



City Research Online

City St George's, University of London

Citation: Bettenay, C., Ridley, A. M., Henry, L. & Crane, L. (2014). Cross-examination: The Testimony of Children With and Without Intellectual Disabilities. *Applied Cognitive Psychology*, 28(2), pp. 204-214. doi: 10.1002/acp.2979

This is the unspecified version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://city-ref-test.eprints-hosting.org/id/eprint/3824/>

Link to published version: <https://doi.org/10.1002/acp.2979>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

Cross-examination of children with and without intellectual disability

Applied Cognitive Psychology, 28 (2), 204-214

Cross-examination: The testimony of children with and without intellectual disabilities

Bettenay, C., Ridley, A., Henry, L.A. & Crane, L.

Abstract

The present study assessed how children with a range of cognitive abilities fared during a mock cross-examination. Ninety children (aged 4 to 11 years; 18 with intellectual disabilities [ID], 13 with borderline intellectual disabilities [BID], and 59 who were typically developing [TD]) witnessed a staged event, participated in an initial forensic interview (a few days later), and were cross-examined by a barrister-in-training (ten months later). During cross-examination, 98% of all children changed at least one response from their initial interview when challenged. However, group differences in performance (total number of changed responses, 'resistance' to challenges), controlling for age and memory for event details, were not significant or did not prove reliable at the level of individual group contrasts. Overall, little robust evidence for group differences in performance on cross-examination could be identified, and *memory for event details* was the most reliable predictor of performance.

Keywords: Child witnesses; cross-examination; intellectual disabilities

Cross-examination of children with and without intellectual disabilities

In an adversarial system of justice such as that of the UK, Australia and the USA, there is a strong emphasis on oral testimony from witnesses about the facts of a disputed case (Ellison, 2001). In court, following the presentation of direct evidence by the prosecution (testimony from the victim, witness or defendant), cross-examination is undertaken by opposing counsel to challenge the reliability of a witness's evidence and, ostensibly, to search for the truth (Wellman, 1986; Yarmey, 1979). Effective cross-examination highlights inconsistencies in witness testimony. Yet the techniques employed to do this, such as pressing the witness to change their response (Zajac, Gross & Hayne, 2003), accusing the witness of lying (Davies, Henderson & Seymour, 1997; Plotnikoff & Woolfson, 2009; Spencer, 2012), repetitive and complex questioning (Plotnikoff & Woolfson, 2012; Zajac, 2009) and deliberately setting the sequence of questioning to confuse the witness (Glissan, 1991) are, in reality, concerned with discrediting a witness (Henderson, 2002). The demands on a witness to produce reliable oral evidence, often many months or even years, after an event are high, and witnesses find the process stressful, aggressive and anxiety provoking (Plotnikoff & Woolfson, 2009, 2012; Zajac, 2009).

Legal reforms to procedures during evidence gathering and submission of evidence-in-chief to courts have taken place in a number of countries over the last 20 years, to protect children and other vulnerable witnesses. In England and Wales, guidelines such as the Memorandum of Good Practice (Home Office, 1992) and Achieving Best Evidence (ABE: Ministry of Justice, 2011) as well as other highly regarded protocols (e.g. NICHD: Kuehnle & Connell, 2009; Lamb, Herskowitz, Orbach & Esplin, 2008; Lamb, La Rooy, Malloy & Katz, 2011) maximise the use of open-ended prompts and ensure that interviewing techniques avoid suggesting information about the events under discussion. As well as ensuring that a child's 'best evidence' is obtained, these guidelines aim to reduce stress and delay (Zajac,

2009). If a case comes to court in England and Wales, video recordings of forensic interviews should be used to replace (in full or in part) direct examination by the prosecution. However, the questions (and questioning style) used during cross-examination often run counter to available research/protocols, and the use of ABE guidelines does not extend to the process of cross-examination. In fact, legal professionals show considerable resistance to proposals to alter cross-examination procedures (see Spencer & Lamb, 2012; Zajac, O'Neill, & Hayne, 2012, for a discussion). Consequently, children are exposed to the same cross-examination techniques that are used on adults (Zajac, 2009).

The negative impact of cross-examination on children's testimony was noted by Zajac, Gross and Hayne (2003), who examined the real-life court transcripts of 5-13 year old children who were key witnesses in sexual abuse trials. During cross-examination, 75% of these children changed at least one aspect of their testimony, with some withdrawing allegations of abuse completely. Children rarely asked for clarification, and also attempted to answer questions that they did not understand. However, the nature of this study meant that it was not possible to provide an objective 'ground truth' benchmark for the accuracy of the children's original responses.

In an empirical investigation of cross-examination performance addressing this issue, Turtle and Wells (1988) showed 8-12 year old children (and a sample of adults) a short video on which they were subsequently interviewed and subjected to cross-examination. Perhaps unsurprisingly, they found that all participants were less accurate during cross-examination than during the initial interview, an effect that was particularly pronounced for younger children. However, only a few cross-examination questions were asked and the nature of these questions is unclear (Zajac, 2009). In addition, there was only a 24 hour period between the initial interview and cross-examination, which is not representative of the average delay in actual proceedings. Current estimates of delay vary, and depend on the type of court and

the jurisdiction. However, in Crown Courts in England and Wales, delays can be up to 30 months (from the time of the defendant's first court appearance rather than disclosure by the child), with an average of around 8 months (Plotnikoff & Woolfson, 2012). Delays are considerably longer elsewhere, for example, in Northern Ireland and New Zealand (Henderson, 2012). Therefore, the length of time between a child providing their initial evidence to the police and their appearance in court can be considerable.

In an experimental study with realistic time delays, Zajac and Hayne (2003) conducted initial interviews six weeks after 5-6 year old children viewed a live event (as opposed to a video). Cross-examinations took place nine months later (ten months after the event). The researchers found that 85% of children made at least one change to their previous statements and one third changed all of their original responses (also see Zajac, Jury, & O'Neill, 2009). A less pronounced, but equally worrying, pattern was subsequently observed in older children (aged 9-10 years), with 70% changing at least one response during cross-examination (Zajac & Hayne, 2006). Of particular concern, 43% of this older group changed their originally correct answers to incorrect ones.

For children with an intellectual disability (ID), who are likely to have poorer cognitive abilities (Brown & Geiselman, 1990) and greater levels of suggestibility (e.g. London, Henry, Conradt & Corser, 2013) than typically developing (TD) children of the same chronological age, cross-examination may prove even more problematic (Zajac et al., 2012). ID is '*the most common developmental disorder and the most handicapping of the disorders beginning in childhood*' (Harris, 2006). It is characterised by significant cognitive deficits (an IQ < 70) that have an onset before the age of 18, and significant difficulties with adaptive functioning (American Psychiatric Association, 2000). Those children who display IQs between the range for children with ID and TD children (i.e. those with IQs of 70-84) are described as having 'borderline ID' (Alloway, 2010). Children and adults with ID are a

heterogeneous group (Kebbell & Hatton, 1999) and their overall IQ scores alone tell us relatively little about their cognitive and behavioural profiles. Nevertheless, they are a vulnerable population (Westcott & Jones, 1999) who are at increased risk of maltreatment, abuse, and sexual violence (Brown & Stein, 1998; Hershkowitz, Lamb & Horowitz, 2007; Lin, Yen, Kuo, Wu & Lin, 2009; Sullivan & Knutson, 2000; Westcott, 1991), as well as often being the only witnesses to others' crimes against those with ID (Milne, 1999).

Individuals with ID may not participate fully in the legal system as victims and witnesses (Kebbell & Hatton, 1999; Zajac et al., 2012) for several reasons, including a lack of identification as victims by authorities, as well as communication problems (Kendall-Tackett, Lyon, Taliaferro, & Little, 2005). Individuals with ID are also perceived as less credible witnesses than their typical peers (Henry, Ridley, Perry, & Crane, 2011; Peled, Iarocci, & Connolly, 2004) and their access to justice may be hindered by the unwarranted assumption that they are inherently unreliable as witnesses (Peled et al., 2004). In actual fact, children with ID are under-researched in the eyewitness testimony literature and little is known about the competencies of these children within the adversarial system and how they would cope with cross-examination.

In relation to initial questioning following a witnessed event, children with ID have been found to produce limited detail in response to free recall instructions, but the information that they do provide is nevertheless very accurate (Henry & Gudjonsson, 1999; 2004; Michel, Gordon, Ornstein & Simpson, 2000; although see Agnew & Powell, 2004). Children with ID are no more suggestible than children of a similar mental age (Henry & Gudjonsson, 1999; Jens, Gordon and Shaddock, 1990; Michel et al., 2000) and although they generally show greater suggestibility than children of a similar chronological age, this is not always the case (Agnew & Powell, 2004; Henry & Gudjonsson, 2003). However, children with ID require more questioning to elicit information (Agnew & Powell, 2004) and are

slightly more vulnerable to misleading ‘yes/no’ questions (Henry & Gudjonsson, 1999; 2003). They also find repeated questioning problematic, with a substantial number changing their responses to such questions (Cederborg, Danielsson, La Rooy, & Lamb, 2009; Henry & Gudjonsson, 2003). Performance appears to be related to level of ID: children with moderate ID experience a greater number of difficulties than children with mild/borderline ID (Brown, Lewis, Lamb & Stephens, 2012; Henry & Gudjonsson, 2003; Michel et al., 2000). However, little is known about the capabilities of children with borderline ID (IQs of 70-84) in this regard. Brown et al. (2012) assessed children with mild-borderline ID but did not report data separately for those in the borderline group, therefore one of the aims of the current study was to distinguish between those with mild/moderate and borderline ID.

Despite several research studies providing an insight into the strengths and difficulties of children with ID during initial interviews, there has been a lack of empirical research undertaken to assess how children with ID fare during cross-examination (Kebbell, Hatton, & Johnson, 2004; Zajac et al., 2012). As children with ID experience difficulties with exactly the types of questions that characterise the cross-examination process (e.g., repeated, complex and suggestive questions), cross-examination may pose greater problems for children with ID than for TD children (Zajac et al., 2012). Further, lawyers and judges are unlikely to adapt their practices to meet the needs of individuals with ID (Kebbell et al., 2004), with court transcripts revealing that the technique of cross-examination is ‘particularly poor’ for eliciting accurate information from witnesses with ID (Kebbell et al., 2004). This is an extremely important issue to address bearing in mind the current debate across many countries that use an adversarial system; namely, should very young children, or individuals with ID, be exempted from cross-examination in court? (Spencer, 2012).

Based in the UK, the current study explored how 4-11-year-old children with and

without ID fared during a mock cross-examination. To achieve this, all children viewed a structured live event (a magic show), and participated in an initial interview (following ABE guidelines) 3-6 days later. Ten months later, they underwent an ecologically valid cross-examination by a barrister-in-training. In line with previous research, it was predicted that performance during the initial interview would be affected by the child's level of ID; the more severe the ID, the more likely they would be to produce fewer items of information (although what they did recall was predicted to be just as accurate). In relation to cross-examination, predictions were more tentative given the absence of existing literature. It was hypothesised that all children would change a large proportion of their answers, in line with work on TD children, but that the number of changes may differ as a function of ID level: the greater the level of ID, the less resilient to cross-examination challenges the child would be (changing a higher number of responses and ceding at an earlier stage of the cross-examination challenge). We were also careful to control for two important variables in all analyses before examining group differences. These were *age* (it was assumed that younger children would be more vulnerable to cross-examination) and *recall for details of the event* (we included a set of 'unchallenged' questions about basic information from the witnessed event; these questions were interspersed between the cross-examination challenges to provide a measure of 'recall for unchallenged details').

Method

Participants

The sample comprised 90 children (40 males) aged 4 years 7 months to 11 years 1 month (mean = 8 years 9 months, SD = 1 year 8 months). This age range was selected as it allowed an examination of the role of age on cross-examination performance (encompassing a range of ages utilised in previous research), but was restricted enough to ensure that the

staged event was suitable for all of the children. All participants attended either an ‘inclusive’ mixed ability mainstream primary school or a ‘special school’ for children with learning disabilities, both of which were located in Greater London. Using a measure of intellectual ability (the Stanford Binet Version 5; Roid, 2003), those children with an IQ between 35 and 69 were classified as having a mild to moderate level of ID (n=18); those with an IQ between 70 and 84 were classified as having Borderline ID (n=13); and those with an IQ of 85 and higher were classified as being typically developing (TD; n=59). The TD group included a number of children who were younger than the ID and BID groups, but whose mental ages were equivalent to those with ID. As such, the TD group included children with both *chronological* and *mental* ages of a similar range to those in the ID sample. This is in line with a ‘developmental trajectories’ approach (Thomas, Annaz, Ansari, Scerif, Jarrold, & Karmiloff-Smith, 2009), which ensures that comparison samples reflect the range of abilities of the target sample, as opposed to being individually matched on specified variables.

The children with ID and BID were of mixed aetiology with no specific diagnoses made available to the researchers. The TD group had no special needs classifications and did not attend any additional special classes. See Table 1 for participant information.

[Place Table 1 about here]

Materials and Procedure

This study was conducted in three phases.

Phase 1 – Staged event. The children viewed one of seven identical live, scripted magic shows at their school. In an attempt to minimise schema driven memories of magic shows, where a script might include a man in a traditional black cape and top hat, the magician was female (‘Auntie Julie’) and dressed in a colourful outfit. The show consisted of

eight tricks (presented in the same order each time) and lasted 20 minutes. To encourage the children (particularly the older ones) to attend fully to the content of the show, the magician explained at the start that she needed their help in testing out a new show for young children in hospital and that, at the close of the show, a 'vote' would take place to see what age range the show was suitable for (5 year olds, 10 year olds, or everybody). To further maximize the children's attention to the event, the show included a number of tricks requiring all children to interact with the magician both verbally ('call out') and non-verbally ('point'). A small number of the children (two per show) were also asked to assist the magician with a trick but, as their experience was qualitatively different to that of their peers in the audience (in that they viewed the show from a different perspective), their data were not included. At the end of the show, the children did not receive an instruction regarding whether or not they could discuss the show with their classmates; this would not necessarily occur in an actual criminal investigation and, given the number of children who viewed the show, it was not possible to reliably enforce such an instruction.

Phase 2 – Initial interview. All children were interviewed 3-6 days after the event by one of two female researchers who were not present at the magic show. Interviewer 1 was a former police officer with specialised training in interviewing children: Interviewer 1 trained Interviewer 2 prior to the study. The format for the interviews was discussed and agreed beforehand and all interviews were conducted according to Achieving Best Evidence (ABE) guidelines in place in England and Wales at the time (Home Office, 2007). Interviews were overtly video and audio taped. They lasted approximately 30-40 minutes, but varied for each child depending on how much they could remember (particularly during the initial free recall). For details of the initial interview please see Appendix 1.

Free recall was coded by giving children one point for every correct piece of information about the show, and these were recorded (e.g. 'the lady [1 point] did tricks [1

point]’, ‘the lady made Harry Potter’s wand fall off the table [5 points]’ and ‘I saw a show [1 point]’). Prompted general recall was coded in the same way: children received one point for each correct piece of information over and above that provided in free recall. Prompted specific recall was coded in the same way except that children received credit for information already recalled during free or prompted general recall. Errors, coded across categories of prompt were defined as: mistakes of detail (e.g. getting the colour of the magician’s hair wrong); and, confabulated information concerning details added by the child that did not occur (e.g. the magician fell over and banged her head). Responses such as ‘I’m not sure’ or ‘I can’t remember’ were combined as ‘don’t know’ responses. A random sample of 25% of these interviews were coded for consistency (across all items; free and prompted general and specific recall) by two independent raters ($r = .89$).

Phase 3 – Cross-examination interview. To reflect current court delays in England and Wales, all children underwent a mock cross-examination at their school ten months after the initial interviews. These were conducted by nine barristers-in-training, who each volunteered to perform multiple interviews (range = 5 to 23 interviews; mode = 6). Each child was taken individually from their class to a quiet room and given a brief explanation of the running order of the session. They were advised that they would be meeting a barrister, and an age-appropriate explanation of the barrister’s role was given. The children were asked if they were happy to proceed, and all were.

After conveying and introducing each child to the barrister, the first author left the room and the barrister said ‘*I’m going to show you a video that you made a few months ago with a lady called (name inserted here). You need to watch the video carefully and listen to what you said to her. I’m then going to ask you some questions about what you said, and I need you to listen carefully and then answer me truthfully, ok?*’

Children watched the video of their initial interview alongside the barrister-in-training (excluding the rapport and truth/lies phase). This was in order to adhere to required practice in England and Wales, whereby children's evidence is provided to the court as pre-recorded evidence-in-chief under 'Special Measures' (in England and Wales), prior to cross-examination (Home Office, 2007; HM Crown Prosecution Service Inspectorate/HM Inspectorate of Constabulary, 2012). Barristers were instructed to direct the child back to the video if their attention wandered, as may occur in an actual criminal case. The barrister stopped the video evidence after they had both viewed it in full and said to the child *'Now I am going to ask you some questions about this video.'*

In court, cross-examination interviews usually flow in a manner dependent on what arises during questioning. As this study needed to balance ecological validity with experimental rigor, the cross-examination questions were structured to allow children's performance to be compared directly. To achieve this, elements of the magic show on which all children had been able to answer questions in the initial interview were identified. A set of draft questions common to all the children, and easily adapted to take into account individual variations in actual testimony, was formulated. Four-part structured challenges were then drafted, designed to exert increasing pressure upon the child to change their responses from their earlier testimony. Barristers-in-training were directed to complete all four parts of each of the challenge questions unless, and until, the child ceded to a challenge. At that point, they immediately moved on to the next question. If the child said that they did not know the answer, the barrister moved on to the next question. For examples please see Appendix 2.

Although the child participants were under the impression that the questions were entirely derived from their evidence, in fact the questions were closely scripted to ensure that each child was asked identical topic-related questions. A number of the topic-related questions were tailored to reflect the child's actual testimony, adding to the effect that the

child was being challenged on their own evidence (e.g., all children were asked about the magician's three coloured handkerchiefs, but the colours used in the question reflected those that they had specified in the initial interview).

The full cross-examination interview included 23 questions, of which 12 were challenges to their earlier evidence. A further 11 were straightforward unchallenged questions from the initial interview (not misleading or designed to confuse the child), and scores on these questions served as an important control measure in the analyses, namely, 'recall for unchallenged details of the event'. The unchallenged questions were alternated with the challenged questions to produce a more realistic flow to the interview and also to allow the child breathing space between the larger four-part challenging questions. Challenges were on a range of general ($n = 4$) and specific ($n = 8$) details, and covered both events that did happen ($n = 8$) and events that did not happen ($n = 4$). Examples from the interview protocol are included in Appendix 1. The entire cross-examination session lasted about 45 minutes (this included the time taken for the children to watch their initial interview; the actual cross-examination questioning lasted around 20-25 minutes). All children were instructed not to discuss the session with their classmates.

The data were coded on a number of parameters. First, the overall number of changed responses (out of the 12 four-point challenges put to the children) was calculated, to measure the child's resistance to cross-examination (referred to as the 'total number of cedes'). A score of zero was assigned if the child did not change a response, and a score of one was assigned if they ceded to cross-examination pressures (max score = 12). An outright change to their testimony (Q: *'Are you sure the magician did tricks?'*, A: *'No'*), agreement with the barrister on any challenge (Q: *'I don't believe you did see the book change, did you?'*, A: *'I didn't'*), or an acceptance of a different explanation (Q: *'If your friends said the magician didn't do tricks, they would be right, wouldn't they?'*, A: *'They might be'*) would all be

accepted as ceding to the challenge. '*I don't know*' or similar responses were *not* taken as ceding to a challenge. Where the child offered their own words to back up their version of events, rather than answer the challenge directly, this was also coded as not ceding. The child had to *explicitly* give way during the question in order to be classified as having ceded to cross-examination. When the barrister deemed the child to have ceded to cross-examination, he/she moved onto the next question or challenge (N.B. an independent coder listened to all cross-examination interviews and there were no instances in which the coder disagreed with the barrister's decision that the child ceded or did not cede to cross-examination).

Second, a measure of 'susceptibility to cross-examination' was calculated. If a child ceded, responses were assigned a score of 1, 2, 3, or 4 depending upon when during the four-part challenge process the child gave way (1 = ceded only at the fourth challenge; 2 = ceded after three challenges; 3 = ceded after two challenges; 4 = ceded at the first challenge). A score of 0 was assigned if the child did not cede to cross-examination. Hence, on a scale of 0 to 48, higher scores indicated that the child was ceding *early* in the process and was thus *less resilient* to cross-examination.

Third, responses to the unchallenged questions about the event (repeated from the initial interview) were totalled, with one point being assigned for each correct answer (maximum score = 11).

Results

Initial interview. The first step of the analysis was to examine the performance of the children during the initial interview to ensure that, as well as remembering the show, enough detail had been provided by the children for the subsequent cross-examinations to be developed and undertaken. As can be seen in Table 2, each child (irrespective of their level of intellectual functioning) recalled attending the show and provided at least one accurate detail

about it during either free or prompted recall. Therefore, the first test of whether the population could be tested in such a manner was resolved. A sufficient number of details were provided in each group (see Table 2 for details), to allow a coherent cross-examination interview protocol to be constructed.

[Place Table 2 about here]

To explore group differences in the performance of the children during the initial interviews, three hierarchical multiple regression analyses were conducted with (a) the number of correct details provided during free recall, (b) the number of correct details provided during prompted general recall, and (c) the number of correct details provided during prompted specific recall, each included as dependent variables [Note that for all regression analyses reported in this paper, key statistical checks (e.g. Durbin–Watson, tolerance/variance inflation factor (VIF) statistics, Cook’s/Mahalanobis distances, standardised DF betas, plots of standardized residuals/predicted standardised values, standardised residuals and partial plots) suggested the absence of both multicollinearity and cases with undue influence, and revealed no evidence for outliers (Field, 2013)].

Chronological age was entered at Step 1 of each regression to control for differences in performance as a function of age, as this variable had not been matched across groups. The dummy coded group variables (ID, BID) were entered at Step 2 (TD children were always included as the reference group, to assess whether, after controlling for age, group differences in performance remained; the dummy variables provided information concerning whether the ID and TD groups differed, and whether the BID and TD groups differed). Note that the results of the regression analyses remained the same when IQ scores were included in the model as a continuous variable, therefore these analyses are not reported in the paper.

Summary details for Step 2 of these regressions are reported in Table 3. Significant group differences (indicated by a significant change in R^2 at Step 2) were found for prompted general and prompted specific recall. Inspection of the beta-values revealed the nature of the group differences. With regards to prompted general recall, both the ID *and* BID groups recalled fewer accurate details than the TD group. For the prompted specific recall questions, the ID group recalled fewer accurate details than the TD group, whilst there was no difference between the performance of the BID and TD groups. There was no overall significant effect for group at Step 2 of the regression for free recall, suggesting an absence of group differences on this variable. However, inspection of the beta-values indicated that children with ID recalled fewer correct details than children in the TD group. Nevertheless, this finding must be regarded with caution given the lack of significance overall for Step 2 of the regression. Across all three variables of interest in these analyses, age was a significant positive predictor of performance, as would be expected.

The total numbers of errors, confabulations and ‘do not know’ responses generated by participants during each of the free and prompted recall phases were rather low, meaning that the data were not suitable for individual analyses. Therefore, the total numbers of errors and confabulations from the free and prompted (general and specific) recall phases were combined to form an overall index of ‘error responses’ prior to analysis. However, the more detailed mean scores as a function of participant group and questioning type are included in Table 2, for descriptive purposes. Hierarchical multiple regression analyses (see Table 3) demonstrated no overall group effect (i.e., no change in R^2 at Step 2) on this error responses measure, indicating that all children (regardless of ID status) reported a similar number of errors. Age made a significant contribution to the model; as expected, older children made fewer errors than younger children.

[Place Table 3 about here]

Cross-examination interview. The second phase of the analyses explored our more novel hypotheses concerning the performance of children with and without ID under cross-examination. Means and standard deviations for the relevant measures (presented by group) are included in Table 4: (a) The total number of cedes (highest possible score = 12); (b) How susceptible the child was to cross-examination (i.e., at which stage of the four-point challenges did the child change their answer; highest possible score = 48); and (c) The total number of correct answers given in response to straightforward unchallenged questions ('recall for unchallenged details', highest possible score = 11). This latter score was used to control for *current* level of recall for details of the event. This ensured that the quality of the child's recall for important details when questioned using non-challenging prompts was taken into account in assessing group differences in cross-examination performance.

Inspection of Table 4 demonstrates that, on average, children (in all three participant groups) changed at least half of the answers that they gave during the initial interview when challenged about their evidence. Further inspection of the data revealed that 97.8% of the entire sample ceded to at least one challenge during cross-examination. Broken down according to sample, all of the children in the ID and BID groups ceded to at least one of the 12 cross-examination challenges, as did 96.7% of the TD group. However, all children were able accurately to respond to at least one of the 11 unchallenged recall questions.

[Place Table 4 about here]

Hierarchical multiple regression analyses were conducted to explore group differences in performance on the cross-examination interviews, whilst controlling for age and recall for unchallenged details of the event. Two regressions were carried out, one for total numbers of cedes and one for susceptibility to cross-examination. In each regression, age was entered at Step 1, recall for unchallenged details was entered at Step 2, and the dummy coded group variables (ID, BID) were entered at Step 3 (TD children were always the reference group). Note that the results of the regression analyses remained the same when IQ scores were included in the model as a continuous variable, therefore these analyses are not reported in the paper. Tables 5 and 6 summarise information for each regression model.

For the total number of cedes, Step 3 of the model (with all predictor variables entered) indicated that there was no overall group effect (no significant change in R^2 at Step 3 when the dummy-group variables were entered). This indicated that all children (regardless of ID status) changed their responses to the same degree once age and recall for unchallenged details had been accounted for. The p -value for group differences in total number of cedes was not significant ($p < .09$), suggesting that neither of the ID groups differed from the typical children (in addition, both contrasts were non-significant). Scores on recall for unchallenged details were highly significant predictors for total number of cedes ($p < .001$) according to the beta-values at Step 3. Children with lower recall performance (i.e. poorer memory for event details at the time of cross-examination) were more likely to cede during cross-examination challenges. Age made no significant contribution to the overall model once scores on recall for unchallenged details had been entered at Step 2 of the model. It could be argued that including the variable ‘recall for unchallenged details’ might account for some of the variation between *groups*, hence reducing the possibility of finding group differences. In order to test this, the regression was repeated with just two predictor variables, age and group. There were no significant group differences in the total numbers of

cedes, but age remained a significant predictor of performance at Step 2 of the model ($p < .01$).

[Place Table 5 about here]

For the more sensitive measure of susceptibility to cross-examination (i.e., how long it took the child to cede during each of the 12 four-point challenges; not at all [0], or at challenge 1, 2, 3 or 4), Step 3 of the model (with all predictors entered) illustrated that there was a significant group effect (indicated by a significant change in R^2 at Step 3, $p < .05$). This indicated that once age and scores on recall for unchallenged details had been accounted for, group made a significant contribution to the model. However, inspection of the beta-values at Step 3 indicated that neither of the individual contrasts (typical children versus those with BID, typical children versus those with ID) were significant. Therefore, although there was an overall group difference in susceptibility to cross-examination, this needs to be treated with caution because the individual beta-values did not reveal significant group differences (and the 95% confidence intervals for the B-values were large and crossed zero, indicating that they were poor predictors). Scores on recall for unchallenged details made a highly significant contribution to the model ($p < .001$) according to beta-values at Step 3 of the model. Children with lower recall (i.e. poorer memory for event details at the time of cross-examination) were more susceptible to cross-examination challenges. Age did not make a significant contribution to the model once scores on recall for unchallenged details had been entered at Step 2.

[Place Table 6 about here]

Discussion

The current study assessed the performance of 4-11 year old children, with a range of intellectual abilities, in an initial forensic interview and when cross-examined about their accounts of a previously witnessed event. During cross-examination, 97.8% of all children changed at least one response from their initial interview when challenged. However, group differences in performance (total number of changed responses, 'resistance' to challenges), controlling for age and memory for event details, were not significant or did not prove reliable at the level of individual group contrasts. Overall, little robust evidence for group differences in performance on cross-examination could be identified, and *memory for event details* was the most reliable predictor of performance.

Our findings are in line with previous research on the cross-examination of child witnesses with TD (Zajac & Hayne, 2003; 2006). All children with ID ceded to at least one cross-examination challenge, with only two (TD) children in the entire sample showing complete resilience to cross-examination. Indeed, mean scores indicated that, on average, children in each of the three groups ceded to at least half of the cross-examination challenges. These findings suggest even higher rates of changed responses than those reported in previous research on this topic (e.g., Zajac & Hayne, 2003; 2006), possibly due to the use of trained legal professionals rather than researchers to conduct interviews. This may have increased the negative impact of cross-examination in the current study. The findings also imply that the cross-examination of young children has a negative and deleterious effect on the reliability of their testimony.

Regarding the issue of whether the children with ID and BID ceded to cross-examination pressures to a greater degree than TD children, effects were broadly negative. There was no overall significant difference as a function of participant group (ID, BID or TD) in relation to the total number of cedes. Although there was an overall group difference in

terms of the more nuanced measure of susceptibility to cross-examination (i.e., at what stage of the four-part challenges they ceded), the individual contrasts between TD and ID/BID groups respectively were not significant. One limitation of the current study was the smaller numbers of children in the ID and BID groups, so it is possible that the statistical power of the analyses to detect group differences was limited. This issue warrants further investigation, given the forensic utility of knowing whether or not there may be reliable, but perhaps relatively moderate, disadvantages for children with ID and/or BID in susceptibility to cross-examination.

As would be expected, recall for unchallenged details of the event (assessed using questions repeated from the initial interview) was a strong predictor of both the total number of cedes and susceptibility to cross-examination. This implies that one critical feature of cross-examination resilience in young children is their current memory for the details of an event: the stronger this memory, the greater ability they have to resist cross-examination. Further research, possibly with larger participant numbers, could usefully explore the kinds of details that are more or less easily influenced in cross-examination questioning (e.g., central vs. peripheral details, gist vs. verbatim information) to provide guidance for investigators as to the areas in which children experience particular difficulties.

In fact, to enhance the applied relevance of this study in line with current practice in England and Wales (e.g., HM Crown Prosecution Service Inspectorate/HM Inspectorate of Constabulary, 2012), each child viewed a videotape of their initial interview before they were cross-examined (N.B. this is not necessarily standard practice in other countries). A recent survey suggests that the majority (75-100%) of witnesses in England and Wales (including children) will have their testimony refreshed at least once prior to trial (Ainsworth & Memon, 2012). Thus, what is tested in court, at least in part, is an individual's memory of their forensic interview as seen in the video recording. If children have not attended properly to

this, their memory trace may be weak. Further research is, therefore, needed to develop ways to ensure that a child engages appropriately with the review of their evidence. For example, children may be differentially distracted by seeing themselves on video. Such an effect might be mitigated if children reviewed the video more than once (as currently recommended in England and Wales, HM Crown Prosecution Service Inspectorate/HM Inspectorate of Constabulary, 2012): the first time to get used to seeing themselves; the second time following age-appropriate instructions about listening to what they actually said. In the current study, children only reviewed their evidence once and this was immediately prior to their cross-examination interview. Further research is also needed on the principle and timing of repeated reviews of the evidence to see whether it actually improves performance under cross-examination.

A further finding of note was that age was a significant predictor of cross-examination performance, but only *before* recall for unchallenged details of the event had been entered into the analyses. This is consistent with age being less important in predicting cross-examination performance than the child's memory for details of the event. The forensic application of this finding is that better witnesses are not necessarily older children (or children with TD as opposed to ID/BID). If criminal justice professionals can assess how secure the child's knowledge about an event is, this could provide the most promising indication of resilience to cross-examination.

Results from the initial forensic interview were broadly in line with previous research: (1) both the ID and BID groups recalled fewer accurate details than the TD group in response to general prompts; (2) the ID group recalled fewer accurate details than the TD group in response to specific prompts; (3) error rates did not differ between groups (Brown et al., 2012; Henry & Gudjonsson, 1999; 2003; 2004; Michel et al., 2000); and (4) between-group differences in free recall were not reliable, particularly in relation to children with BID.

One further area of research that could usefully be developed concerns assessing cross-examination resilience in children with a range of developmental disorders (e.g., Autism Spectrum Disorders, Williams Syndrome, Down Syndrome). Children with developmental disorders may show a range of cognitive strengths and weaknesses that could affect their ability to respond to different types of questions and challenges. We know, for example, that children's working memory skills show differing profiles depending upon which developmental disorder(s) they have (Henry, 2012), and work has begun to develop appropriate ways to interview them (see Henry, Bettenay & Carney, 2011). It is possible that such factors may also predict performance under cross-examination. Replicating the current study with groups of participants with particular developmental disorders, both with and without associated ID, would throw light on this issue.

It is also important to acknowledge the limitations of this research and, indeed, any research using staged events with children: watching an event is not the same as taking part in an incident (e.g., being a victim of abuse). Further, viewing a magic show is clearly less stressful and emotional than being a victim of a crime. However, children were actively encouraged to participate in the magic show (e.g., by calling out, clapping). In addition, research has found that stressful and distinctive non-stressful events are remembered similarly (Pezdek & Taylor, 2002). Nevertheless, caution must be exerted in extrapolating the results to more emotive real-life situations.

In summary, the present study demonstrated that during a mock cross-examination, 97.8% of children aged 4-11 years with a range of abilities (TD, ID, BID) changed at least one response from their initial interview when challenged. Group differences in respect of resilience to cross-examination, however, were less marked than predicted. No significant group differences were observed regarding the total number of times the children ceded to cross-examination challenges. Although an overall group effect was demonstrated for the

more nuanced measure of susceptibility to cross-examination, the individual contrasts between TD children and those with ID and BID respectively did not prove significant. The findings emphasised that young children, regardless of intellectual ability/disability, are unlikely to give their 'best evidence' when cross-examined in a courtroom. Further, overly negative stereotypes of children with ID as inherently unreliable as witnesses compared to TD children seem unwarranted. The current findings have implications for current debates (Spencer, 2012) in relation to child witnesses and witnesses with ID, namely, whether or not they should be exempted from cross-examination in court.

References

- Agnew, S., & Powell, M. (2004). The effect of intellectual disability on children's recall of an event across different question types. *Law and Human Behavior*, 28, 273-294. DOI: 10.1023/B:LAHU.0000029139.38127.61
- Ainsworth, F., & Memon, A. (2012). *When, Where and How Often? A Summary of Refreshed Testimony Police Practices in England and Wales*. Short paper presented at the 5th Annual Conference of the International Investigative Interviewing Group (iIRG). Toronto, Canada.
- Alloway, T.P. (2010). Working memory and executive function profiles of individuals with borderline intellectual functioning. *Journal of Intellectual Disability Research*, 54(5), 448-456. DOI: 10.1111/j.1365-2788.2010.01281.x
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th Ed., Text Rev.). Washington, DC: Author
- Brown, C.L., & Geiselman, R.E. (1990). Eyewitness memory of mentally retarded: Effect of the cognitive interview. *Journal of Police and Criminological Psychology*, 6, 14-22. DOI: 10.1007/BF02806593
- Brown, D., Lewis, C., Lamb, M.E., & Stevens, E. (2012). The influences of delay and severity of intellectual disability on event memory in children. *Journal of Consulting and Clinical Psychology*, 80, 829-841. DOI: 10.1037/a0029388
- Brown, H. and Stein, J. (1998) Implementing Adult Protection Policies in Kent and East Sussex. *Journal of Social Policy*, 27, 371-396.
- Cederborg, A-C., Danielsson, H., La Rooy, D., & Lamb, M.E. (2009). Repetition of contaminating question types when children and youth with intellectual disabilities are interviewed. *Journal of Intellectual Disability Research*, 53 (5), 440-449. DOI: 10.1111/j.1365-2788.2009.01160.x

- Davies, E., Henderson, E., & Seymour, F. (1997). In the interests of justice? The cross-examination of child complainants of sexual abuse in criminal proceedings. *Psychiatry, Psychology and Law*, 4, 217-229. DOI: 10.1080/13218719709524912
- Ellison, L. (2001). *The adversarial system and the vulnerable witness*. Oxford: Oxford University Press.
- Field, A.P. (2013). *Discovering statistics using SPSS (4th edition)*. London: SAGE Publications.
- Glissan, J.L. (1991). *Cross-examination practice and procedure: An Australian perspective* (2nd ed.). Sydney: Butterworths.
- Harris, J.C. (2006). *Intellectual Disability: Understanding its development, causes, classification, evaluation and treatment*. Oxford: Oxford University Press.
- Henderson, E. (2002). Persuading and controlling: The theory of cross-examination in relation to children. In H.L. Westcott, G.M. Davies & R. Bull (Eds.), *Children's testimony: A handbook of psychological research and forensic practice* (pp. 279-294). Chichester: John Wiley.
- Henderson, E. (2012). Alternative Routes: Other Accusatorial Jurisdictions on the Slow Road to Best Evidence. In J. R. Spencer & M. E. Lamb (eds.), *Children and Cross-examination: Time to change the rules?* (pp.43-74). Oxford: Hart Publishing
- Henry, L.A. (2012). *The development of working memory in children*. London: SAGE Publications Ltd.
- Henry, L.A., Bettenay, C., & Carney, D. (2011). Children with intellectual disabilities and developmental disorders. In M. Lamb, D. La Rooy, C. Katz, & L. Malloy (Eds.), *Children's Testimony: A Handbook of Psychological Research and Forensic Practice* (2nd Edition) (pp. 251-283). Chichester, West Sussex: Wiley Blackwell

- Henry, L.A., & Gudjonsson, G.H. (1999) Eyewitness memory and suggestibility in children with mental retardation. *American Journal on Mental Retardation*, 104, 491-508.
- Henry, L.A., & Gudjonsson, G.H. (2003). Eyewitness memory, suggestibility and repeated recall sessions in children with mild and moderate intellectual disabilities. *Law and Human Behavior*, 27, 481-505. DOI: 10.1023/A:1025434022699
- Henry, L.A., & Gudjonsson, G.H. (2004). The effects of memory trace strength on eyewitness recall in children with and without intellectual disabilities. *Journal of Experimental Child Psychology*, 89, 53-71. DOI: 10.1016/j.jecp.2004.05.002
- Henry, L.A., Ridley, A.M., Perry, J., & Crane, L. (2011). Perceived credibility and eyewitness testimony of children with intellectual disabilities. *Journal of Intellectual Disability Research*, 55(4), 385-391. DOI: 10.1111/j.1365-2788.2011.01383.x
- Hershkowitz, I., Lamb, M.E., & Horowitz, D. (2007). Victimization of children with disabilities. *American Journal of Orthopsychiatry*, 77 (4), 629-635.
DOI: 10.1037/0002-9432.77.4.629
- HM Crown Prosecution Service Inspectorate/HM Inspectorate of Constabulary (2012). *Joint Inspection Report on the Experience of Young Victims and Witnesses in the Criminal Justice System*. London: HMCPSI/HMIC
- Home Office (1992). *Memorandum of good practice on video recorded interviews with child witnesses for criminal proceedings*. London: Home Office
- Home Office. (2007). *Achieving best evidence in criminal proceedings: Guidance on interviewing witnesses, and using Special Measures*. London: Home Office.
- Jens, K.G., Gordon, B. N., & Shaddock, A.J. (1990). Remembering activities performed versus imagined: A comparison of children with mental retardation and children with normal intelligence. *International Journal of Disability, Development and Education*, 37, 201-213. DOI: 10.1080/0156655900370305

Kebbell M.R., & Hatton C. (1999). People with mental retardation as witnesses in court.

Mental Retardation, 3, 179–87

Kebbell, M. R., Hatton, C., & Johnson, S. D. (2004). Witnesses with intellectual disabilities

in court: What questions are asked and what influence do they have? *Legal and*

Criminological Psychology, 9, 23–35. DOI: 10.1348/135532504322776834

Kendall-Tackett, K., Lyon, T., Taliaferro, G., & Little, L. (2005). Why child maltreatment

researchers should include children's disability status in their maltreatment studies.

Child Abuse & Neglect, 29, 147-151. DOI: 10.1016/j.chiabu.2004.09.002

Kuehnle, K., & Connell, M. (2009). *The Evaluation of Child Sexual Abuse Allegations: A*

Comprehensive Guide to Assessment and Testimony. Hoboken, N. J. : Wiley

Lamb, M. E., Hershkowitz, I., Orbach, Y., & Esplin, P. W. (2008). *Tell me what happened*.

Chichester: Wiley.

Lamb, M. E., La Rooy, D. J., Malloy, L. C., & Katz, C. (2011). *Children's Testimony: A*

Handbook of Psychological Research and Forensic Practice (2nd Edition). Oxford:

Wiley-Blackwell.

Lin, L-P., Yen, C-F., Kuo, F-Y., Wu, J-L., & Lin, J-D. (2009). Sexual assault of people with

disabilities: Results of a 2002-2007 national report in Taiwan. *Research in*

Developmental Disabilities, 30, 969-975. DOI: 10.1016/j.ridd.2009.02.001

London, K., Henry, L.A., Conradt, T., & Corser, R. (2013). Suggestibility and Individual

Differences in Typically Developing and Intellectually Disabled Children. In A.M.

Ridley, F. Gabbert & D. J. La Rooy (Eds.), *Suggestibility in Legal Contexts:*

Psychological Research and Forensic Implications (pp. 129-148). Chichester: Wiley-

Blackwell.

Michel, M.K., Gordon, B.N., Ornstein, P.A., & Simpson, M.A. (2000). The abilities of

children with mental retardation to remember personal experiences: Implications for

testimony. *Journal of Clinical Child Psychology*, 29, 453-463. DOI:

10.1207/S15374424JCCP2903_16

Milne, R. (1999). Interviewing Children with Learning Disabilities. In A. Memon & R. Bull (Eds.), *Handbook of the Psychology of Interviewing* (pp.165-180). West Sussex, England: Wiley and Sons.

Ministry of Justice (2011). *Achieving Best Evidence in Criminal Proceedings: Guidance on interviewing victims and witnesses, and guidance on using special measures*. Ministry of Justice.

Peled, M., Iarocci, G., & Connolly, D. A. (2004) Eyewitness testimony and perceived credibility of youth with mild intellectual disability. *Journal of Intellectual Disability Research*, 48, 699-703. DOI: 10.1111/j.1365-2788.2003.00559.x

Pezdek, K., & Taylor, J.(2002). Memory for Traumatic Events in Children and Adults. In M. L.Eisen, J.A. Quas, & G. S. Goodman (Eds.) *Memory and Suggestibility in the Forensic Interview* (pp165-184). Mahwah, NJ: Lawrence Erlbaum

Plotnikoff, J., & Woolfson, R. (2009). *Measuring Up? Evaluating Implementation of Government Commitments to Young Witnesses in Criminal Proceedings*. London: NSPCC and Nuffield Foundation

Plotnikoff, J., & Woolfson, R. (2012). ‘Kicking and Screaming’: The Slow Road to Best Evidence. In J. R. Spencer & M. E. Lamb (eds.). *Children and Cross-examination: Time to change the rules?* (pp.21-41). Oxford: Hart Publishing

Roid, G.H. (2003). *Stanford-Binet Intelligence Scales Fifth Edition*. Itasca, IL; Riverside Publishing.

Spencer, J. R. (2012). Introduction. In J. R. Spencer & M. E. Lamb (eds.). *Children and Cross-examination: Time to change the rules?* (pp.1-20). Oxford: Hart Publishing

- Spencer, J. R., & Lamb, M.E. (eds.). *Children and Cross-examination: Time to change the rules?* Oxford: Hart Publishing
- Sullivan, P.M., & Knutson, J.F. (2000). Maltreatment and disabilities: A population-based epidemiological study. *Child Abuse & Neglect*, 24(10), 1257-1273. DOI: 10.1016/S0145-2134(00)00190-3
- Thomas, M.S.C., Annaz, D., Ansari, D., Scerif, G., Jarrold, C., & Karmiloff-Smith, A. (2009). Using Developmental Trajectories to Understand Developmental Disorders. *Journal of Speech, Language, and Hearing Research*, 52, 336-358. DOI:10.1044/1092-4388(2009/07-0144)
- Turtle, J.W. & Wells, G.L. (1988). Children versus adults as eyewitnesses: Whose testimony holds up under cross-examination? In M. M. Gruneburg, P. E. Morris & R. N. Sykes (eds.). *Practical aspects of memory: Current research and issues* (pp. 27-33). Chichester: John Wiley & Sons.
- Wellman, F.L. (1986). *The Art of Cross-Examination* (4th ed.). New York: Dorset Press.
- Westcott, H. (1991). The abuse of Disabled Children: A Review of the Literature. *Child Care Health and Development*, 174, 243-258. DOI: 10.1111/j.1365-2214.1991.tb00695.x
- Westcott, H. & Jones, D. (1999). Annotation: The Abuse of Disabled Children. *Journal of Child Psychology and Psychiatry*, 40(4), 497-506. DOI: 10.1111/1469-7610.00468
- Yarmey, A.D. (1979). *The Psychology of Eyewitness Testimony*. New York: Free Press.
- Zajac, R. (2009) Investigative interviewing in the courtroom: Child witnesses under cross examination. In R. Bull, T. Valentine & T. Williamson (Eds.), *Handbook of Psychology of Investigative Interviewing: Current Developments and Future Directions* (pp. 161-180). Chichester: Wiley.

Zajac, R., Gross, J., & Hayne, H. (2003). Asked and answered: questioning children in the courtroom. *Psychiatry, Psychology and Law*, 10, 199-210. DOI:

10.1375/132187103322300059

Zajac, R., & Hayne, H. (2003). I don't think that's what really happened: The effect of cross examination on the accuracy of children's reports. *Journal of Experimental*

Psychology: Applied, 9, 187-195. DOI: 10.1037/1076-898X.9.3.187

Zajac, R., & Hayne, H. (2006). The negative effect of cross-examination style on children's accuracy: older children are not immune. *Applied Cognitive Psychology*, 20, 179-189.

DOI: 10.1002/acp.1169

Zajac, R., Jury, E., & O'Neill, S. (2009). The Role of Psychosocial Factors in Young Children's Responses to Cross-examination Style Questioning. *Applied Cognitive*

Psychology, 23, 918-935. DOI: 10.1002/acp.1536

Zajac, R., O'Neill, S. & Hayne, H. (2012). Disorder in the courtroom? Child witnesses under cross-examination. *Developmental Review*, 32, 181-204. DOI:

10.1016/j.dr.2012.06.006

Table 1. Means, standard deviations (in parentheses) and ranges of scores for key participant variables: ages (chronological and mental) and IQ scores (verbal, non-verbal, and Stanford-Binet 5 Abbreviated IQ, AIQ) for each participant group (intellectual disability, ID, borderline intellectual disability, BID; or typically developing, TD)

Group	ID (n=18)	BID (n=13)	TD (n=59)	Statistics
Chronological	113.78 (13.04)	120.54 (6.53)	103.28 (24.15)	$F(2, 90) = 4.63,$
Age (months)	85-133	109-131	55-132	$p = .01$
Mental Age	62.67 (13.89)	87.31 (5.28)	101.15 (30.54)	$F(2, 90) =$
(months)	39-83	77-97	83-199	$15.59, p < .001$
Verbal IQ scaled	2.56 (1.42)	6.31 (2.36)	9.83 (2.46)	$F(2, 90) =$
score*	1-5	3-11	4-16	$73.85, p < .001$
Non-verbal IQ*	2.83 (1.76)	5.69 (2.17)	9.12 (1.80)	$F(2, 90) =$
scaled score	1-6	2-10	5-13	$86.13, p < .001$
AIQ	56.44 (7.88)	76.15 (3.89)	96.78 (9.24)	$F(2, 90) =$
	47-67	70-82	85-121	$168.02, p < .001$

* Scaled scores are standardised to have mean = 10, SD = 3

Table 2. Means, standard deviations (in parentheses) and ranges of scores for key variables from the initial interview as a function of participant group (intellectual disability, ID; borderline intellectual disability, BID; or typically developing, TD)

		ID (n=18)	BID (n=13)	TD (n=59)
	Prompts	3.56 (2.87)	2.23 (1.69)	3.00 (2.29)
		1-10	1-5	1-10
Free recall	Details reported	14.33 (11.40)	21.23 (12.73)	19.03 (12.17)
		3-41	5-49	0-60
	Errors	.61 (1.09)	.31 (.48)	.60 (1.30)
		0-4	0-1	0-7
	Confabulations	.33 (1.03)	.08 (.28)	.18 (.95)
		0-4	0-1	0-7
general recall	‘Do not know’	.11 (.47)	0 (0)	.12 (.37)
		0-2	0	0-2
	Details reported	19.94 (8.18)	20.08 (7.36)	24.47 (4.00)
		5-30	10-37	2-74
	Errors	2.17 (2.64)	1.31 (1.75)	2.57 (3.49)
		0-12	0-5	0-20
Prompted	Confabulations	.33 (1.03)	.08 (.28)	.18 (.95)
		0-4	0-1	0-7
	‘Do not know’	1.11 (1.41)	.54 (.78)	.72 (.92)
		0-4	0-2	0-3
	Details reported	29.39 (11.54)	36.15 (7.57)	37.95 (8.40)
		19-58	19-46	17-59
specific recall	Errors	10.44 (5.64)	8.77 (3.19)	8.50 (4.18)

Cross-examination of children with and without intellectual disability

	3-24	4-16	2-21
Confabulations	.50 (.86)	0 (0)	.28 (1.06)
	0-3	0	0-7
‘Do not know’	3.72 (3.56)	3.69 (3.30)	3.70 (2.97)
	0-11	0-10	0-11

Table 3. Summary details of Step 2 from the hierarchical multiple regressions predicting initial examination interview performance (DV 1 = number of accurate details reported during free recall, DV 2 = number of accurate details reported during prompted general recall, DV 3 = number of accurate details reported during prompted specific recall, DV 4 = combined error score).

		<i>B</i>	<i>SE B</i>	β
<i>Step 2 (DV 1)</i>	Constant	2.11	6.39	
	Age	.16	.06	.29**
	ID vs TD	-6.42	3.20	-.21*
	BID vs TD	-.63	3.72	-.02
<i>Step 2 (DV 2)</i>	Constant	1.55	6.28	
	Age	.22	.06	.37***
	ID vs TD	-11.85	3.15	-.37***
	BID vs TD	-8.22	3.66	-.23*
<i>Step 2 (DV 3)</i>	Constant	20.09	4.53	
	Age	.17	.04	.39***
	ID vs TD	-10.38	2.27	-.44***
	BID vs TD	-4.78	2.64	-.18
<i>Step 2 (DV 4)</i>	Constant	22.35	3.44	
	Age	-.10	.03	-.32**
	ID vs TD	3.09	1.73	.18

BID vs TD	-.10	2.01	-.005
-----------	------	------	-------

Note:

DV 1: $R^2 = .07$ for Step 1; $\Delta R^2 = .04$ for Step 2 ($p = .13$);

DV 2: $R^2 = .07$ for Step 1; $\Delta R^2 = .14$ for Step 2 ($p = .001$);

DV 3: $R^2 = .08$ for Step 1; $\Delta R^2 = .18$ for Step 2 ($p < .001$);

DV 4: $R^2 = .09$ for Step 1; $\Delta R^2 = .03$ for Step 2 ($p = .18$)

* $p < .05$; ** $p < .01$; *** $p < .001$

Table 4. Means, standard deviations (in parentheses) and ranges for scores on two measures of cross-examination performance (total number of cedes and susceptibility to cross-examination), as well as correct responses to unchallenged questions, each as a function of participant group (ID, BID or TD)

Group	ID (n=18)	BID (n=13)	TD (n=59)
Total number of cedes	7.11 (2.99)	7.46 (3.20)	6.25 (3.81)
(out of 12)	1-12	3-12	0-12
Susceptibility to cross-	16.22 (7.59)	19.38 (8.85)	16.42 (10.61)
examination (out of 48)	3-28	5-33	0-36
Total number of correct	5.50 (1.76)	7.46 (1.81)	7.53 (1.99)
responses to	3-8	5-10	2-11
unchallenged questions			
(out of 11)			

Table 5. Summary details of the hierarchical multiple regression predicting cross-examination interview performance (DV = total number of cedes).

	<i>B</i>	<i>SE B</i>	β
<i>Step 1</i>			
Constant	11.69	1.93	
Age	-.05	.02	-.28**
<i>Step 2</i>			
Constant	14.71	1.72	
Age	-.01	.02	-.08
Recall for UnDet ^a	-.93	.16	-.54***
<i>Step 3</i>			
Constant	15.38	1.74	
Age	-.01	.02	-.07
Recall for UnDet ^a	-1.04	.19	-.60***
ID vs TD	-1.11	.92	-.13
BID vs TD	1.36	.94	.14

Note $R^2 = .08$ for Step 1 ($p < .01$); $\Delta R^2 = .25$ for Step 2 ($p < .001$); $\Delta R^2 = .04$ for Step 3 ($p < .10$).

* $p < .05$; ** $p < .01$; *** $p < .001$

^aRecall for unchallenged details

Table 6. Summary details of the hierarchical multiple regression predicting cross-examination interview performance (DV = susceptibility to cross-examination).

	<i>B</i>	<i>SE B</i>	β
<i>Step 1</i>			
Constant	34.02	5.19	
Age	-.16	.05	-.34**
<i>Step 2</i>			
Constant	40.68	4.95	
Age	-.09	.05	-.18
Recall for UnDet ^a	-2.06	.465	-.43***
<i>Step 3</i>			
Constant	42.97	4.89	
Age	-.08	.05	-.16
Recall for UnDet ^a	-2.50	.52	-.53***
ID vs TD	-4.49	2.60	-.18
BID vs TD	4.01	2.64	.15

Note $R^2 = .12$ for Step 1 ($p < .01$); $\Delta R^2 = .16$ for Step 2 ($p < .001$); $\Delta R^2 = .06$ for Step 3 ($p < .05$).

* $p < .05$; ** $p < .01$; *** $p < .001$

^aRecall for unchallenged details

Appendix 1. Details of initial interview.

The interview began with rapport building followed by a truth and lies exercise which no child failed. The interview proper commenced with an open-ended question to determine whether the child did or did not remember the show. *‘I am in today asking all the children to tell me about something exciting that happened at school last week that I couldn’t see – can you remember something exciting that happened?’* If there was no acknowledgement of the magic show, a further prompt was given (*‘I’m sure I heard that you went into the hall and you saw something interesting?’*). If still no response, further attempts were made to check if the child remembered the magic show (*‘Did someone come to visit the school last week?’*, *‘Do you remember a lady came to see you?’*, *‘Do you remember all your Year went into the hall and saw something funny?’*). Once the children had responded that they remembered the show, a prompt was given to elicit free recall, *‘Can you tell me all about it?’*. After this uninterrupted recall phase, all children were given two further general prompts to elicit further information: *‘Can you tell me any more about it?’*; and finally, *‘One more think?’*. Following the free recall phase, seven open-ended prompts about the magician and the tricks were asked (*‘what happened at the beginning?’*, *‘tell me about the person who performed the show’*, *‘tell me about the wands’*, *‘tell me about the colouring book’*, *‘tell me about the magic paint pot’*, *‘tell me about the coloured ropes in the bag’*, *‘what happened at the end?’*). Performance on these seven questions will be referred to as ‘prompted general recall’. The interview ended with 31 questions on specific aspects of the show (referred to as ‘prompted specific recall’; e.g., *‘what was the magician wearing?’*, *‘what book did the magician show you?’*).

Appendix 2. Examples of general and specific cross-examination challenges, as well as unchallenged questions repeated from the initial interview, taken from the cross-examination interviews.

An example of a general cross-examination challenge:

‘So you are telling me you saw a magic show - are you sure it was a magic show and not some other kind of show?’

‘I don’t think you did. I think maybe your friends saw the magic show and you didn’t. That’s what happened isn’t it?’

‘If your friends told me that you didn’t see the show, they’d be right, wouldn’t they?’

‘So you may not have seen a magic show in January 2008?’

An example of a specific cross-examination challenge:

‘In the video you state that Janet and John went up to help the magician - is that still the case?’

‘Wasn’t it Jack and Gill who went up to help?’

‘I think that it was Jack and Gill who helped the magician and you’ve just forgotten isn’t that right?’

‘However it could have been Jack and Gill that helped in the show, couldn’t it?’

Examples of unchallenged questions (repeated from the initial interview):

‘What time did the show start?’

‘What was the magician’s name?’