)	4.98 (0.42)	4.33 (0.75)	4.62 (0.75)	4.68 (0.63)
Suggestibility	4.07 (0.69)	3.88 (0.68)	3.70 (0.72)	4.22 (0.55)	3.98 (0.74)	3.97 (0.64)

*Note.* Ratings were made on 6-point scales (1 = low, 6 = high). Except for suggestibility, a high score reflects superior ratings.

23) = 11.65,  $\eta^2 = .34$ ], and more credible [ $F(1, 23) = 11.86$ ,  $\eta^2 = .34$ ] than repeat-event children. Older children were judged to be more cognitively competent [ $F(1, 23) = 9.52$ ,  $\eta^2 = .29$ ] and more credible [ $F(1, 23) = 6.39$ ,  $\eta^2 = .22$ ] than younger children.

### Report characteristics

Event frequency affected all measures of report characteristics in the predicted direction. Single-event children were judged to be more consistent [ $F(1, 23) = 7.50$ ,  $\eta^2 = .25$ ], more confident [ $F(1, 23) = 6.24$ ,  $\eta^2 = .21$ ], more likeable [ $F(1, 23) = 9.05$ ,  $\eta^2 = .28$ ], and less suggestible [ $F(1, 23) = 4.70$ ,  $\eta^2 = .17$ ] than repeat-event children. Older children were judged to be more confident [ $F(1, 23) = 4.62$ ,  $\eta^2 = .17$ ] and more likeable [ $F(1, 23) = 5.41$ ,  $\eta^2 = .19$ ], than younger children.

### Discussion

Event frequency affected every rating in the predicted direction. Repeat-event children were judged to be less cognitively competent, less honest and less credible than single-event children. Given that repeat-event children were also rated as less consistent, less confident, less likeable, and more suggestible, we speculate that these characteristics may help to explain the effects of event frequency on credibility. We discuss this more fully in the General Discussion, after we report a replication of this study and an analysis of the content of children's reports.

Predictably, older children were judged to be more cognitively competent, credible, and confident than younger children. Credibility has been defined along two dimensions, honesty and cognitive competence (Bottoms, 1993; Ross et al., 2003). Generally, older children are judged to be more cognitively competent whereas younger children are judged to be more honest (Bottoms, 1993). In the present study, the children experienced either one or four unfamiliar play sessions that contained particular details they were asked to report in free and cued recall. This emphasis on details may have led participants to place more emphasis on cognitive competence than honesty when estimating credibility. We were a bit surprised that older children were seen as more likeable than younger children. Perhaps, because we directed participant's attention to the child's credibility, we also primed them to favor more credible children.

Participants did not distinguish between children classified as high- and low-accurate. Power was calculated at .80 if the effect size is at least medium and there are 16 observations in each cell. In the current study, there were 31 videotapes in total, sufficient to detect a medium effect size, if an effect was present. That our participants were insensitive to actual accuracy is consistent



with a substantial body of research that finds discrimination between children's accurate and inaccurate reports of a unique event to be poor (Ceci, Crotteau-Huffman et al., 1994; Ceci, Loftus et al., 1994; Goodman et al., 2002; Leippe et al., 1992). We extend that finding to include an inability to discriminate accurate from inaccurate reports of instances of repeated events.

To the best of our knowledge, this is the first demonstration of a deleterious effect of event frequency on ratings of credibility and report characteristics that are associated with credibility. To further generalize the effects observed in Experiment 1, community members were recruited to participate in Experiment 2.

## Experiment 2

### Method

#### *Participants*

Participants were 186 community members recruited from kiosks at two shopping malls (91 males,  $M = 36.45$  years,  $SD = 14.09$  years). Participation was restricted to jury-eligible persons: over age 18 and a citizen of the country. Most participants (84.4%) reported that English was their first language. Of the remaining participants, the average number of years they had spoken English was 18.57 ( $SD = 9.31$  years). Participants were randomly assigned to view one videotape in one of the cells of a  $2 \times 2 \times 2$  matrix described in Experiment 1. There were 24 observations in each cell with one exception. There were 18 observations in the cell labeled young child, single-event, and high-accurate.

#### *Materials, design, and procedure*

The materials and design were identical to that used in Experiment 1 except that participants viewed the videotape from a laptop computer with a headset used for the audio.

### Results

Data were analyzed as in Experiment 1. Given the robust effects of event frequency reported in Experiment 1, the main effects of frequency were run as one-tailed tests. All other main effects and all interactions were run as two-tailed tests. The means for main effects of age, event frequency, and accuracy are in Table 3.

#### *Credibility ratings*

Children in the single-event condition were judged to be marginally more cognitively competent [ $F(1, 24) = 2.65, p = .06, \eta^2 = .10$ ] and marginally more credible [ $F(1, 24) = 2.93, p = .05, \eta^2 = .11$ ] than children in the repeat-event condition. There was an interaction between age and frequency [ $F(1, 24) = 4.48, \eta^2 = .16$ ] on honesty ratings. Among younger children, those reporting a single event were rated to be as honest ( $M = 3.90, SD = 0.36$ ) as those reporting an instance of a repeated event ( $M = 4.08, SD = 0.34$ ),  $F = 1$ . Among older children, those reporting a unique event were rated as marginally more honest ( $M = 4.17, SD = 0.24$ ) than those reporting an instance of a repeated event ( $M = 3.89, SD = 0.23$ ),  $F(1, 14) = 5.91, p = .07, \eta^2 = .30$ .

**Table 3** Mean ratings (*SD* in parentheses) of cognitive competence, honesty, credibility, consistency, confidence, likeability, and suggestibility as a function of age, event frequency, and accuracy in Experiment 2

	Age		Event frequency		Accuracy	
	4–5	6–7	Single	Repeat	Low	High
Cog. comp.	3.92 (0.50)	4.17 (0.41)	4.18 (0.47)	3.93 (0.45)	4.27 (0.44)	3.79 (0.36)
Honesty	4.00 (0.35)	4.03 (0.27)	4.04 (0.32)	3.99 (0.30)	4.09 (0.30)	3.93 (0.30)
Credibility	4.14 (0.57)	4.42 (0.47)	4.46 (0.55)	4.12 (0.48)	4.39 (0.50)	4.15 (0.56)
Consistency	4.06 (0.60)	4.27 (0.52)	4.35 (0.64)	4.01 (0.45)	4.24 (0.57)	4.09 (0.57)
Confidence	3.66 (0.75)	4.12 (0.66)	4.08 (0.78)	3.72 (0.67)	4.14 (0.58)	3.61 (0.80)
Likeability	4.99 (0.47)	4.85 (0.40)	5.00 (0.44)	4.85 (0.43)	5.02 (0.40)	4.81 (0.46)
Suggestibility	3.90 (0.63)	3.49 (0.62)	3.43 (0.49)	3.93 (0.69)	3.67 (0.67)	3.73 (0.64)

*Note.* Ratings were made on 6-point scales (1 = low, 6 = high). Except for suggestibility, a high score reflects superior ratings.

There was a main effect of accuracy classification on ratings of cognitive competence. Children classified as low-accurate were judged to be more cognitively competent than children classified as high-accurate,  $F(1, 24) = 13.27$ ,  $\eta^2 = .36$ .

#### Report characteristics

Children in the single-event condition were judged to be marginally more consistent [ $F(1, 24) = 2.41$ ,  $p = .07$ ,  $\eta^2 = .09$ ], and less suggestible [ $F(1, 24) = 5.64$ ,  $\eta^2 = .19$ ] than repeat-event children. There was an interaction between frequency and accuracy on confidence that approached significance,  $F(1, 24) = 4.08$ ,  $p = .05$ ,  $\eta^2 = .14$ . Among children classified as low-accurate, single-event children were rated as more confident ( $M = 4.55$ ,  $SD = .34$ ) than repeat-event children ( $M = 3.78$ ,  $SD = .51$ ), [ $F(1, 15) = 13.13$ ,  $\eta^2 = .47$ ]. In the analysis of children classified as high-accurate, event frequency did not influence ratings of confidence ( $M = 3.55$ ,  $SD = .82$  and  $M = 3.66$ ,  $SD = .85$ , for single and repeat-event children, respectively;  $F < 1$ ). The interaction between frequency and accuracy also approached significance,  $F(1, 24) = 3.43$ ,  $p = .08$ ,  $\eta^2 = .12$ . There was also a reliable 3-way interaction between age, event frequency, and accuracy. In the analysis of older children, there were no reliable effects. In the analysis of younger children, among those classified as low-accurate, single-event children were rated as more likeable ( $M = 5.46$ ,  $SD = .25$ ) than repeat-event children ( $M = 4.73$ ,  $SD = .37$ ), [ $F(1, 5) = 11.33$ ,  $\eta^2 = .37$ ]. There were no differences in ratings of likeability as a function of event frequency among younger children classified as high-accurate ( $M = 4.74$ ,  $SD = .40$  and  $M = 4.88$ ,  $SD = .53$ , for single and repeat-event children, respectively;  $F < 1$ ).

Older children were judged to be more confident than younger children,  $F(1, 24) = 4.78$ ,  $\eta^2 = .17$ . Children classified as high-accurate were judged to be less confident than children classified as low-accurate,  $F(1, 24) = 7.00$ ,  $\eta^2 = .23$ .

#### Discussion

Event frequency influenced ratings of cognitive competence, honesty, credibility, consistency, confidence, likeability, and suggestibility, either as a main effect or in an interaction. Relative to repeat-event children, single-event children were judged to be marginally more cognitively competent, more credible, more consistent, and less suggestible. Among older children, single-event children were judged to be more honest than repeat-event children and among children classified as low-accurate, those who reported a unique event were judged to be more confident

and, for younger children, more likeable than those who reported the same event that was the last in a sequence of four similar events. These data replicate most of the effects of event frequency observed in Experiment 1.

We replicated only one age effect reported in Experiment 1: Younger children were judged to be less confident than older children. Similar to Experiment 1, community participants did not give higher ratings to the reports of children who were classified as accurate relative to those classified as inaccurate. In fact, unlike the undergraduate students, community members rated low-accurate children to be more cognitively competent and more confident than high-accurate children. Thus, in Experiment 2, as in Experiment 1, we were able to extend the findings that adults do not effectively discriminate children's high- and low-accurate reports of unique events to include an inability to correctly discriminate children's high- and low-accurate reports of instances of repeated events.

In the next study we investigated report content and presentation style to study possible sources of the effects observed in Experiments 1 and 2. As discussed in the introduction, there is evidence that children's reports of an instance of a repeated event are different in many ways from children's reports of a unique event and that these report characteristics have been linked to perceptions of credibility. In Study 3, each child's videotape was coded for syntax, powerfulness, and internal consistency. We hypothesized that repeat-event children would evince a syntactic structure that suggests that the child is unsure of the details that occurred during the particular target play session; less powerful speech and fewer internally consistent responses.

### Study 3

#### Method

Verbatim transcripts from the 31 videotapes shown to participants in Experiments 1 and 2 were coded for syntax, powerfulness, and internal consistency to determine whether these verbal behaviors differed as a function of age, event frequency, and accuracy. First, raters parsed children's free recall narratives into words and assigned them to syntax and powerfulness categories. Syntax categories accounted for the way in which children described the event and were derived from literature suggesting differences in the language used by children describing a unique event versus an instance of a repeated event. Specifically, reports of a repeated event typically contain more temporal markers, lists, impersonal pronoun use, present tense, conditional tense, and fewer specific details (syntax categories adapted from Fivush, Haden, & Adam, 1995; Fivush et al., 1984; Hudson & Nelson, 1986). There were four powerfulness categories, hedges, intensifiers, hesitations, and gestures, the presence of which indicates *less* powerful speech (powerfulness categories adapted from Leippe, Romanczyk, & Manion, 1991; O'Barr, 1982). Second, the information children provided about each of the 16 critical details in free recall was compared to their responses to the 16 cued recall questions about the critical details, yielding one response pair per critical detail. These response pairs were placed into an internal consistency category (adapted from Brewer et al., 1999) that measured the consistency with which children provided information about the critical details across repeated questioning. See Table 4 for specific definitions of syntax, powerfulness, and internal consistency coding categories. Interrater reliability was calculated as percent agreement ( $\text{agreements} / (\text{agreements} + \text{disagreements}) \times 100$ ). Raters came to a consensus for items on which they disagreed. Two raters each coded 4 transcripts (roughly 10% of the sample) independently and then one rater coded the remaining transcripts. See Table 4 for the percent agreement values for each coding category.

**Table 4** Definitions and percent agreement for syntax, powerfulness, and internal consistency coding

<i>Syntax</i>	Percent agreement = 85.9%
Temporal markers	References to chronological time, conditional states, causal relations, or optional states ( <i>then, when, because, first</i> )
Particular qualifiers	References to specific people, times, and places
Lists	Child used <i>and/or</i> to list options of experienced details
Pronouns	Child referenced him/herself using the impersonal pronoun “ <i>You</i> ”
Tense	Used of present or conditional tense (“ <i>We play/would play tennis</i> ”)
<i>Powerfulness</i>	Percent agreement = 86.7%
Hedges	Words that reduce force of child’s speech (“ <i>It was sort of fun</i> ”)
Intensifiers	Words that increase force of child’s speech (“ <i>It was very fun</i> ”)
Hesitations	Pause fillers and meaningless particles (“ <i>Well, it was um fun</i> ”)
Gestures	Spoken directions (“ <i>We had fun over there</i> ”)
<i>Internal consistency</i>	Percent agreement = 86.5%
Consistencies	Child provided the same information in free and cued recall
Contradictions	Child’s responses were contradictory (“ <i>We played pretend tennis</i> ” vs. “ <i>We played pretend bowling</i> ”)
Partially inconsistencies	Part of child’s responses overlapped (“ <i>We played pretend tennis and bowling</i> ” vs. “ <i>We played pretend bowling</i> ”)
Increase in detail	Child provided less specific information in free than cued recall (“ <i>We played pretend sports</i> ” vs. “ <i>We played pretend tennis</i> ”)
Decrease in detail	Child provided more specific information in free than in cued recall

## Results

### *Multivariate analyses of variance*

A 2 (age: 4- to 5-years, 6- to 7-years)  $\times$  2 (event frequency: 1, 4)  $\times$  2 (accuracy: high, low) univariate ANOVA was conducted to determine whether there were differences in free recall narrative word count as a function of the independent variables. Older children’s free recall narratives contained significantly more words ( $M = 378.75$ ,  $SD = 252.45$ ) than younger children’s free recall narratives ( $M = 170.87$ ,  $SD = 99.15$ ), thus, percentages were used for subsequent analyses. For syntax and powerfulness categories, percentages were based on free recall narrative total word count with the exception of three categories: impersonal pronoun use (a percentage of first person and impersonal pronoun use), and present and conditional tense use (percentage of all verbs: conjugated into past, present, or conditional tense). Internal consistency category percentages were derived from the total number of response pairs in free and cued recall. Three separate 2 (age: 4- to 5- years, 6- to 7-years)  $\times$  2 (event frequency: 1, 4)  $\times$  2 (accuracy: high, low) between-subjects multivariate analyses of variance (MANOVA) were performed: one on the six syntax categories, one on the four powerfulness categories, and one on the five consistency categories. Refer to Table 5 for all percentage means and standard deviations.

*Syntax.* There was a main effect of age, Wilks’  $\lambda = .31$  [ $F(6, 10) = 3.66$ ,  $\eta^2 = .69$ ]. Follow-up univariate ANOVAs showed that, compared to younger children, older children’s free recall narratives contained a higher percentage of temporal markers [ $F(1, 29) = 7.80$ ,  $\eta^2 = .21$ ], particular qualifiers [ $F(1, 29) = 6.86$ ,  $\eta^2 = .19$ ], and conditional tense [ $F(1, 29) = 4.34$ ,  $\eta^2 = .13$ ]. The overall MANOVAs for event frequency [Wilks’  $\lambda = .43$ ,  $F(6, 10) = 2.20$ ,  $p = .13$ ,  $\eta^2 = .57$ ] and accuracy [Wilks’  $\lambda = .53$ ,  $F(6, 10) = 1.47$ ,  $p = .28$ ,  $\eta^2 = .47$ ] were non-significant.

**Table 5** Mean percentage (*SD* in parentheses) of verbal behaviors for age and event frequency

	Age		Event frequency	
	4- to 5-years	6- to 7-years	Single	Repeat
<i>Syntax</i>				
Temporal markers	1.18 (1.27)	3.23 (2.56)	3.09 (2.59)	1.44 (1.60)
Particular qualifiers	.25 (.41)	1.00 (1.03)	.91 (.93)	.39 (.75)
Lists	1.13 (1.25)	.80 (1.07)	.33 (.31)	1.55 (1.35)
Impersonal pronoun	14.86 (23.77)	12.00 (26.49)	16.31 (27.55)	9.63 (22.55)
Present tense	19.71 (16.72)	11.01 (20.03)	9.89 (10.20)	20.22 (23.43)
Conditional tense	.74 (2.21)	2.55 (2.61)	2.16 (2.78)	1.21 (2.32)
<i>Powerfulness</i>				
Hedges	.24 (.52)	.17 (.34)	.31 (.54)	.10 (.29)
Intensifiers	.50 (.74)	.37 (.47)	.55 (.64)	.33 (.57)
Hesitations	3.90 (3.59)	3.50 (2.91)	2.63 (2.11)	4.69 (3.77)
Gestures	.33 (.71)	.64 (.76)	.37 (.46)	.60 (.94)
<i>Internal consistency</i>				
Consistencies	30.07 (22.45)	40.02 (18.15)	48.16 (17.53)	23.01 (15.46)
Contradictions	6.93 (6.15)	1.37 (3.94)	4.21 (5.07)	3.93 (6.54)
Partial inconsistencies	4.51 (6.44)	7.14 (12.73)	2.97 (5.38)	8.59 (12.69)
Increase in detail	47.32 (23.88)	2.46 (3.96)	37.61 (12.97)	58.12 (20.29)
Decrease in detail	11.17 (10.26)	49.01 (15.87)	7.05 (9.74)	6.32 (8.01)

*Powerfulness.* There was an Age  $\times$  Frequency  $\times$  Accuracy interaction [Wilks'  $\lambda = .56$ ,  $F(4, 20) = 1.47$ ,  $\eta^2 = .47$ ] for hedges. However, this interaction could not be examined inferentially as children in some cells did not exhibit any hedges. Thus, we provide descriptive information only. Among low-accurate younger children, those describing a unique event had a higher percentage of hedges ( $M = 0.15$ ,  $SD = 0.17$ ) than those describing a repeated event ( $M = 0.00$ ,  $SD = 0.00$ ). For low-accurate older children, the percentage of hedges appeared to be the same when describing a unique event ( $M = 0.27$ ,  $SD = 0.42$ ) and a repeated event ( $M = 0.21$ ,  $SD = 0.42$ ). Among high-accurate younger children, those describing a unique event had a higher percentage of hedges ( $M = 1.00$ ,  $SD = 0.86$ ) than those describing a repeated event ( $M = 0.00$ ,  $SD = 0.00$ ). This pattern was reversed for high-accurate older children; narratives of a repeated event contained a higher percentage of hedges ( $M = 0.21$ ,  $SD = 0.41$ ) than narratives of a unique event ( $M = 0.00$ ,  $SD = 0.00$ ).

Univariate tests also revealed a 3-way interaction effect for hesitations [ $F(1, 23) = 9.14$ ,  $\eta^2 = .28$ ]. For low-accurate protocols, only the main effect of age was reliable [ $F(1, 12) = 5.16$ ,  $\eta^2 = .30$ ]; younger low-accurate children included more hesitations ( $M = 6.05$ ,  $SD = 3.34$ ) than older low-accurate children ( $M = 3.02$ ,  $SD = 2.57$ ). For high-accurate protocols, the main effect of age [ $F(1, 11) = 5.24$ ,  $\eta^2 = .32$ ] was qualified by an Age  $\times$  Frequency interaction [ $F(1, 11) = 9.51$ ,  $\eta^2 = .46$ ]. There was no effect of event frequency for younger-accurate children [ $t(5) = 1.05$ ,  $p = .34$ ]. However, among older-accurate children, those reporting an instance of a repeated event exhibited a higher percentage of hesitations ( $M = 6.48$ ,  $SD = 2.77$ ) than those describing a unique event ( $M = 1.48$ ,  $SD = 1.18$ ),  $t(6) = 3.32$ .

*Internal consistency.* There was a main effect of event frequency [Wilks'  $\lambda = .51$ ,  $F(4, 20) = 4.79$ ,  $\eta^2 = .49$ ]. Follow-up univariate ANOVAs revealed that, compared to children describing a unique event, children describing an instance of a repeated event had a smaller percentage of consistent response pairs [ $F(1, 29) = 17.66$ ,  $\eta^2 = .43$ ], and a greater percentage of response pairs that increased in detail from free to cued recall [ $F(1, 29) = 12.73$ ,  $\eta^2 = .36$ ].

The overall MANOVA for age was also significant [Wilks'  $\lambda = .48$ ,  $F(4, 20) = 5.49$ ,  $\eta^2 = .52$ ]. Follow-up univariate ANOVAs revealed that, compared to older children, younger children had a greater percentage of contradictory response pairs [ $F(1, 29) = 8.64$ ,  $\eta^2 = .27$ ], and response pairs that decreased in detail from free to cued recall [ $F(1, 29) = 11.42$ ,  $\eta^2 = .33$ ]. The overall MANOVA for accuracy was non-significant [Wilks'  $\lambda = .88$ ,  $F(4, 20) = 0.68$ ,  $p = .62$ ,  $\eta^2 = .12$ ].

### Mediation analyses

A mediation model was used to test whether verbal behavior mediated the relationship between age, event frequency, and accuracy classification on the one hand and credibility ratings on the other. Due to the multilevel nature of the verbal behavior data, and to minimize the number of mediation analyses conducted, overall credibility, script-like language, and internal consistency scores were calculated. Specifically, responses to each question on the credibility questionnaire used in Experiments 1 and 2 were summed to create a single credibility score that served as the dependent measure in the mediation analyses. As discussed in the introduction, previous research suggests that descriptions of an instance of a repeated event may be less specific, less powerful, and contain more optionals, present tense, impersonal pronouns, and temporal markers (e.g., Fivush, 1984; Hudson & Nelson, 1986). Thus, a script-like language score was calculated by summing the percentage values for the syntax and powerfulness categories. However, given that particular qualifiers refer to specific information, these percentages are more likely to be attributed to descriptions of unique events and were thus subtracted in the computation of the script-like language score. To compute an internal consistency score, consistent response pairs were assigned a value of two; changes in detail (partial inconsistencies, increase in detail, and decrease in detail) were assigned a value of one; and contradictions were given a value of zero. These values were then summed. To account for different numbers of response pairs across children (not all children provided information for all critical details in the play session), we divided children's internal consistency scores by the total number of response pairs. Thus, overall consistency scores ranged from 0 to 2 (see Brewer et al. (1999) for a similar computation).

What follows is a series of four mediation analyses, two each for Experiments 1 and 2: (1) script language as the mediator through which event frequency, age, and accuracy influence overall credibility scores, and (2) internal consistency as the mediator through which event frequency, age, and accuracy influence overall credibility scores. Mediation analyses followed a four step procedure outlined in Baron and Kenny (1986) and Judd and Kenny (1981).

*Undergraduate sample.* Correlations can be found in Table 6. The first assumption was partially met. Event frequency and age were significantly correlated with overall credibility, but accuracy classification was not. Thus, accuracy classification was excluded from the mediational models. The second and third criteria were not supported when script-like language was the mediator. Script-like language was not significantly correlated with event frequency or age. Further, script-like language was not significantly correlated with overall credibility scores. The lack of a relationship between the mediator and dependent variable prohibited further analyses including script-like language. However, internal consistency scores were significantly correlated with event frequency, age, and overall credibility scores. Regression analyses were then conducted to determine whether internal consistency mediated the relationship between event frequency/age and credibility. The direct effect of event frequency and age on credibility scores was moderate,  $R^2_{\text{adj}} = .19$ ,  $p < .001$ . In a second regression analysis with internal consistency included in the model to control for its effect, age/event frequency contributed significantly to variance accounted for in overall credibility scores (an incremental change to  $R^2$  of .04,  $p < .05$ ). However, the relationship between the independent variables and dependent variable decreased when internal

**Table 6** Correlations between the independent variables (age, event frequency, accuracy), the dependent variable (overall credibility ratings), and the mediators (script-like language and internal consistency) for the mediation analyses in Experiment 1

Variable	1	2	3	4	5	6
1. Age	–					
2. Event frequency	– .03	–				
3. Accuracy	.04	.03	–			
4. Credibility	.24**	– .39**	.06	–		
5. Language	– .07	.08	– .23**	– .13	–	
6. Consistency	.33**	– .56**	.13	.45**	–	–

\* $p < .05$ .\*\* $p < .01$ .

consistency was controlled (age change in  $B = 4.80$  to  $2.96$ ,  $SE = 1.69$  to  $1.78$ ; event frequency change in  $B = -2.62$  to  $1.55$ ,  $SE = 0.56$  to  $0.68$ ). The decrease in the effect of age was significant according to the Sobel test ( $t = 2.67$ ,  $p < .01$ ), suggesting that internal consistency partially mediated the relationship between the independent variables and the dependent variable.

*Community sample.* Correlations can be found in Table 7. Event frequency, age, and accuracy correlated significantly with overall credibility scores in the community sample. The second and third criteria were only partially satisfied when exploring script-like language as the mediator. Script-like language was correlated significantly with accuracy, however, it was not significantly correlated with age, event frequency, or credibility. Thus, script-like language did not mediate the effect of the independent variables on overall credibility ratings. As with the undergraduate sample, the second and third criteria were fulfilled when internal consistency scores were considered as a mediator: They were significantly correlated with event frequency, age, accuracy, and overall credibility scores. Regression analyses were then conducted to determine whether internal consistency mediated the relationship between age/event frequency/accuracy and credibility. The direct effect of event frequency, age, and accuracy on credibility scores was small,  $R^2_{adj} = .04$ ,  $p < .001$ . In a second regression analysis with internal consistency included in the model to control for its effect, age/event frequency/accuracy did not contribute significantly to variance accounted for in overall credibility scores (an incremental change to  $R^2$  of .01,  $p = .51$ ), suggesting that internal consistency fully mediated the relationship between age/event frequency/accuracy and overall credibility ratings.

**Table 7** Correlations between the independent variables (age, event frequency, accuracy), the dependent variable (overall credibility ratings), and the mediators (script-like language and internal consistency) for the mediation analyses in Experiment 2

Variable	1	2	3	4	5	6
1. Age	–					
2. Event frequency	– .03	–				
3. Accuracy	.03	.03	–			
4. Credibility	.15*	– .16*	– .15*	–		
5. Language	– .08	.08	– .19*	– .01	–	
6. Consistency	.34**	– .54**	.19*	.24**	–	–

\* $p < .05$ .\*\* $p < .01$ .

## Discussion

The content analysis involving syntax and powerfulness revealed only one feature that discriminated single- and repeat-event reports, which was qualified by an interaction. Among older children classified as high-accurate, there were more hesitations in repeat-event children's reports than single-event children's reports. It is possible that hedges also discriminate reports of single- and repeat-events, but our values were too low to confidently complete inferential statistics. Given our review of literature indicating that event frequency influences the content and presentation of memory reports and the findings of Experiments 1 and 2 that event frequency had a significant and broad effect on ratings of credibility, we were surprised that the content analyses did not reveal more differences. However, it is instructive that syntax and powerfulness, at least when combined, were not correlated with credibility ratings.

On measures of syntax and powerfulness, there were several differences as a function of child's age. Younger children included a smaller percentage of temporal markers, particular qualifiers, and conditional tense. Also, among children classified as low-accurate, younger children included a higher percentage of hesitations in their reports than older children. Although the content analysis revealed several differences in reports of younger and older children, the results of Experiments 1 and 2 revealed very few differences in credibility ratings as a function of age. As noted above, this is consistent with the correlation analyses that syntax and powerfulness, when combined, were not correlated with credibility.

These two findings independently and collaboratively support the notion that neither syntax nor powerfulness influenced ratings of credibility in the present study. When differences in syntax and powerfulness were absent, as with event frequency, credibility ratings varied significantly. When there were differences in syntax and powerfulness, as across age, there were few differences in credibility ratings.

The content analyses involving internal consistency discriminated groups on the basis of event frequency and on the basis of age. Specifically, relative to reports of repeat-event children, reports of single-event children contained more consistencies and fewer instances where more specific information was provided in cued than free recall. And, compared to older children, younger children's reports contained more contradictions and were more likely to include less specific information in cued than free recall. Importantly, these differences in internal consistency mediated the relationship between age, event frequency, and accuracy on the one hand and credibility ratings on the other hand in the community sample and partially explained the relationship in the undergraduate sample. This agrees with the literature that inconsistent reports are often seen as inaccurate (Berman et al., 1995; Brewer et al., 1999; Myers et al., 1999).

We did not identify a single characteristic that discriminated children classified as high-accurate from those classified as low-accurate. This is not surprising based on the literature and supported by the data reported in Experiments 1 and 2, that mock jurors are not skilled at discriminating children's reports on the basis of accuracy. These data suggest that, at least among the features included in our content analysis, there are few differences that could be used to make such a discrimination.

## General discussion

### Perceptions of credibility and event frequency

Across two experiments, children who reported details of an instance of a repeated event that varied across experiences were judged to be less competent rememberers than children who



reported details of the same event that was experienced once. This pattern was observed in spite of the fact that actual accuracy was comparable across the two frequency conditions and the studies involved 15 different children who reported the event as unique and 16 different children who reported the same event as the last instance in a sequence of four similar events.

These results were predicted based on research that demonstrates that, compared to reports of unique events, the content and presentation style of children's reports of instances of repeated events that vary across instances is distinct in ways that can decrease perceptions of credibility. In particular, reports of instances of repeated events can be less consistent (e.g., Connolly & Price, 2006; Powell & Roberts, 2002; Powell & Thomson, 1996, 2003), less likely to be distinctly associated with a particular instance (e.g., when reports include the impersonal pronoun "you" and the present tense; Fivush, 1984, 1997; Hudson & Nelson, 1986), possibly reported with less confidence (Roberts & Powell, 2005), more suggestible (Connolly & Lindsay, 2001; Connolly & Price, 2006; Powell & Roberts, 2002; Price & Connolly, 2004), and reported in language that may appear less powerful. We provide some evidence that perceptions of these characteristics discriminated our single- and repeat-event children. In Experiment 1, repeat-event children were perceived to be less consistent, less confident, less likeable, and more suggestible. Similarly, in Experiment 2, compared to single-event children, repeat-event children were seen as marginally less consistent, more suggestible less confident (inaccurate only), and less likeable (younger low-accurate only).

Importantly, the data from Experiments 1 and 2 reflect *perceptions* rather than actual occurrence in children's reports. We offer empirical support for the accuracy of some of these perceptions. In the report of the suggestibility study from which these videotapes were obtained, we found that repeat-event children were sometimes more suggestible than the single-event children (Connolly & Price, 2006). And, in Study 3 we found that, compared to single-event children, repeat-event children were indeed less consistent and this mediated the relationship between event frequency and perceptions of credibility.

In summary, repeat-event children were seen as less credible than single-event children, possibly because they were perceived to be less consistent, sometimes less confident, and more suggestible than single-event children. The accuracy of participants' perceptions of heightened suggestibility is supported in Connolly and Price (2006) and perceptions of lower consistency is supported in Study 3.

To the best of our knowledge, there are only three published reports of the effect of prior experience with an event on credibility ratings (Blandon-Gitlin et al., 2005; Pezdek et al., 2004; Strömwall et al., 2004). In all three studies, prior experience with the event increased ratings of credibility (as measured by CBCA scores and, in Strömwall et al., Reality Monitoring scores). Importantly, in each of the published studies, either the observation (Pezdek et al., 2004) or manipulation (Blandon-Gitlin et al., 2005; Strömwall et al., 2004) of prior experience primarily involved details common to all experiences (i.e., fixed details). In the present report, the repeated event contained 16 details that changed across instances (i.e., variable details) and children's attention was drawn to the details that varied. Although a report of an instance of a repeated event that contains primarily fixed details may be judged to be more credible than a report of the same event that was experienced once, this may not be true of reports that contain many variable details.

#### Perceptions of credibility and accuracy

Neither undergraduates nor community members were able to discriminate between high- and low-accurate reports of single- or repeat-event children. In fact, community members rated low-accurate children to be more cognitively competent and more confident than high-accurate

children. The lack of discrimination of high- and low-accurate children is not unusual in the extant literature. Although most studies report no difference in credibility scores of accurate and inaccurate children (Ceci, Crotteau-Huffman et al., 1994; Ceci, Loftus et al., 1994; Goodman et al., 2002; Leippe et al., 1992, Exp. 2), among studies that have reported a difference, it is more likely to be in the wrong direction. That is, inaccurate children are rated as more credible than accurate children (Goodman et al., 1989; Leichtman & Ceci, 1995; Talwar & Lee, 2002; Tye et al., 1999).

We have not been able to find an explanation in the literature for below-chance discrimination; however, an explanation for chance performance has been offered. Some scholars have suggested that children who are inaccurate come to believe their reports and do not exhibit perceptible signs of deception (e.g., Leichtman & Ceci, 1995; Ceci, Crotteau-Huffman et al., 1994). If there are no perceptible signs of deception among children who come to believe their inaccurate reports, there should be no effect of actual accuracy on ratings of credibility. In agreement with this possibility, in the present report, we did not find an effect of accuracy classification on perceptions of report characteristics that predict credibility (i.e., consistency, likeability, and suggestibility) nor were we able to discriminate accurate from inaccurate reports on the basis of syntax, speech powerfulness, or consistency.

#### Perceptions of credibility and age

Only in Experiment 1 were younger children rated as less competent rememberers than older children. To the extent that ratings of credibility were based on judgments of cognitive competence rather than judgments of honesty, this finding is entirely consistent with previous literature (e.g., Bottoms, 1993). In the present study, children were directed to recall particular details of the unfamiliar target event. Under these conditions, it is expected that cognitive competence would be a more salient characteristic of credibility and, therefore, more emphasis placed on it. Indeed, it was on ratings of cognitive competence and credibility in Experiment 1 that younger children received ratings lower than those given to older children, not on ratings of honesty.

Although measures of perceptions of report characteristics that predict credibility did not provide much insight into why younger children were rated below older children, (i.e., younger children were judged to be less confident than older children but not less consistent or more suggestible) the content analysis was informative. Younger and older children's reports differed on several of the measures of content (fewer temporal markers, fewer particular qualifiers, a smaller percentage of conditional tense, more contradictions, and more incidents of decreasing amount of detail reported in cued compared to free recall). The results of the mediation analysis revealed that differences on measures of consistency mediated the relationship between age and overall credibility ratings.

#### Limitations and future directions

There are several limitations to the present experiments. First, children in the present studies reported neutral and/or positive events. It is possible that these findings will not generalize to situations in which children testify about more personally salient events. Second, credibility ratings may not predict verdict in high-stakes situations such as criminal court cases. However, if even a small reduction in credibility ratings raises a reasonable doubt about the accuracy of the allegation in a small number of cases, the findings are important and relevant. Third, participants were not told about the frequency manipulation. In actual court cases, jurors likely know if the allegation involves a single or repeated event and judgments of credibility may be adjusted accordingly. However, the direction of the adjustment is unclear and likely depends

on participants' intuition about the nature of repeated events as well as their characterization of a credible report. If participants believe that repeated events contain largely fixed details, they may reason that memory for particular instances should be strong, leading to a decrease in credibility ratings of a report that contains internal inconsistencies. Or, if participants believe that repeated events contain largely variable details, they may expect some confusion and view internal inconsistencies as uninformative in terms of assessing credibility. Alternatively, participants may ignore event frequency. That is, they may reason, for the purposes of evaluating credibility, what matters is *if* a child knows what occurred on a particular occasion rather than *why* a child appears not to know precisely what happened. Fourth, in the present study free recall was followed by very particular questions about details of the play session. This is contrary to recommendations for good interviewing techniques and could impact perceptions of credibility (Tubb et al., 1999). Although in-court testimony is likely to include specific questions, it is unlikely to contain a fairly long sequence of very specific questions about particular details. One consequence of our methodology could be to blur the distinction between single- and repeated-event reports because differences in language and presentation style would be minimized in cued recall. Given that we found a difference, we do not think that this compromised our findings. Had the children in our study provided longer free recall narratives and not been presented with cued recall questions, our results may have been stronger.

## Conclusion

Often, when children testify in court as witnesses/victims, they do so as victims of a repeated offence and may be required to report details of an instance or a subset of instances of the repeated offence. Recently, researchers have found that under some circumstances, children are more suggestible about details of a repeated event compared to details of a unique event (Connolly & Lindsay, 2001; Connolly & Price, 2006; Powell & Roberts, 2002; Price & Connolly, 2004). This study highlights another forensically important difference between reports of single and repeated events. Children's reports of an instance of a repeated event may appear less credible than an equally accurate report from a child who experienced the event once. Accordingly, children who experience repeated abuse may have several obstacles to overcome: The need to provide particular details in accordance with the specificity principle, heightened suggestibility, and the possible deleterious effect of the manner of reporting on perceptions of credibility.

**Acknowledgments** This research was supported by a Social Sciences and Humanities Research Council (SSHRC) operating grant to the first author, a Natural Sciences and Engineering Research Council graduate scholarship to the second author, and SSHRC graduate scholarships to the third and fourth authors. The authors thank Catharine Martin, Pam Johnson, Scott Currie, Jef Grainger, Alexa Dunbar, Christine McBeth, Melissa Northcott, and Michelle Houle for their assistance with data collection.

## References

- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182.
- Berman, G. L., & Cutler, B. L. (1996). Effects of inconsistencies in eyewitness testimony and mock-juror judgments. *Journal of Applied Psychology*, *81*, 170–177.
- Berman, G. L., Narby, D. J., & Cutler, B. L. (1995). Effects of inconsistent eyewitness statements on mock-jurors' evaluations of the eyewitness, perceptions of defendant culpability and verdicts. *Law and Human Behavior*, *19*, 79–88.

- Blandon-Gitlin, I., Pezdek, K., Rogers, M., & Brodie, L. (2005). Detecting deception in children: An experimental study of the effect of event familiarity on CBCA Ratings. *Law and Human Behavior, 29*, 187–197.
- 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.
- Brewer, N., Potter, B., Fisher, R. P., Bond, N., & Luszcz, M. A. (1999). Beliefs and data on the relationships between consistency and accuracy on eyewitness testimony. *Applied Cognitive Psychology, 13*, 297–313.
- Bruck, M., Ceci, S. J., & Hembrooke, H. (2001). Reliability and credibility of young children's reports: From research to policy and practice. In R. Bull (Ed.), *Children and the law: The essential readings* (pp. 87–123). MA, US: Blackwell.
- 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.
- Ceci, S. J., Crotteau-Huffman, M. L., Smith, E., & Loftus, E. F. (1994). Repeatedly thinking about a non-event. Source misattributions among preschoolers. *Consciousness and Cognition: An International Journal, 3*, 388–407.
- Ceci, S. J., Loftus, E. F., Leichtman, M. D., & Bruck, M. (1994). The possible role of source misattributions in the creation of false beliefs among preschoolers. *International Journal of Clinical and Experimental Hypnosis, 42*, 304–320.
- Connolly, D. A., & Lindsay, D. S. (2001). The influence of suggestions on children's reports of a unique experience versus an instance of a repeated experience. *Applied Cognitive Psychology, 15*, 205–223.
- Connolly, D. A., & Price, H. L. (2006). Children's suggestibility for an instance of a repeated event versus a unique event: The effect of degree of association between variable options. *Journal of Experimental Child Psychology, 93*, 207–223.
- Conte, J. R., Sorenson, E., Fogarty, L., & Rosa, J. D. (1991). Evaluating children's reports of sexual abuse: Results from a survey of professionals. *American Journal of Orthopsychiatry, 78*, 428–437.
- Cutler, B., Penrod, S., & Stuve, T. (1988) Juror decision making in eyewitness identification cases. *Law and Human Behavior, 12*, 41–55.
- Farrar, M. J., & Goodman, G. S. (1990). Developmental differences in the relation between scripts and episodic memory: Do they exist? In R. Fivush & J. A. Hudson (Eds.), *Knowing and remembering in young children* (pp. 30–64). New York: Cambridge University Press.
- Fivush, R. (1984). Learning about school: The development of kindergartner's school script. *Child Development, 55*, 1697–1709.
- Fivush, R. (1997). Event memory in early childhood. In N. Cowan (Ed.), *The development of memory in childhood* (pp. 139–161). UK: Psychology Press.
- Fivush, R., Haden, C., & Adam, S. (1995). Structure and coherence of preschoolers' personal narratives over time: Implications for childhood amnesia. *Journal of Experimental Child Psychology, 60*, 32–56.
- Fivush, R., Hudson, J., & Nelson, K. (1984). Children's long-term memory for a novel event: An exploratory study. *Merrill-Palmer Quarterly, 30*, 303–316.
- 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.
- Hudson, J. A. (1990). Constructive processes in children's event memory. *Developmental Psychology, 26*, 180–187.
- Hudson, J., & Nelson, K. (1986). Repeated encounters of a similar kind: Effects of familiarity on children's autobiographic memory. *Cognitive Development, 1*, 253–271.
- Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review, 5*, 602–619.
- Lamb, M. E. (1999). Child witnesses: Recent research on children's accounts of forensically relevant experiences. *Applied Developmental Science, 3*, 2–5.
- Leichtman, M. D., & Ceci, S. J. (1995). The effects of stereotypes and suggestions on preschoolers' reports. *Developmental Psychology, 31*, 568–578.
- Leippe, M. R., Manion, A. P., & Romanczyk, A. (1992). Eyewitness persuasion: How and how well do fact finders judge the accuracy of adults' and children's memory reports? *Journal of Personality and Social Psychology, 63*, 181–197.
- Leippe, M. R., & Romanczyk, A. (1989). Reactions to child (versus adult) eyewitnesses: The influence of jurors' preconceptions and witness behavior. *Law and Human Behavior, 13*, 103–132.
- Leippe, M. R., Romanczyk, A., & Manion, A. P. (1991). Eyewitness memory for a touching experience: Accuracy differences between child and adult witnesses. *Journal of Applied Psychology, 76*, 367–379.

- Lindsay, R. C. L., Wells, G. L., & O'Connor, F. (1989). Mock juror belief of accurate and inaccurate witnesses: A replication. *Law and Human Behavior*, *13*, 333–340.
- Luus, C. A. E., & Wells, G. L. (1992). The perceived credibility of child eyewitnesses. In H. Dent, & R. Flin (Eds.), *Children as witnesses* (pp. 73–92). Chichester, UK: John Wiley & Sons.
- Luus, C. A. E., & Wells, G. L. (1994). The malleability of eyewitness confidence: Co-witness and perseverance effects. *Journal of Applied Psychology*, *79*, 714–724.
- 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.
- Myers, J. E. B., Redlich, A. D., Goodman, G. S., Prizmich, L. P., & Imwinkelreid, E. (1999). Jurors' perceptions of hearsay in child sexual abuse cases. *Psychology, Public Policy and Law*, *5*, 388–419.
- O'Barr, W. M. (1982). *Linguistic evidence: Language, power, and strategy in the courtroom*. New York: Academic Press.
- Pezdek, K., Morrow, A., Blandon-Gitlin, I., Goodman, G. S., Quas, J. A., Saywitz, K. J., Bidrose, S., Pipe, M-E., Rogers, M., & Brodie, L. (2004). Deception detection in children: Event familiarity affects criterion-based content analysis ratings. *Journal of Applied Psychology*, *89*, 119–126.
- Powell, M. B., & Roberts, K. P. (2002). The effect of repeated experience on children's suggestibility across two question types. *Applied Cognitive Psychology*, *16*, 367–386.
- Powell, M. B., & Thomson, D. M. (1996). Children's memory of an occurrence of a repeated event: Effects of age, repetition, and retention interval across three question types. *Child Development*, *67*, 1988–2004.
- Powell, M. B., & Thomson, D. M. (2003). Improving children's recall of an occurrence of a repeated event: Is it a matter of helping them to generate options? *Law and Human Behavior*, *27*, 365–384.
- Price, H. L., & Connolly, D. A. (2004). Event frequency and children's suggestibility: A study of cued recall responses. *Applied Cognitive Psychology*, *18*, 809–821.
- R. v. B.(G.)*, (1990) 2 S.C.R. 3.
- Roberts, K. P., & Powell, M. B. (2005). Evidence of metacognitive awareness in young children who have experienced a repeated event. *Applied Cognitive Psychology*, *19*, 1019–1031.
- Ross, D. F., Dunning, D., Tolia, 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.
- Strömwall, L. A., Bengtsson, L., Leander, L., & Granhag, P. A. (2004). Assessing children's statements: The impact of a repeated experience on CBCA and RM ratings. *Applied Cognitive Psychology*, *18*, 653–668.
- Talwar, V., & Lee, K. (2002). Development of lying to conceal a transgression. Children's control of expressive behaviour during verbal deception. *International Journal of Behavioural Development*, *26*, 436–444.
- Talwar, V., Lee K., Bala N., & Lindsay, R. C. L. (2006). Adults' judgments of children's coached reports. *Law and Human Behavior*, *30*, 561–570.
- 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.
- Tye, M. C., Amato, S. L., Honts, C. R., Devitt, M. K., & Peters, D. (1999). The willingness of children to lie and the assessment of credibility in an ecologically relevant laboratory setting. *Applied Developmental Science*, *3*, 92–109.