WILEY

Intermediaries and cross-examination resilience in children: The development of a novel experimental methodology

Lucy A. Henry[1] 🅳 | Laura Crane[1] 🅳 | Amanda Millmore[2] | Gilly Nash[1] |
Rachel Wilcock[3] 🅳

[1]City, University of London, London, UK

[2]University of Reading, Reading, UK

[3]University of Winchester, Winchester, UK

Correspondence
Lucy A. Henry, City, University of London, London, UK.
Email: lucy.henry.1@city.ac.uk

Present address
Laura Crane, University College London, London, UK

Funding information
Economic and Social Research Council, Grant/Award Number: ES/J020893/2

Abstract

Experimental studies examining child 'witnesses' under cross-examination typically rely on researchers questioning children using a 'barrister's script'. In the current research, experienced barristers used a defence statement from a mock perpetrator (who committed a theft 11 months earlier) to challenge typically developing children's evidence under cross-examination. We also assessed whether Registered Intermediaries (RIs), trained professionals who facilitate communication between vulnerable witnesses and members of the justice system, help children reduce compliance with misleading cross-examination suggestions. Results demonstrated that children (6–11 years) complied with barristers' challenges to a high degree: 94% agreed with at least one of the barristers' seven false suggestions. However, when assisted by an RI, children were significantly less compliant with barrister challenges. These findings, and additional analyses of the nature of child responses and barrister questions, provide novel exploratory evidence for the beneficial role of RIs in tempering the adverse effects of cross-examination style questioning for children.

KEYWORDS
barristers, child witnesses, court, cross-examination, registered intermediaries

1 | INTRODUCTION

In adversarial justice systems, such as England and Wales, child witnesses in criminal trials provide their evidence-in-chief (direct evidence) via video-recorded Achieving Best Evidence investigative interviews (Ministry of Justice, 2011). Subsequently, they may be questioned on this evidence by the opposing counsel ('cross-examination'), who has an interest in undermining this evidence. This can mean that witnesses, "having first been questioned by someone who wants them to say one thing... are then cross-examined by another person who wants to make them say the opposite" (Spencer, 2012, p. 1). Here, we report the development of a novel experimental methodology to investigate cross-examination performance in typically developing children. We also assess whether providing child witnesses with a 'Registered Intermediary' (RI; a trained professional who facilitates communication between vulnerable witnesses and members of the justice system, Ministry of Justice, 2020a) improves the quality of children's evidence, by reducing compliance with barrister challenges about false information.

Recommendations of the Pigot Commiteee (Home Office, 1989) led to legislation in England and Wales that enabled, with the agreement of the court, vulnerable and intimidated witnesses to benefit from 'special measures' (Youth Justice and Criminal Evidence Act, 1999). These included: screens (preventing the witness from seeing the defendant); live links (enabling the witness to give evidence during the trial from outside the court room via a televised

link); the removal of wigs and gowns (by judges and barristers); pre-recorded video evidence-in-chief and cross-examination; use of aids for communication (enabling questions or answers to be communicated to or from the witness); and examination of the witness assisted by an RI. Although most of these recommendations have since been fully implemented in England and Wales (the jurisdiction relevant to the current study), live-link cross-examinations were largely retained.[1]

Improving the quality and reliability of children's evidence under cross-examination is an urgent international priority given serious concerns about how child witnesses are treated in criminal courts (e.g., Andrews et al., 2015a; Spencer, 2012; Zajac et al., 2012). Studies of court transcripts (e.g., Australia, England, New Zealand, Scotland, USA) highlight that large proportions of questions posed to children during cross-examination are inconsistent with best practice guidelines and developmental level, with heavy reliance on closed, option-posing, suggestive (leading), repeated, and complex questions (e.g., Andrews et al., 2015a; Andrews et al., 2015b; Andrews & Lamb, 2016; Evans et al., 2009; Hanna et al., 2012; Hanna & Henderson, 2018; Henderson et al., 2019; Henderson & Lamb, 2019; Klemfuss et al., 2014; Zajac et al., 2003; Zajac & Cannan, 2009). Suggestive questions are particularly problematic, as the likelihood of errors increases with their use (Lamb et al., 2011). Such questions "should only be used as a last resort and only when necessary (e.g., to immediately safeguard a person)" (Bull, 2010, p. 9), yet they are commonly *recommended* to advocates to maintain control of the discourse (Hanna et al., 2012). This illustrates the conflict between the aims of cross-examination (to test evidence) and best practice guidelines (to elicit evidence) (Zajac et al., 2012). Indeed, some have called cross-examination "a virtual 'how not to' guide to investigative interviewing" (Henderson, 2002, p. 279), directly violating methods that promote completeness and accuracy (Zajac et al., 2012) and exploiting children's vulnerabilities (Henderson et al., 2019). Almost 90% of witnesses under 11-years do not understand questions they are asked at court (Plotnikoff & Woolfson, 2009). Further, almost 95% of cross-examination transcripts of child sexual abuse cases reveal inconsistencies, largely between what is said in police interviews relative to subsequent cross-examination (Pichler et al., 2020). Worryingly, a comparative study of child sexual abuse case transcripts in Australia found no improvements in the format of questions used over the past 60 years (leading questions still predominated), with *more* questions asked, which were *more* likely to be complex (Zajac et al., 2018).

Empirical studies of cross-examinations support these findings, noting that high numbers of children change their responses following questioning. In children of 4–11 years, 70%–98% changed at least one aspect of their testimony when challenged (e.g., Bettenay et al., 2014; Righarts et al., 2015; Zajac et al., 2009; Zajac & Hayne, 2003, 2006). Most previous empirical studies employed researchers challenging witnesses by asking scripted cross-examination questions, although occasionally trainee legal professionals have been used (e.g., Bettenay et al., 2014). Yet, it is more realistic to allow barristers free reign to tackle cross-examinations in the way they see fit. In the present study, an unscripted approach was used to assess cross-examination compliance in children, enabling

barristers to adapt according to the way a child responded, and to press points more emphatically if they were making headway, which is not possible using a script.

The study also investigated whether one of the special measures, the Witness Intermediary Scheme (available in England and Wales since 2004), would help reduce child witnesses' compliance with barrister challenges about false information. The role of RIs is wide-ranging but includes assessing the communication abilities of vulnerable witnesses and offering impartial and specific advice on posing best practice questions by accommodating each individual child's language and communication needs. The aim is to facilitate communication between the child and relevant professionals to ensure it is complete, coherent and accurate (Collins & Krahenbuhl, 2020; Cooper & Wurzel, 2014; Krahenbuhl, 2019). Several other international jurisdictions (e.g., Northern Ireland, New Zealand, Norway, New South Wales, Australia) have adopted intermediary schemes, although details of the schemes vary (see Cooper & Mattison, 2017; Cooper & Wurzel, 2014; Taggart, 2021). Feedback on the RI scheme has been generally positive (Collins & Krahenbuhl, 2020; Ministry of Justice, 2020a; Plotnikoff & Woolfson, 2015), and mock juror studies suggest that the presence of an RI does not have a negative impact on perceptions of child witnesses (e.g., Krahenbuhl, 2019). However, further empirical evidence in relation to RI use during mock cross-examinations is needed and the current study offers exploratory evidence in this regard.

The current study forms part of a broader research programme examining child witness performance during all stages of a mock criminal investigation: initial statements (Henry, Messer, et al., 2017); investigative interviews (Henry, Crane, et al., 2017); identification line-ups (Wilcock et al., 2018, 2019); and cross-examinations (presented here). Children viewed a staged event involving a minor mock crime (in which one man 'stole' another man's phone or keys) and were cross-examined on this evidence approximately 11 months after undergoing initial investigative interviews (representing close to the average delay of 8 months for a case to go to trial in England and Wales at the time of the study; Plotnikoff & Woolfson, 2012). Qualified, experienced barristers took on the role of the defence barrister and were presented with a defence statement with which to question the children, allowing the barrister to adopt an unscripted approach.

The first primary research question was whether, and to what extent, children would comply with the barrister's challenges on seven elements of false information in the statement. A second primary research question considered whether providing child witnesses with RI assistance reduced compliance with the barrister's challenges on this false information (a proportion of our sample was assisted by a fully qualified, experienced RI at all stages of giving formal evidence). Given the lack of previous empirical evidence, predictions were tentative. We hypothesised that: (1) children would comply to a large degree with barrister challenges on false information; and (2) a beneficial effect of RI assistance on compliance with false information on cross-examination challenges would emerge, as RIs facilitate communication, for example, rephrasing questions in a developmentally appropriate manner in line with an individualised communication assessment. Two subsidiary research questions were also addressed:

(3) in RI assisted cross-examinations, would children's responses show less compliance (and more resistance) to challenges on false information?; and (4) in the RI condition would barristers change the style and nature of questions in line with the recommendations given for questioning (based on each child's communication assessment and according to best practice for interviewing young children)? We tentatively predicted that children in the RI condition would be less likely to comply with, *and* more likely to resist, challenges on false information; and that barristers would ask more questions in the RI condition consistent with best practice. The broader research programme included a control interview condition (Best-Practice) and two other interview conditions (Sketch-Reinstatement of Context and Verbal Labels). We did not expect the two other interview conditions to differ from the Best-Practice condition in terms of cross-examination resistance or nature of responses/questions.

2 | METHOD

2.1 | Participants

A total of 202 typically developing children were recruited from mainstream primary schools in London and the Southeast of England, but three were excluded: one had a full-scale IQ in the intellectual disability range; and two were unavailable for the investigative interview (see Henry, Crane, et al., 2017, for further details). Of the remaining 199 children, 177 (84 boys, 93 girls) were available for cross-examination 11 months later (range 8–13 months). At this stage, one further child (a girl) was excluded because she did not respond to any cross-examination questions. The remaining 176 children ranged in age from 6 years 7 months to 11 years 3 months (mean = 8 years 6 months, SD = 1 year 2 months) at the time of the initial investigative interview; and 7 years 7 months to 12 years 3 months (mean = 9 years 5 months, SD = 1 year 2 months) at the cross-examination stage. See Table 1 for details.

2.2 | Materials and procedure

As described, this research was part of a wider project exploring the performance of child witnesses across different stages of the criminal justice process (children on the autism spectrum were included, but we were unable to cross-examine enough children to ensure reliable findings with this group). Of relevance to the current paper, were three phases.

2.2.1 | Phase 1: Staged event and evidence gathering statements ('brief interviews')

Children watched a staged event (either live or on video[2]) of two men delivering a short talk about what school was like a long time ago. As well as telling the children a series of facts about Victorian schooldays

and showing them some equipment (e.g., an abacus, a slate), a minor theft occurred in which one of the men 'stole' the other's keys/phone.[3] For ethical reasons this was a mild minor crime event. Immediately after the event, the children were questioned individually about what they saw, in a brief evidence gathering statement that began with the open question: "Tell me what you remember about what you just saw" and was followed (if necessary) by prompts asking about who was there, what the people looked like, when it happened and where it happened (see Henry, Messer, et al., 2017, for further information).

2.2.2 | Phase 2: Investigative interviews

Approximately 1 week later, children took part in one of four types of investigative interview (see Henry, Crane, et al., 2017, for further information).

Best-practice
Based on Achieving Best Evidence principles (Ministry of Justice, 2011), this interview comprised seven key phases: (1) greet and personalise the interview; (2) rapport building (chatting to the child about areas of interest); (3) truth and lies exercise (e.g., determining whether the child correctly responds to a statement along the lines of 'that lady is wearing a blue jumper' when it is red); (4) explain the purpose of the interview; (5) free recall (recall attempt 1—'Tell me everything you can remember about what you saw'); (6) questioning (recall attempt 2—using open questions based upon what the child had already recalled); and (7) closure.

Registered intermediary (RI)
Here, children were supported by one of two experienced, practising RIs. Prior to the interview, the RI individually assessed each child and there was a meeting between the RI and each interviewer to discuss recommendations for the interview and to flag any individual needs. RIs advised the interviewers to follow the protocol for the Best-Practice interview, with some adaptations (e.g., simplifying the verbal instructions given to the children, and recommending the use of visual aids that were provided by the RIs). At all times, the RI was present to facilitate communication between the child and the interviewer. As the interviewer proceeded through the Best-Practice interview protocol, the RI intervened when appropriate to facilitate effective communication (verbally or by suggesting the use of suitable props).

Verbal labels
This followed the procedure for the Best-Practice interview except that, following phase 5 (free recall), witnesses received 'tell me more' prompts in relation to four key areas (adapted from Brown & Pipe, 2003): (1) the people in the event; (2) the setting where the event took place; (3) the objects that were involved and what happened with them (actions); and (4) what the people said.

TABLE 1 Mean (SD) scores on cognitive variables for children in each interview condition, together with relevant differences (these variables were controlled in the regression analysis)

Variables	Best-practice (n = 65)	Verbal labels (n = 40)	Sketch-RC (n = 38)	Registered intermediary (n = 33)	Group differences[c] (indicated in bold)
Age at cross-exam (months)	114.95 (12.84)	110.55 (12.24)	108.92 (13.96)	118.73 (14.44)	$F_{(3, 172)} = 4.13$, $p = .007$*** **RI > S-RC**
WASI-II[a] (IQ)	109.89 (12.97)	106.50 (12.42)	109.39 (13.99)	100.94 (14.20)	$F_{(3, 172)} = 3.70$, $p = .01$* **RI < BP**
TOMAL-2 composite[a] (Memory)	113.95 (15.63)	111.75 (14.18)	112.53 (12.47)	108.73 (16.25)	$F_{(3, 172)} = .93$, $p = .43$
TOMAL-2 Verbal[a] (Verbal Memory)	114.43 (16.17)	113.95 (15.29)	110.47 (14.29)	104.94 (16.74)	$F_{(3, 172)} = 3.05$, $p = .03$* **RI < BP**
TOMAL-2 non-verbal[a] (non-verbal memory)	110.20 (18.18)	106.73 (15.42)	111.68 (13.68)	110.76 (20.11)	$F_{(3, 172)} = .63$, $p = .60$
BPVS-3[a] (receptive vocabulary)	95.65 (13.02)	94.73 (12.79)	94.87 (13.17)	87.52 (15.30)	$F_{(3, 172)} = 2.95$, $p = .03$* **RI < BP**
ELT-2 sequencing[a] (Narrative ability)	109.83 (9.06)	107.70 (9.27)	112.11 (8.43)	109.12 (6.91)	$F_{(3, 172)} = 1.76$, $p = .16$
ELT-2 Grammar and Syntax[a] (grammatical morphology)	106.92 (10.40)	106.97 (10.42)	108.82 (9.45)	103.79 (11.49)	$F_{(3, 171)} = 1.40$, $p = .25$
CELF-4-UK recalling sentences[b] (grammatical understanding/production)	10.58 (3.41)	11.70 (2.19)	11.26 (2.46)	10.85 (3.23)	$F_{(3, 172)} = 1.31$, $p = .27$
CELF-4-UK formulated sentences[b] (Sentence formulation/production)	9.28 (3.28)	10.15 (2.81)	10.58 (2.75)	8.94 (3.34)	$F_{(3, 172)} = 2.39$, $p = .07$
TEA-Ch sky search[b] (selective attention)	9.35 (2.64)	9.25 (2.59)	8.89 (2.85)	8.97 (3.45)	$F_{(3, 172)} = .27$, $p = .84$
TEA-Ch score![b] (sustained attention)	9.03 (3.36)	8.85 (3.30)	9.32 (3.80)	8.91 (3.53)	$F_{(3, 172)} = .13$, $p = .94$
TEA-Ch dual task[b] (sustained-divided attention)	6.91 (3.57)	6.80 (3.46)	6.03 (3.81)	5.15 (3.67)	$F_{(3, 172)} = 2.03$, $p = .11$
Memory trace score (max = 18)	4.61 (3.02)	5.78 (2.99)	5.89 (3.25)	6.94 (2.61)	$F_{(3, 172)} = 4.79$, $p = .003$** **RI > BP**
Delay to cross-exam. (months)	10.15 (1.78)	11.20 (1.47)	11.34 (1.56)	12.36 (0.55)	$F_{(3, 172)} = 16.78$, $p < .001$*** **BP < all others** **RI > all others**

Note: * $p < .05$, ** $p < .01$, *** $p < .001$.

Abbreviations: BPVS-2, British Picture Vocabulary Scale, second edition; CELF-4-UK, Clinical Evaluation of Language Fundamentals, fourth edition, UK version; ELT-2, Expressive Language Test, second edition; TEA-Ch, Test of Everyday Attention for Children; TOMAL-2, Test of Memory and Learning, second edition; WASI-II, Wechsler Abbreviated Scale of Intelligence, second edition.

[a]Standardised scores (mean 100, SD 15).

[b]Scaled scores (mean 10, SD 3).

[c]For paired comparisons after Bonferroni corrections.

Sketch-reinstatement of context (sketch-RC)

This followed the procedure of the Best-Practice interview except that, prior to phase 5 (free recall), witnesses were instructed to think about the event and draw whatever reminded them about it, as well as what happened. Witnesses were asked to explain to the interviewer what they were drawing. After finishing their sketch, children were asked to give a free recall account of what happened (as per the Best-Practice interview) and were told they could use their drawing to point out or explain things (Dando et al., 2009).

2.2.3 | Phase 3: Cross-examination

Prior to the cross-examination, children were 'refreshed' on their evidence as per Achieving Best Evidence guidance (Ministry of Justice, 2011) and the Registered Intermediary Procedural Guidance Manual (Ministry of Justice, 2015). This is standard practice for witnesses in advance of cross-examination within courts in England and Wales. Therefore, as in real-life, cross-examination performance may draw upon original memories of the event and recent memories of the refreshed interview. The researcher visited the child to explain that, in the next day or so, they would be speaking to a barrister who would ask them some questions about the staged event they previously saw. The researcher explained that the child would be listening to the audio of their interview,[4] to remind them of the event and what they had said. After refreshing of the evidence, the researcher again reminded the child about the forthcoming cross-examination.

A team of six barristers was recruited for the cross-examinations, comprising four men and two women. Five were currently practising barristers, whilst one was no longer practising but had their own legal business. Barristers had between 5–21 years of criminal law experience (mean = 15.2 years).

Cross-examination: A new methodological approach

For the cross-examination, a 'defence statement' was developed for each version of the staged event, which the barristers were asked to put to the children. This created a more realistic situation in which the barrister was representing a defendant in relation to a charge of theft. The defence statement (and the cross-examination protocol) was developed with the advice and guidance of an experienced barrister. The first two items in the statement included correct information designed to set the scene, establish rapport with the child witness, and make them feel at ease. The remaining points contained an element of untruthfulness (except for points 6 and 7, which were included so children did not feel that they were disagreeing with all the points the barrister was raising). Table 2 provides a sample defence statement for one version of the event.

Barristers were asked to challenge the child on all seven of the 'false' points (e.g., "I think you've got a little bit mixed up because it wasn't the phone that Adam put in his pocket, it was the keys, wasn't it?") a maximum of four times (a decision, in consultation with one of the barristers, to avoid ethical concerns). As there was variability in this (based on barrister judgement), scores only reflect whether a child complied immediately, following challenge/s, or not at all. If the child complied with the challenge on first time of asking, they received a resistance score of 0; if they complied with a challenge on the second or subsequent time of asking, they received a resistance score of 1; and if they did not comply at all, they received a maximum resistance score of 2. Average resistance scores on each of the seven false points could range from 0–2, with higher scores indicating higher cross-examination resilience (i.e., lower compliance with false statements). On a few occasions, barristers judged that it was not necessary to pose all challenges to the children. In real life, barristers make judgements about how much/little to press a witness and do not take a fixed approach, so the present study aimed to reflect this. Therefore, mean resistance scores were calculated for each child based on the total number of challenges given.

We were careful to code the child's original recall of information pertaining to each of the seven false points (taken from the investigative interview), so this score could be controlled in the analyses. These 'memory trace' scores were allocated for full (3), moderate (2), partial (1) or no (0) knowledge about six of the false points in terms of degree of information recalled in the investigative interview. For one other point (false item 5), this was a complete confabulation about something that did not happen at all in the event, therefore, a score of 0 was allocated for all children because it was not possible to code this item in terms of original recall of information (maximum memory trace score = 18: see Table 1 for mean memory trace scores and Appendix S1 for full details of the coding scheme). Fifteen percent of the transcripts were independently coded by a second rater for memory trace scores and intra-class correlations for information pertaining to each of the challenges ranged from .89 to 1.00, indicating excellent inter-rater reliability.

Cross-examination: The study protocol

One special measure available to support vulnerable witnesses in courts in England and Wales is the 'live link'. The child is not present in the courtroom with the barristers, judge or jury, but is in a separate room. Those in the courtroom see the child via a television screen, and the child can see the judge or barrister on his/her screen. To mimic this, cross-examinations were performed using video conferencing software (Skype). A female researcher was in a room with the child at their school and partially took on the role of 'judge'. We could not entirely replicate the judge role as we had no facility for the child to view the judge only via the screen—and for ethical reasons the researcher had to be with the child—so this aspect of the study must be viewed as approximate to real-life. There was a brief 'ground rules hearing' between the judge and the barrister prior to each individual cross-examination (with or without an RI) where the judge explained any important considerations to the barrister (e.g., age of child, any additional needs they had). As a prelude to the cross-examination, the judge explained to the child that they: (1) needed to tell the truth—must not guess or make anything up; (2) could say that they 'do not know' or 'cannot remember'; (3) should say if they do not understand something the barrister says; (4) could tell the barrister if they get something wrong; and (5) should say if there is a problem of any kind

TABLE 2 Sample defence statement from one of the two versions of the event (including the 'truth' and the seven 'false' statements)

	Points from the defence statement	The 'ground truth'—From the event
1	One morning last year, Max and I visited a school to give a talk about the Victorians to the children and their teachers	True
2	Max was wearing a blue top and has short brown hair. I was wearing a grey top and had long blond hair tied back in a ponytail.	True
3	When we arrived, a woman helped us by setting up the video camera at the back which recorded the talk.	False item 1—Adam set up the video camera. There was no woman involved in the event.
4	We told the children some rules that Victorian children had to obey, for instance, we said that boys must learn needlework	False item 2—whilst the children were told about rules, this specific example is incorrect—the children were told that girls (not boys) had to learn needlework.
5	We showed the children a slate and Max showed them how to write the letters of the alphabet on it with chalk.	False item 3—the children were shown a slate, but Max wrote a sum on the slate (not the alphabet).
6	Max is very forgetful and during the talk he asked the children to remind him not to forget his phone at the end of the talk.	True
7	Max then put his phone on the chair in the hall.	True
8	Max says that I stole his phone by taking it and putting it in my pocket—I did not do this. Max's phone was on the chair the whole time. I did not go near the chair at any time during or after the talk.	False item 4—Adam did take Max's phone and put it in his pocket.
9	I did borrow Max's keys during the talk and put them in my pocket.	False item 5—there were no keys involved in the staged event.
10	At the end of the talk, Max forgot his coat.	False item 6—Max forgot his jumper (which he spoke about at the start of the talk).
11	When Max forgot his coat, I had to go back to get it.	False item 7—Max (not Adam) returned after he had left, to collect the forgotten item.

Note: Whilst the other version of the event was very similar, points 4–11 on the defence statement differed: for example, there were slightly different names (Mark and Alex) for the key actors; children saw the theft of a set of keys, but the barrister had to put to them that it was, in fact, a phone; and the children were told that boys had to learn technical drawing (with the barristers suggesting to them that this was girls).

(as per the Judicial College Bench Checklist: Young Witness Cases, 2012). The judge also described the role of the barrister, explaining that they would be asking the child questions about what happened during the staged event. The judge added that the job of the barrister was to test the evidence, so they may ask questions that challenge what the child has said, but all the child needed to do was tell the truth about what they could remember or say if they did not know the answer. Whilst judges are advised to explain how often breaks are planned, and to inform the child that the judge can always see them via live link (even if they cannot see the judge), these elements were not incorporated in the instructions as: (a) the cross-examinations were short, and breaks would not be needed; (b) the judge was already in the room with the child.

Once the child and barrister were introduced, they listened to the child's audio of their investigative interview together, so everyone could hear it (barristers were provided with a transcript of the children's testimony, as well as basic demographic information, in advance of the cross-examination, to enable them to prepare their questions; in real-life, they would have access to the child's evidence-in-chief in advance of the refreshing of the evidence). The barrister then began questioning the child, with the only stipulations being that they were to cover all points on the defence statement (unless the child appeared to show any signs of distress), and that—for ethical reasons—they were not to excessively challenge the child on their testimony (no more than four challenges per point).

At three time points (before, during and after the cross-examination), children were presented with a 10-point visual analogue rating scale. This enabled us to monitor how worried or anxious the children were (1 = no anxiety; 10 = high anxiety) and to offer additional support or reassurance if their responses highlighted that they were affected by the cross-examinations. Note that these anxiety ratings were not study variables but introduced for ethical reasons. Most children were not highly anxious at any point. Before the cross-examination, 7 children (4%) had scores at the top end of the anxiety scale (8, 9, 10); during the cross-examination this figure was 9 children (5%); after the cross-examination nearly all (171 children, 97%) had the lowest anxiety scores of 1, 2 or 3 (and the remaining 5 children had moderate scores of 4, 5 or 6). Cross-examinations were, on average, 8.56 min long (SD = 2.24 min, range 3.53–16.25 min).

Cross-examination protocol: The RI condition
The protocol for the cross-examinations was the same across three interview conditions (Best-Practice, Sketch-RC and Verbal Labels), but there were some differences for the RI condition. As per

recommendations for best practice in England and Wales at the time of the study (Registered Intermediary Procedural Guidance Manual, Ministry of Justice, 2015), children received RI assistance both at their initial interview and again at cross-examination. Of the 33 children in the RI condition, 18 were assisted by the same RI at both stages, which is also recommended best practice, and 15 had a different RI at cross-examination (although using exactly the same protocol). In real cases there is also likely to be some variability in whether the same RI is available for both stages. RI assistance involved the following: Prior to the cross-examination, all children were re-assessed by the RI to ensure that information about the child's communication needs (originally collected 8–13 months previously) was up-to-date and accurate. This re-assessment took place at least a week before the cross-examination and consisted of: (1) re-establishing rapport with the children; (2) explaining what would happen in the cross-examination; (3) checking the children could say they 'do not know' or 'cannot remember', and could state whether the barrister (adult) was wrong or right; (4) checking the children could respond to questions beginning with, for example, 'when' or 'how'; and (5) preparing simplified instructions for the judge to present during the preamble before the cross-examination (to make them easier to follow and remember). The barristers and RIs also met together for a dedicated 'ground rules hearing' (see Cooper et al., 2015, for further details) prior to all RI cross-examinations, in which the RIs explained what their role was and discussed their recommendations with the barristers. In real-life, ground rules hearings would take place for each individual child. However, the RIs noted that many of their recommendations would be the same for most children in the study, so one overall ground rules hearing was conducted (with RIs flagging individual cases where necessary). (Note that this was in addition to the 'short' ground rules hearing for each individual child just before the cross-examination (regardless of interview condition).)

At the ground rules hearing, RIs discussed the principles of questioning and gave barristers a written summary of their suggestions. The summary included advice to: practice the live link prior to the child coming into room; use a short and simple preamble; be careful about references to do with time (e.g., when, how long), or questions requiring a number in the answer (e.g., how many); use a slow pace; allow thinking time; use short sentences with only one point per question; use basic vocabulary and sentence structure; and use names the child knows people by. Question types were discussed and RIs recommended avoiding questions that: were negatively phrased; were statements with a questioning intonation; were tagged (e.g., 'Max forgot his coat, *didn't he*?'); had an answer implied; and were repeats of already asked questions. The RIs additionally: reviewed each barrister's list of cross-examination questions and highlighted the specific needs of individual children prior to cross-examination sessions (discussions by phone or email); reminded barristers that visual materials were available if needed to support expressive language (drawing materials, small world figures/furniture) and sequencing of events (post-it notes, timelines); and brought along calming objects so they were available to the children if necessary. Importantly, RIs did not

TABLE 3 Types of child responses during cross-examinations with explanations

Type of response	Explanation
Complies (true)	When a child complies with what the barrister has said, in relation to a true (correct) statement
Complies (false)	When a child complies with what the barrister has said, in relation to a false (incorrect) statement
Resists (true)	When a child has resisted what the barrister has said, in relation to a true (correct) statement
Resists (false)	When a child has resisted what the barrister has said, in relation to a false (incorrect) statement
No response	The child has not given a response
Open response	When a child has given a response to a barrister's open question (they cannot comply or resist, as the child is given the opportunity to tell their version of events)
Seeks clarification	The child seeks Clarification (e.g., "I do not know what you mean")

intervene about the content of the questions but rather the format (Ministry of Justice, 2015), for example, "[Barrister's name], could that question be rephrased, as you know it's a tagged question" or if they thought the child would not understand the question, for example, "I am not sure [child's name] will understand that complex question". In the RI condition, an RI was present alongside each child for every cross-examination, simplified the instructions given to the children by the judge, and made interventions during the cross-examinations as required. For example, if the barrister moved away from planned questions or began to use statements with tags, the RI would remind the barrister of best practice. The RI also intervened if the child appeared not to understand or follow the questioning.

Coding child responses and barrister questions
Children's responses were coded into mutually exclusive categories reflecting whether they complied, resisted, did not respond, responded with an open question, or sought clarification (see Table 3). When a child responded with an acknowledgement (e.g., 'okay'), this was not coded as a response to the question. If the child said they were not sure, this did not mean they had complied: children were instructed to say 'do not know' if this was the case, so they were resisting the barrister's attempts to get them to agree with them.

Barrister questions were coded into one of seven overarching mutually exclusive primary categories (see Table 4 for details). All questions (as well as non-content-based utterances which were given the code 'other') were coded separately, even if they occurred, sequentially, e.g. "*That's really helpful, thank you very much* (code=other). *Okay, now they talked to you about Victorian schools* (code=assertion, true). *Did they tell you lots of things about what*

TABLE 4 The seven overarching primary codes for barrister questions during cross-examinations, with explanations and examples

Type of question	Explanation	Example
Invitation open	A question that invites the witness to offer their account and does not declare the answer (or have a correct answer)	"Who set up the video camera?"
Invitation closed (true)	A question that invites a yes or no response, or asks for confirmation—includes true (correct) information	"Did Alex set up the video camera?"
Invitation closed (false)	A question that invites a yes or no response, or asks for confirmation—includes false (incorrect) information	"Did Mark set up the video camera?"
Assertion (true)	Questions in the form of a statement, which is true (correct); or a statement of the child's previous response	"Alex set up the video camera?"
Assertion (false)	Questions in the form of a statement, which is false (incorrect)	"Alex did not set up the video camera?"
Option posing	Questions in the form of two or more options (that may include the option to choose 'something else')	"Had he got his back to you, front, side, something else?" "Was it blond or brown hair?"
Other	Utterances that were not content-based questions (e.g. signpost, credibility, praise, clarification and reassurance)	"Lovely, thank you so much B."; "Can I ask you some questions about that because that's really helpful?"

happened in Victorian times? (code=invitation closed, true)" would attract three codes as indicated. Barrister questions were additionally coded for each instance of 17 other secondary features (see Table 5), which were not mutually exclusive categories, that is, a question could challenge credibility as well as contain a tag. The coding systems were developed by looking at guidance on questioning available at the time (May 2015) in The Advocate's Gateway (Toolkit 6, 2015), the Judicial College Bench Checklist: Young Witness Cases (2012), and the Equal Treatment Bench Book (Judicial College, 2013). We also used an iterative process of discussion and reflection on the coding process to capture all question types in one overarching primary code, yet additionally reflect other relevant question features within the secondary codes. The classification system was designed to be as comprehensive and informative as possible, although it could not capture more subtle features such as intonation.

Reliability of coding

To establish coder agreement, 10% of scripts were coded independently by a second coder. Overall percentage agreement was 91% (range 86%–100%) for the child codes, 89% (range 82%–92%) for the barrister primary codes and 88% (range 81%–100%) for the barrister secondary codes, all of which represented moderately high agreement.

2.2.4 | Control measures

Around the time that the children took part in Phases 1 and 2 of the study, several cognitive measures (intelligence, language, memory, attention) were administered to ensure factors that may affect eyewitness recall and cross-examination were controlled or matched between interview groups (see Table 1 for differences between conditions that were controlled for statistically).

Intelligence

Two subtests (Vocabulary and Matrix Reasoning) of the second edition of the Wechsler Abbreviated Scale of Intelligence (WASI-II; Wechsler & Zhou, 2011) were used to provide an assessment of intellectual ability and to establish suitability for entry into the study.

Language

The British Picture Vocabulary Scale Third Edition (BPVS-3; Dunn et al., 2009) was used to provide a measure of receptive vocabulary. Two subtests (Sequencing, and Grammar and Syntax) of the Expressive Language Test 2 (ELT-2, Bowers et al., 2010) assessed narrative ability and grammatical morphology, respectively. Two subtests (Recalling Sentences and Formulated Sentences) of the Clinical Evaluation of Language Fundamentals, 4th edition (CELF-4 UK; Semel et al., 2006) provided an assessment of the ability to recall and formulate grammatically correct, meaningful sentences.

Memory

Subtests from the Test of Memory and Learning 2 (TOMAL-2; Reynolds & Voress, 2007) were used to provide a composite memory measure, comprising both verbal ('Memory for Stories' and 'Paired Recall') and non-verbal ('Facial Memory' and 'Visual Sequential Memory') memory.

Attention

The Test of Everyday Attention for Children (Tea-Ch; Manly et al., 1999) was used to assess a range of relevant attention skills: selective/focused attention (the 'Sky Search' subtest); sustained attention (the 'Score!' subtest); and sustained-divided attention (the 'Sky Search Dual Task' subtest).

2.3 | General procedure

Ethical approval was obtained from the relevant university Research Ethics Committee. Prior to participation, written consent was obtained from parents, and children also gave their own written assent to participate. At the start of Phase 1, children viewed the

TABLE 5 Further secondary classifications of features of the barrister's questions during cross-examinations, with explanations and examples

Classification	Explanation	Example
Tag	A question asking for confirmation, suggestive as it communicates the expected response	"Mark picked up the keys, *didn't he*?"
Credibility	A question that challenges the integrity or credibility of the witness, or their memory	"You think they did. You say you think, did you actually see them do it or are you guessing?"
Negatives	A question containing a negative	"Didn't Mark pick up the keys, not Alex?"
Repetition	Repeating the same question, even if interspersed by others	"Did Alex take the keys?" A: "No". "Did Alex take the keys?"
Confirmation	The advocate confirms the answer the child has given, in a best practice way—a permissible and gentle way of checking evidence	"I want to make sure I understand what you said…" "so they showed you the slate but they did not do any writing, is that what you are saying?"
Clarification	The advocate checks that the answer the child has given is what was intended	"You nodded, so is that a yes, brilliant, thank you very much."
Social influence of another person	The barrister suggests that 'someone else' told them that what the child has said happened did not really happen	"Alex told me he did not take the keys"
Possibility	A question that suggests that what the barrister is putting to them might be true (even if the witness is unsure)—possibility is introduced	"And was there maybe a lady helping out?"
Complex	A question that is linguistically complex, because of the large number of instructions contained in it, because of ambiguity or because it has conjunctions making it long-winded	"But I hope that if I ask you some questions, and I know you have, you have gone through what you said in your, um, your interview about it, uh, if I ask you some questions, we might be able to work out together, um, exactly what happened when those two people came to school, okay?"
Idiom	Phrase with a figurative or literal meaning	"Now let us go back to square one"
Do you remember…?	Questions asking the witness if they remember what they said on a previous occasion are particularly frowned upon	"Do you remember any other adults in the room?" "Can you remember that?"
Lying	Directly accuses the witness of lying	Note: an example is not given as there were no examples of accusing the child of lying in the current study.
Signpost	Explaining or signposting changes of subject (includes references to original evidence, e.g., "in your interview, you said that…")	"Now we are going to talk about the other man, the man with the long hair called Adam."
Praise	Thanking or commending the child in an encouraging way	"That's brilliant, thank you for that. I've only got one more thing to ask you…"
Filler	Irrelevant questions	"The men who came to your school, were they funny?"
Name	The advocate uses the child's name	"That's really helpful, you have got a very good memory here N."
Reassurance	The advocate provides reassurance that the child is doing okay	"That's okay, not to worry, so you cannot help me with who set it up if you do not remember."

staged event and immediately took part in the brief interviews (see Henry, Messer, et al., 2017). Phase 2, investigative interviews (see Henry, Crane, et al., 2017) and identification line-ups (see Wilcock et al., 2018), took place around 1 week later. Cognitive testing also took place around this time, which was split over several sessions to fit in with school timetables and to ensure children remained engaged with tasks. Phase 3, the cross-examinations, took place 8–13 months (Mean = 11.06 months, SD 1.69 months) after viewing the staged event. As some variability in this delay emerged across conditions (see Table 1) due to timing of school holidays and availability of RIs/barristers, we controlled for delay in the primary statistical analyses. All children were refreshed on their evidence in one session with the researcher, before the researcher returned at least 1 day later to conduct the cross-examination with the barrister. Children in the RI

condition were re-assessed in a session prior to the refreshing of their evidence (on a different, earlier day). The RI was always present at the cross-examination and, beforehand, used a visual aid to explain to the child that they should only say what really happened, that if the barrister got something wrong, they could tell them, and equally that it was OK to say that the barrister 'got it right'. In addition, the children were told, using the visual aid, that it was OK to say 'I don't know', 'I can't remember', or 'I don't understand'.

3 | RESULTS

The key outcome measures for the primary research questions concerned: (1) children's cross-examination resistance scores on seven

cross-examination challenges pertaining to false elements from the defence statement; and (2) whether RI assistance during cross-examinations reduced children's compliance with these challenges on false information.

Table 6 shows mean resistance scores (SDs). Ten children resisted all seven challenges on false information that the barrister put to them (5.7%), meaning that 94.3% of children complied with at least one challenge. Five children complied with all seven challenges on false information (2.8%).

Hierarchical multiple regression was used to examine whether cross-examination resistance scores on the seven false information challenges differed between children in the RI condition versus other conditions (note that we had no reason to expect cross-examination differences for the Sketch-RC and Verbal Labels conditions as they involved adaptations to investigative interview protocols). At step 1, three background variables showing differences between interview conditions (see Table 1 for details), namely age at cross-examination, IQ, and Verbal Memory, were controlled (BPVS scores also differed between interview conditions, but IQ and BPVS scores were highly correlated, $r = .66$, so only IQ was controlled). Three further control variables included: memory trace scores (concerning relevant information pertaining to the false information challenges) as children in the RI condition had higher memory trace scores (they had benefitted from RI intervention at the investigative interview stage) (Henry, Crane, et al., 2017); event version (A or B); and length of delay before cross-examination (this differed across condition—see Table 1). At step 2, three dummy-coded interview condition variables were included to test for differences between conditions in cross-

examination resistance. Best-Practice was the reference (control) group to which the other three conditions were compared: RI, Sketch-RC and Verbal Labels. The dependent variable was average cross-examination resistance score (see Table 6). With nine predictor variables in total, Green (1991) would recommend a sample size of at least 122, thus for the current regression our sample size exceeded the minimum numbers recommended. Key statistical checks (multicollinearity, Durbin-Watson, tolerance and VIF statistics, Cook's and Mahalanobis distances, standardised DFbetas, leverage values, plots of standardised residuals and predicted standardised values, standardised residuals, partial plots) were within acceptable limits (Field, 2013).

Table 7 gives details of the regression. The full regression model was significant, $F(9, 166) = 5.37$, $p < .001$, accounting for 22.5% (18.3% adjusted) of the variance in cross-examination resistance scores. Step 1 was significant (R^2 change = 7.7%; $F[6, 169] = 2.35$, $p = .03$), indicating that the six control variables accounted for a small proportion of the variance when entered on their own (although only memory trace was significant when inspecting standardised β-values, $\beta = .16$, $p = .04$). Crucially, Step 2 was also significant (R^2 change = 14.8%; $F[3, 166] = 10.61$, $p < .001$), indicating interview condition differences in cross-examination resistance. Inspection of the standardised β-values at Step 2 showed that only the contrast between the RI and Best-Practice interview conditions was significant ($\beta = .47$, $p < .001$). As tentatively predicted, children in the RI condition were less compliant with cross-examination challenges than children in the Best-Practice condition, with higher resistance scores (an average of .63 higher with a 95% CI of .37–.88), once all other

TABLE 6 Resistance scores for children in each interview condition (highest average resistance score is 2, lowest is 0), total numbers of child responses, and proportional (prop.) scores for different types of responses for each interview condition

Scores	Best-practice ($n = 65$)	Verbal labels ($n = 40$)	Sketch-RC ($n = 38$)	Registered intermediary ($n = 33$)
Cross-examination resistance score (average over 7 'false' defence statement elements)	**.85** (.49) .71 (.00–2.00)	**.80** (.41) .86 (.00–1.67)	**.94** (.56) .84 (.00–2.00)	**1.42** (.45) 1.43 (.43–2.00)
Total number of child responses across full cross-examination	**43.82** (11.40) 42 (26–71)	**45.35** (16.41) 42 (12–78)	**46.37** (16.2) 41 (16–78)	**52.82** (13.05) 56 (33–76)
Prop. complies with true statement	**.38** (.10) .37 (.15–.60)	**.38** (.13) .34 (.00–.66)	**.36** (.10) .34 (.21–.53)	**.29** (.12) .24 (.11–.61)
Prop. complies with false statement	**.13** (.07) .13 (.00–.37)	**.15** (.10) .11 (.03–.39)	**.13** (.11) .12 (.00–.56)	**.05** (.05) .05 (.00–.17)
Prop. resists true statement	**.12** (.07) .12 (.00–.28)	**.09** (.08) .08 (.00–.33)	**.11** (.08) .09 (.00–.35)	**.15** (.09) .14 (.03–.37)
Prop. resists false statement	**.18** (.10) .16 (.03–.50)	**.20** (.10) .21 (.00–.41)	**.20** (.10) .18 (.00–.39)	**.20** (.06) .20 (.06–.32)
Prop. 'no response'	**.03** (.05) .00 (.00–.30)	**.03** (.04) .03 (.00–.16)	**.04** (.05) .02 (.00–.16)	**.01** (.02) .01 (.00–.05)
Prop. open response	**.16** (.11) .17 (.00–.38)	**.12** (.12) .06 (.00–.35)	**.15** (.13) .15 (.00–.45)	**.29** (.15) .32 (.03–.50)
Prop. seeks clarification	**.01** (.02) .00 (.00–.09)	**.02** (.04) .00 (.00–.20)	**.02** (.02) .00 (.00–.08)	**.01** (.03) .00 (.00–.12)

Note: Mean scores (SDs) are given on line 1 (means are in bold), medians (ranges) on line 2.

TABLE 7 Summary of the multiple regression predicting average cross-examination resistance

Step	B	SE B	β	p
Step 1				
Constant	−.70	.67		.29
Age	.005	.003	.14	.08
IQ	.00	.003	.01	.90
Verbal memory	.002	.003	.07	.44
Memory trace	.03	.01	.16	**.04***
Performance version (A or B)	.08	.09	.08	.37
Cross-exam delay (months)	.04	.03	.14	.10
Step 2				
Constant	−.08	.63		.90
Age	.003	.003	.07	.33
IQ	.004	.003	.10	.25
Verbal memory	.003	.003	.09	.29
Memory trace	.01	.01	.08	.32
Performance version (A or B)	.01	.08	.01	.92
Cross-exam delay (months)	−.02	.03	−.07	.47
Best-practice-v-verbal labels	−.02	.11	−.015	.86
Best-practice-v-sketch-RC	.12	.11	.10	.25
Best-practice-v-RI	.63	.13	.47	**<.001***

Note: Significant predictors are indicated in bold.

variables had been accounted for. All other variables were non-significant predictors at Step 2. To check whether initial viewing of the event live or via video affected the findings, this regression was repeated with only children who had seen the event live ($n = 144$). The results were identical in all respects, except that memory trace score at Step 1 just missed significance ($p = .055$).[5]

3.1 | Children's responses

The first subsidiary research question had two components: first, whether the numbers of compliant responses by children to barrister challenges on false information would be lower in RI interviews; and second, whether the numbers of resistant responses by children to barrister challenges on false information would be higher in RI interviews. Whilst children gave, on average, 46.40 (SD = 14.31) responses across the cross-examination, this differed across interview conditions, $F(3, 172) = 3.10$, $p = .03$, partial $\eta^2 = .05$. Bonferroni corrected paired comparisons indicated that children gave significantly more responses in the RI condition (mean = 52.82, SD = 13.05) than in the Best-Practice condition (mean = 43.82, SD = 11.40) ($p = .02$), but no other comparisons were significant. Given this, subsequent analyses were carried out on proportional scores (proportions of each type of response in relation to total number of responses for each child). Table 6 includes mean proportions of the seven types of responses.

Proportional data were not all normally distributed, so Kruskal-Wallis tests were used to explore whether there were differences between interview conditions for each type of response, with a Bonferroni adjusted significance level of $p < .007$ (for seven tests). Bonferroni corrected follow-up paired comparisons were used to explore any differences between interview conditions. Values of η^2 represent large (>.14), medium (.06–.14) or small (.01–.06) effect sizes.

Two analyses were of relevance to predictions as follows. For *Complies (with false information)* responses, a significant interview condition effect was present, $H(3) = 34.04$, $p < .001$, $\eta^2 = .18$. Follow-up comparisons indicated that, as predicted, proportions of Complies (false) responses were lower in the RI condition than in all other conditions: Best-Practice ($z = 5.39$, $p < .001$); Verbal Labels ($z = 4.94$, $p < .001$); and Sketch-RC ($z = 4.22$, $p < .001$). For *Resists (false information)* responses, no significant interview condition effect was present, contrary to predictions, $H(3) = 3.09$, $p = .38$, $\eta^2 = 00$.

We did not have specific predictions for the other five question types, but we present these analyses here, for completeness. For Complies (with true information) responses, a significant interview condition effect was present, $H(3) = 18.33$, $p < .001$, $\eta^2 = .09$: proportions of Complies (true) responses were lower in the RI condition than in other conditions: Best-Practice ($z = 4.02$, $p < .001$); Verbal Labels ($z = 3.54$, $p = .002$); and Sketch-RC ($z = 3.05$, $p = .014$). For Open responses, a significant interview condition effect was present, $H(3) = 21.96$, $p < .001$, $\eta^2 = .11$: proportions of Open responses were higher in the RI condition than in other conditions: Best-Practice ($z = −3.48$, $p = .003$); Verbal Labels ($z = −4.48$, $p < .001$); and Sketch-RC ($z = −3.51$, $p = .003$). No other interview condition effects reached significance for child responses: Resists (true information), $H(3) = 10.65$, $p = .014$, $\eta^2 = .04$; No Response, $H(3) = 5.86$, $p = .12$, $\eta^2 = .02$; and Seeks Clarification, $H(3) = 4.44$, $p = .22$, $\eta^2 = .01$.

3.2 | Barrister questions

A second subsidiary research question concerned whether, in the RI condition, the barristers' questions might be more consistent with best practice guidance for cross examination or re-examination. Table 8 shows mean numbers of questions per cross-examination, as well as proportions of each of the seven primary overarching types of questions for each interview condition. Overall, barristers asked an average of 61.39 (SD =18.78) questions per child. A one-way analysis of variance (data were normally distributed) showed a significant effect of interview condition, $F(3, 172) = 3.89$, $p = .01$, partial $\eta^2 = .06$. Bonferroni corrected paired comparisons indicated that barristers asked significantly more questions in the RI condition (mean = 71.09, SD =17.87) than in the Best-Practice (mean = 58.92, SD =16.75) ($p = .01$) and Sketch-RC conditions (mean = 58.26, SD =16.43) ($p = .02$). (This is consistent with real cross examinations: to simplify questions, asking two questions rather than one is often necessary.) The RI and Verbal Labels (mean = 60.35, SD =22.42) conditions did not differ significantly ($p = .08$). Given these differences, further analyses on barrister questions were performed using

TABLE 8 Total number of barrister questions across the full cross-examination, and proportions (prop.) of each of the seven primary overarching types of questions for each interview condition

Scores	Best-practice (n = 65)	Verbal labels (n = 40)	Sketch-RC (n = 38)	Registered intermediary (n = 33)
Total number of barrister questions	**58.92** (16.75) 59 (26–105)	**60.35** (22.42) 65.5 (25–117)	**58.26** (16.43) 56.5 (20–90)	**71.76** (18.03) 70.5 (28–109)
Prop. invitation open	**.05** (.04) .04 (.00–.22)	**.04** (.04) .02 (.00–.20)	**.05** (.05) .04 (.00–.2)1	**.12** (.05) .12 (.02–.21)
Prop. invitation closed (true)	**.24** (.11) .25 (.03–.51)	**.18** (.09) .17 (.04–.38)	**.20** (.10) .20 (.02–.43)	**.33** (.07) .34 (.15–.49)
Prop. invitation closed (false)	**.17** (.07) .16 (.05–.38)	**.15** (.05) .15 (.03–.28)	**.13** (.05) .13 (.02–.25)	**.16** (.05) .16 (.04–.29)
Prop. assertion (true)	**.32** (.10) .32 (.15–.54)	**.36** (.09) .36 (.22–.57)	**.34** (.10) .34 (.17–.58)	**.21** (.06) .20 (.10–.33)
Prop. assertion (false)	**.11** (.10) .07 (.00–.3)7	**.16** (.11) .15 (.02–.41)	**.14** (.11) .11 (.01–.39)	**.06** (.04) .06 (.00–.16)
Prop. option-posing	**.03** (.03) .03 (.00–.12)	**.02** (.03) .00 (.00–.08)	**.02** (.03) .00 (.00–.11)	**.015** (.02) .00 (.00–.06)
Prop. other	**.08** (.07) .06 (.00–.23)	**.09** (.06) .08 (.01–.23)	**.11** (.07) .10 (.02–.24)	**.11** (.04) .11 (.02–.20)

Note: Mean scores (SDs) are given on line 1 (means are in bold) and medians (ranges) on line 2.

proportional scores: the total number of questions in each question-type category were divided by the total number of barrister questions asked per child. These proportional data were not all normally distributed, so Kruskal-Wallis tests were used to explore whether there were interview condition differences on each question type, with a Bonferroni adjusted significance level of $p < .007$ (for seven tests). Bonferroni corrected follow-up paired comparisons were used to explore any differences between interview conditions.

Invitation Open questions differed significantly across interview condition, $H(3) = 45.24$, $p < .001$, $\eta^2 = .25$. Proportions of Invitation Open questions were higher in the RI condition than in other conditions: Best-Practice ($z = -5.63$, $p < .001$); Verbal Labels ($z = -6.18$, $p < .001$); and Sketch-RC ($z = -5.03$, $p < .001$).

Invitation Closed (true information) questions differed significantly across interview condition, $H(3) = 39.91$, $p < .001$, $\eta^2 = .22$. Proportions of Invitation Closed (true) questions were higher in the RI condition than in other conditions: Best-Practice ($z = -3.80$, $p = .002$); Verbal Labels ($z = -5.91$, $p < .001$); and Sketch-RC ($z = -5.00$, $p < .001$). A difference between Verbal Labels and Best-Practice also emerged ($z = 2.87$, $p = .02$).

Assertion (true information) questions differed significantly across interview condition, $H(3) = 48.78$, $p < .001$, $\eta^2 = .27$. Proportions of Assertion (true) questions were lower in the RI condition than in any other condition: Best-Practice ($z = 5.41$, $p < .001$); Verbal Labels ($z = 6.49$, $p < .001$); and Sketch-RC ($z = 5.51$, $p < .001$).

Assertion (false information) questions differed significantly across interview condition, $H(3) = 16.71$, $p < .001$, $\eta^2 = .08$. Proportions of Assertion (false) questions were lower in the RI condition than in the Verbal Labels condition ($z = 3.64$, $p = .001$) and the Sketch-RC condition ($z = 2.81$, $p = .03$); and that they were higher in the Verbal Labels condition than in the Best-Practice condition ($z = -2.80$, $p = .03$).

Option-posing questions differed significantly across interview condition, $H(3) = 11.49$, $p = .009$, $\eta^2 = .05$. Proportions of option-

posing questions were lower in RI than in Best-Practice interviews ($z = 2.65$, $p = .049$). No other paired comparisons were significant.

Invitation Closed (false information) questions ($p = .10$) and *Other* questions ($p = .03$) showed no significant interview condition differences.

Table 9 includes breakdowns of barrister questions into 17 secondary features. These are presented as proportions (i.e., divided by the total number of barrister questions), but will not add up to one given the categories are not mutually exclusive (any question could be classified in one or more ways). (Note: no instances of the barrister saying the child was 'lying' were found; similarly, mean proportions for use of idiom were less than 1%; so these data were excluded.) These proportional data were not all normally distributed, so Kruskal-Wallis tests were used to explore interview condition differences for each question feature, with a Bonferroni adjusted significance level of $p < .003$ (for 15 tests). Bonferroni corrected follow-up paired comparisons were used to explore any differences between interview conditions.

Eight secondary question features showed significant interview condition differences.

Tags, $H(3) = 53.71$, $p < .001$, $\eta^2 = .29$. Proportions of Tags were lower in the RI condition than in any other condition: Best-Practice ($z = 5.58$, $p < .001$); Verbal Labels ($z = 6.54$, $p < .001$); and Sketch-RC ($z = 6.23$, $p = .008$).

Credibility, $H(3) = 30.74$, $p < .001$, $\eta^2 = .16$. Proportions of Credibility challenges were lower in the RI condition than in other conditions: Best-Practice ($z = 5.46$, $p < .001$); Verbal Labels ($z = 3.92$, $p = .001$); and Sketch-RC ($z = 3.03$, $p = .01$).

Repetition, $H(3) = 22.54$, $p < .001$, $\eta^2 = .11$. Proportions of Repeated questions were higher in the RI condition than in the Best-Practice ($z = -4.65$, $p < .001$) and Verbal Labels ($z = -3.17$, $p = .009$) conditions.

TABLE 9 Proportions of features of barrister questions coded into 17 secondary categories for each interview condition

Questionfeature classification	Best-practice ($n = 65$)	Verbal labels ($n = 40$)	Sketch-RC ($n = 38$)	Registered intermediary ($n = 33$)
Tag	**.19** (.18) .12 (.00–.77)	**.28** (.22) .22 (.04–.70)	**.24** (.19) .18 (.00–.70)	**.04** (.03) .05 (00–.11)
Credibility	**.07** (.05) .07 (.00–.19)	**.06** (.04) .05 (.00–.16)	**.05** (.04) .05 (.00–.16)	**.02** (.02) .02 (.00–.08)
Negative	**.03** (.04) .02 (.00–.14)	**.04** (.05) .02 (.00–.14)	**.05** (.06) .03 (.00–.20)	**.04** (.04) .03 (.00–.16)
Repetition	**.05** (.08) .00 (.00–.28)	**.08** (.11) .00 (.00–.35)	**.09** (.09) .08 (.00–.35)	**.12** (.07) .13 (.00–.37)
Confirmation	**.18** (.13) .16 (.00–.62)	**.17** (.11) .16 (.00–.55)	**.17** (.13) .14 (.00–.62)	**.13** (.06) .12 (.02–.25)
Clarification	**.02** (.02) .00 (.00–.10)	**.02** (.02) .00 (.00–.09)	**03** (.04) .02 (.00–.16)	**.02** (.02) .02 (.00–.08)
Social influence of another person	**.03** (.05) .02 (.00–.24)	**.04** (.05) .02 (.00–.16)	**.04** (.05) .02 (.00–.15)	**.09** (.05) .08 (.01–.24)
Possibility	**.04** (.05) .03 (.00–.26)	**.03** (.03) .02 (.00–.11)	**.03** (.03) .02 (.00–.14)	**.007** (.01) .00 (.00–.04)
Complex	**.05** (.06) .03 (.00–.30)	**.06** (.08) .02 (.00–.31)	**.08** (.10) .03 (.00–.40)	**.08** (.06) .08 (.00–.22)
Do you remember	**.14** (.09) .12 (.03–.40)	**.17** (.13) .12 (.00–.47)	**.16** (.10) .15 (.03–.42)	**.12** (.08) .12 (.00–.39)
Signpost	**.13** (.05) .13 (.00–.25)	**.11** (.05) .10 (.00–.23)	**.13** (.05) .12 (.03–.20)	**.13** (.05) .12 (.04–.29)
Praise	**.10** (.05) .09 (.02–.24)	**.10** (.07) .08 (.02–.36)	**.08** (.05) .07 (.01–.20)	**.04** (.04) .03 (.00–.14)
Filler	**.004** (.01) .00 (.00–.05)	**.006** (.01) .00 (.00–.06)	**.01** (.02) .00 (.00–.08)	**.02** (.02) .02 (.00–.07)
Name	**.08** (.05) .07 (.00–.21)	**.07** (.04) .06 (.00–.16)	**.06** (.04) .06 (.00–.16)	**.05** (.04) .05 (.00–.12)
Reassurance	**.04** (.04) .03 (.00–.16)	**.02** (.03) .02 (.00–.13)	**.02** (.02) .02 (.00–.06)	**.006** (.01) .00 (.00–.04)

Note: Categories are not mutually exclusive so overall proportions do not add to 1. Mean proportions (SDs) are given on line 1 (means are in bold), medians (ranges) on line 2.

Social Influence of another person, $H(3) = 28.64$, $p < .001$, $\eta^2 = .15$. Proportional use of Social Influence was higher in the RI condition than in other conditions: Best-Practice ($z = -5.28$, $p < .001$); Verbal Labels ($z = -3.95$, $p < .001$); and Sketch-RC ($z = -3.67$, $p = .001$).

Possibility, $H(3) = 22.30$, $p < .001$, $\eta^2 = .11$. Proportional use of Possibility was lower in the RI condition than in other conditions: Best-Practice ($z = 4.71$, $p < .001$); Verbal Labels ($z = 3.03$, $p = .014$); and Sketch-RC ($z = 2.97$, $p = .018$).

Praise, $H(3) = 26.92$, $p < .001$, $\eta^2 = .14$. Proportions of Praise were lower in the RI condition than in other conditions: Best-Practice ($z = 4.86$, $p < .001$); Verbal Labels ($z = 4.37$, $p < .001$); and Sketch-RC ($z = 3.43$, $p = .004$).

Filler questions, $H(3) = 22.90$, $p < .001$, $\eta^2 = .12$. Proportions of Filler questions were higher in the RI condition than in the Best-Practice ($z = -4.61$, $p < .001$) and Verbal Labels ($z = -3.67$, $p = .001$) conditions.

Reassurance, $H(3) = 24.63$, $p < .001$, $\eta^2 = .13$. Proportions of Reassurance were lower in the RI condition than in other conditions:

Best-practice ($z = 4.95$, $p < .001$); Verbal Labels ($z = 3.15$, $p = .01$); and Sketch-RC ($z = 2.73$, $p = .038$).

4 | DISCUSSION

In this paper, a novel experimental methodology for the cross-examination of vulnerable child witnesses has been presented. Experienced barristers questioned children based on a 'defence statement' containing seven false elements, without recourse to a 'script' (as is typically used in experimental research on cross-examination). As predicted, children complied with barristers' challenges on this false information to a high degree: 94% of children complied with at least one cross-examination challenge on false information, consistent with previous experimental studies using scripted questioning in which compliance rates ranged between 70% and 98% (cf. Bettenay et al., 2014; Righarts et al., 2015; Zajac et al., 2009; Zajac & Hayne, 2003, 2006). Our findings underline concerns about whether

cross-examination is a reliable method for obtaining best evidence from child witnesses, given that lawyers try to 'persuade children to change details in their accounts, often by exploiting their developmental limitations' (Andrews & Lamb, 2016, p. 953).

We also tested, in an exploratory way, whether RI assistance, available in England and Wales, might help children to give better evidence by reducing compliance with barristers' cross-examination challenges on false information. As per recommendations for best practice in England and Wales at the time of the study (Ministry of Justice, 2015, see also current Registered Intermediary Procedural Guidance Manual, Ministry of Justice, 2020b), children received RI assistance at their initial interview and again at cross-examination. As tentatively predicted, RI assistance at cross-examination reduced children's compliance with false information, even after controlling for background cognitive factors, other key factors that could have influenced the findings, and memory for relevant details of the original event. Specifically, when children were challenged to agree with evidence that was 'false' (i.e., the barrister was suggesting that the child should agree with something in the defence statement that was 'false' and the child needed to resist this line of questioning), RI assistance made it less likely that children would comply with the barrister's challenges. This finding highlights the importance of using RIs for typically developing children to ensure that they do not give compliant responses to false information or change their responses when pressurised. For a child to accept that it was "possible", for example, that a woman had helped set up the video camera (when no such woman was present), would be enough to be used by the defence lawyer in undermining the evidence given the burden and standard of proof in criminal trials.[6] Overall, these exploratory findings about RIs support current recommendations in the Equal Treatment Bench Book (Judicial College, 2018, 2020) that: "All young witnesses should ideally have an intermediary assessment as, no matter how advanced they appear, their language comprehension is likely to be less than that of an adult witness" (paragraph 98, p. 60). For typical children, RIs also help improve volume of recall in interviews and accuracy of identification in video line-ups (Henry, Crane, et al., 2017; Plotnikoff & Woolfson, 2012; Wilcock et al., 2018). Overall, providing RIs for primary age typical children may improve the quality of their evidence.

A subsidiary research question concerned whether, when we broke down children's specific responses to barrister questioning, these responses would be less compliant with *and* more resistant to challenges on false information in the RI condition. As tentatively predicted, significantly lower proportions of 'complies with false information' responses were given by children in the RI condition than in other conditions (5% in the RI condition vs. 11%–13% in other conditions): children were less likely to agree with a barrister's false statement in the RI condition. Although the proportions of 'resists false information' responses did not vary with interview condition, as expected, this could be because resisting a false statement is more difficult for a child (i.e., actively saying 'that is not true') than not agreeing with a false statement (this is possible with more passive responses such as 'do not know' or providing no response at all). Overall, these

findings accorded closely with the primary research finding that RI assistance helped children to reduce compliance in response to barrister challenges on false information.

A final subsidiary research question concerned whether barristers would ask questions more aligned with best practice recommendations in the RI condition. In support of this, barristers asked proportionally more Invitation Open questions in the RI condition. Whilst these have been associated with inconsistencies (due to the longer answers they elicit) (Pichler et al., 2020), they are in accord with best practice (Ministry of Justice, 2011), are least likely to lead the witness (Henderson et al., 2019), and are highly valued by practitioners (Magnusson et al., 2020). Invitation Open questions were, nevertheless, relatively rare, as reported in real cases (e.g., Andrews & Lamb, 2016; Pichler et al., 2020; Zajac et al., 2018). Rates here ranged from 4% to 5% in non-RI conditions, to 12% in RI cross-examinations. Also consistent with best practice, barristers asked proportionally fewer Assertion questions in the RI condition. Such questions are risky because they present a strong statement that might be difficult to resist and could, thus, lead the witness (Henderson et al., 2019; Judicial College, 2013; The Council of the Inns of Court, 2019). Proportions of Assertions about true information were significantly lower (21%) in RI interviews than in other interviews (range 32%–36%), although proportions of Assertions about false information did not reveal such consistent group differences (RI = 6%, other conditions = 11%–16%).

Other findings concerning the barrister questions were harder to interpret. Invitation Closed (true information) questions were significantly higher in RI interviews (33%) than in other interviews (range 18%–24%), although no group differences emerged for Invitation Closed (false information) questions. In real cases it may not be apparent whether these yes/no style questions are misleading, if the truth is not known. Yes/no questions for 'true' information may be less risky in terms of leading the witness, whereas yes/no questions for false information could be actively misleading. Finally, the small group difference in Option-Posing questions indicated somewhat fewer of these in the RI condition than the Best-Practice condition, but rates of these questions were low (3% or less in all conditions), so this result should be viewed with caution.

Further detailed classification of the features of barrister questions into secondary categories offered some evidence that they were more aligned with best practice recommendations in the RI condition. First, there were reductions in the use of suggestive tag questions (4% vs. 19%–28%), supporting existing best practice guidance (Judicial College, 2013, 2018; Ministry of Justice, 2011; The Advocate's Gateway, 2015; The Council of the Inns of Court, 2019). Second, there were reductions in challenges to the children's credibility (2% vs. 5%–7%) and fewer suggestions that something 'possibly' happened (<1% vs. 3%–4%). Although these questions were infrequent overall, the lower rates in RI interviews may have increased the child's confidence in themselves as a respondent, particularly as children dislike having their credibility challenged (Plotnikoff & Woolfson, 2012).

More difficult to interpret was the fact that RI interviews showed increases in repetitions compared to most other interviews (12% vs. 5%–9%). Question repetition is not recommended as it could

confuse or exploit the child into changing answers (Andrews et al., 2015b; Judicial College, 2013, 2018; Ministry of Justice, 2011; The Council of the Inns of Court, 2019). In fact, the RIs removed any repeated questions when checking barrister questions before cross-examination, so it is possible that barristers re-introduced them to help children to follow the line of questioning if they lost track, or because they were unable to diverge from the listed questions if they wanted to press a point. Other differences in RI interviews that were unexpected included the use of 'social influence of another person' being more common (9% vs. 3%–4%). This could reflect barristers switching from challenging the children's credibility outright or inferring the 'possibility' of being incorrect, to relying on a gentler approach by suggesting the child was affected by social influence of another person instead. It could also reflect a technique to check the child's ability to challenge the barrister (or the defendant) who expresses a different view. There was also less praise and reassurance (4% vs. 8%–10%, and <1% vs. 2%–4%, respectively) in RI cross-examinations, perhaps because barristers opted to give more praise and reassurance in non-RI interviews to conceal the fact that they were undermining the child's evidence. Finally, there were more irrelevant (filler) questions in RI interviews (although note that the RI vs. S-RC comparison here was not significant and the values were low in all cases: RI 2% and other conditions 1% or less). Overall, despite some areas of uncertainty, these findings suggest that recommendations by RIs regarding the wording of cross-examination challenges could align questioning more closely with best practice recommendations.

The study findings may contribute to internationally available sources of guidance about how lawyers should question children in court, given concerns in this area (e.g., Andrews et al., 2015a). Further training about how to question vulnerable witnesses (e.g., advocates in England and Wales now attend training to acknowledge the '20 Principles of Questioning', The Council of the Inns of Court, 2019), along with pre-trial ground rules hearings as standard (see Henderson et al., 2019), would be useful for all barristers involved in child cases. The Advocate's Gateway provides detailed recommendations for barristers and other legal professionals on questioning a range of vulnerable witnesses, including children (www.theadvocatesgateway.org). Pre-trial guidance aimed at children may also help because practice sessions for responding to cross-examination style questions on an unrelated topic can significantly improve children's overall accuracy during a subsequent cross-examination interview (Irvine et al., 2016; Righarts et al., 2013), provided it is given close to the interview date (O'Neill & Zajac, 2013). Future research could investigate a combination of RI assistance and timely pre-trial preparation (perhaps delivered as part of the RI assessment), as combining these interventions may further improve the quality of children's cross-examination evidence.

One area the study was unable to illuminate was whether the RI assistance impacted on the child's responses, the barrister's questioning technique, or both. We are also uncertain about the mechanisms and exact points through which RI assistance operated, but it is important to note that the overarching role of the RI is to support the child's communication needs (e.g., simplifying instructions, using visual aids) and impact the barrister's questioning to ensure it is appropriate. All of this should help the witness more easily understand what others are saying so that they can communicate better. Further research could unpick the important mechanisms underpinning the interplay between children's responses and barristers' questions. The very nature of cross-examination requires some fluidity in questioning and a good advocate will always be influenced by the child's responses. The exception to this would be to use a rigid script of questions (which is necessary in some extreme cases, but not generally). Otherwise, the barrister can be flexible and adapt in response to the child's answers. This was one of the advantages to the present novel approach to assessing cross-examination empirically, which has, to our knowledge, not been addressed in previous empirical work.

There are some limitations to the study that should be acknowledged. One is that the findings are applicable only to defence barristers, as different lines of questioning may be applied by prosecution barristers (Denne et al., 2020). Another is that children in the RI condition, as per best practice guidance (Ministry of Justice, 2015), had already received RI assistance during previous phases of the mock criminal investigation: this was given at the investigative interview stage (which also included an identification line-up). Therefore, the current conclusions can only be applied to children who have had RI assistance throughout a criminal investigation which, in practice, is not always the case (RIs may sometimes only brought in at trial stage, although this is not recommended). A related issue was that children in the RI condition remembered more about the initial witnessed event, as RI assistance was effective in increasing the volume of accurate recall at investigative interview (Henry, Crane, et al., 2017). This meant that children in the RI condition started their cross-examination with a recall advantage. We mitigated this by controlling for how well the child had recalled key facts about the false information in the defence statement (memory trace scores). Although memory trace was not a significant predictor of cross-examination resistance in the full regression (and many children did not score highly on this measure), future research could match on initial memory of the staged event before instigating cross-examinations in groups with and without RI assistance. This method would mean that no children could be included who had previously undergone an investigative interview assisted by an RI, but such a method would provide evidence about the effectiveness of RI assistance brought in only at the trial stage.

Further limitations are as follows. We used a mild minor crime event that took place in a familiar environment (the children's school), so were unable to replicate the anxiety, unfamiliarity and potential trauma of a real court case, which limits generalisation of the findings. Children were seen by friendly and supportive researchers, and the barristers were also approachable and experienced—they were, partly, chosen on the basis of having previous experience in cross-examining children (for ethical reasons)—again, this might not be so in real-life. Our ground rules hearings for non-RI children were also brief, and more recent guidance now recommends they are included as 'good practice' for all young witnesses (Judicial College, 2018, revisions 2020, Equal Treatment Bench Book, p. 64). Finally, the length of the cross-examinations, for ethical reasons was short (average 8.56 min)

compared to real cases (reported in England and Wales as between 45 min and 3 h, Baverstock, 2016). However, Henderson et al. (2019) reported much shorter video-recorded cross-examinations (16 min) in a pilot trial of this special measure in England, and with new advocate training and guidance, cross-examinations are likely to be more limited in length (e.g., Judicial College, 2018). Similarly, although studies of court transcripts in Scotland, California and New Zealand have emphasised the large numbers of questions (ranging from 160 to 500) posed to children by prosecutors and defence lawyers (Andrews & Lamb, 2016; Andrews et al., 2015a; Klemfluss et al., 2014; Zajac & Cannan, 2009), the number of questions posed during pilot video-recorded cross- and direct-examinations in Henderson et al.'s (2019) study was lower (average = 92). Thus, although the current cross-examinations contained fewer questions (average = 61), the overall numbers of questions may be more aligned with the newer pre-recorded cross-examinations in England. Given that long and complex cross-examinations will likely lead to fatigue, worsening the quality of evidence (e.g., Zajac et al., 2018), changes that encourage shorter questioning should be advantageous.

5 | CONCLUSION

The current study was the first to use a more ecologically valid defence statement as the basis for unscripted cross-examinations. Using this novel method, we found that children complied with a very high number of barrister challenges on false information. However, we also found exploratory evidence that RI assistance reduced children's compliance with barristers' cross-examination challenges on false information. This could be, in part, because the barristers asked questions that were somewhat more aligned with best practice recommendations in the RI condition. These findings extend previous research on the utility of RIs during investigations (evidence-gathering interviews and identification line-ups). They provide additional evidence of the importance of using RIs to ensure typically developing young children can give accurate testimony during the final investigative phase (cross-examination).

CONFLICT OF INTEREST
The authors declare no conflict of interest.

ORCID
Lucy A. Henry https://orcid.org/0000-0001-5422-4358
Laura Crane https://orcid.org/0000-0002-4161-3490
Rachel Wilcock https://orcid.org/0000-0002-6535-8228

ENDNOTES

[1] In 2014, a pilot programme of video-recorded live-link cross-examinations in England was trialled (Baverstock, 2016), involving pre-trial Ground Rules Hearings (which can place restrictions on traditional cross-examination practices to improve witness experiences) and video-recorded cross-examinations (to reduce delays between giving initial evidence and cross-examination in court). The scheme has now been rolled out to all Crown Courts across England and Wales. Henderson et al. (2019) and Henderson and Lamb (2019) evaluated cases with and without pre-trial Ground Rules Hearings prior to pre-recorded children's cross examination. With these measures, fewer suggestive questions were asked, and question complexity was reduced.

[2] One hundred and fourty four children saw the event live and thirty two children saw it via video. A *t*-test on number of correct details recalled in the brief evidence-gathering statement across these two groups was non-significant: Mean live = 33.82 (SD = 14.84); Mean video = 38.94 (SD = 14.17), t(174) = 1.78, p = .08. Nevertheless, we ran our primary analyses on both the full sample and the live-only sample to ensure this variable did not affect the findings.

[3] Two versions of the event differed slightly in terms of names used (Alex/Adam, Max/Mark), objects shown (abacus/slate), and prop 'stolen' (keys/phone). No differences emerged in the number of correct details recalled in the brief evidence gathering statement across these two versions for the current sample: Mean Version A (n = 87) = 34.03 (SD = 12.94); Mean Version B (n = 89) = 35.45 (SD = 16.49), t(174) = .63, p = .53. Nevertheless, we controlled for this variable in our primary analyses.

[4] We did not have permission to video all children, although we did have permission to audio record all children, therefore, audio recordings were used to refresh children on their evidence.

[5] Results were similar when barrister was included as a further control variable—the only significant predictor at Step 2 was the contrast between the RI and Best-Practice interview conditions (p < .001). At Step 1 memory trace (p = .03) and barrister (p = .01) were significant predictors. However, this analysis is only exploratory because not all barristers were evenly spread across conditions.

[6] Although our study specifically looked at compliant responses to 'false information', which are undesirable, in some cases such responses would be appropriate if the information were true.

REFERENCES
Andrews, S. J., & Lamb, M. E. (2016). How do lawyers examine and cross-examine children in Scotland? *Applied Cognitive Psychology*, 30, 953–971. https://doi.org/10.1002/acp.3286

Andrews, S. J., Lamb, M. E., & Lyon, T. D. (2015a). Question types, responsiveness and self-contradictions when prosecutors and defense attorneys question alleged victims of child sexual abuse. *Applied Cognitive Psychology*, 29, 253–261. https://doi.org/10.1002/acp.3103

Andrews, S. J., Lamb, M. E., & Lyon, T. D. (2015b). The effects of question repetition on responses when prosecutors and defense attorneys question children alleging sexual abuse in court. *Law and Human Behavior*, 39, 559–570. https://doi.org/10.1037/lhb0000152

Baverstock, J. (2016). *Process evaluation of pre-recorded cross-examination pilot (section 28)*. Ministry of Justice.

Bettenay, C., Ridley, A. M., Henry, L. A., & Crane, L. (2014). Cross-examination: The testimony of children with and without intellectual disabilities. *Applied Cognitive Psychology*, 28, 204–214. https://doi.org/10.1002/acp.2979

Bowers, L., Huisingh, R., LoGiudice, C., & Orman, J. (2010). *The expressive language test 2*. LinguiSystems.

Brown, D., & Pipe, M.-E. (2003). Individual differences in Children's event memory reports and the narrative elaboration technique. *Journal of Applied Psychology*, 88, 195–206. https://doi.org/10.1037/0021-9010.88.2.195

Bull, R. (2010). The investigative interviewing of children and other vulnerable witnesses: Psychological research and working/professional practice. *Legal and Criminological Psychology*, 15, 5–23. https://doi.org/10.1348/014466509X440160

Collins, K., & Krahenbuhl, S. (2020). Registered intermediaries' assessment of children's communication: An exploration of aims and processes. *Evidence & Proof*, 24(4), 374–395. https://doi.org/10.1177/1365712720952335

Cooper & Wurzel. (2014). Better the second time around? Department of Justice Registered Intermediaries Schemes and lessons from England and Wales. *Northern Ireland Legal Quarterly*, 65, 39–61.

Cooper, P., Backen, P., & Marchant, R. (2015). Getting to grips with ground rules hearings: A checklist for judges, advocates and intermediaries to promote the fair treatment of vulnerable people in court. *Criminal Law Review*, 2015(6), 420–435.

Cooper, P., & Mattison, M. (2017). Intermediaries, vulnerable people and the quality of evidence: An international comparison of three versions of the English intermediary model. *The International Journal of Evidence & Proof*, 21, 351–370. https://doi.org/10.1177/1365712717725534

Dando, C., Wilcock, R., & Milne, R. (2009). The cognitive interview: The efficacy of a modified mental reinstatement of context procedure for frontline police investigators. *Applied Cognitive Psychology*, 23, 138–147. https://doi.org/10.1002/acp.1451

Denne, E., Sullivan, C., Ernest, K., & Stolzenberg, S. N. (2020). Assessing children's credibility in courtroom investigations of alleged child sexual abuse: Suggestibility, plausibility, and consistency. *Child Maltreatment*, 25(2), 224–232. https://doi.org/10.1177/1077559519872825

Dunn, L. M., Dunn, D. M., & Styles, B. (2009). *British picture vocabulary scale* (3rd ed.). GL Assessment.

Evans, A. D., Lee, K., & Lyon, T. D. (2009). Complex questions asked by defense lawyers but not prosecutors predicts convictions in child abuse trials. *Law and Human Behavior*, 33, 258–264. https://doi.org/10.1007/s10979-008-9148-6

Field, A. P. (2013). *Discovering statistics using IBM SPSS statistics: And sex and drugs and rock 'n' roll* (4th ed.). Sage publications.

Green, S. B. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research*, 26, 499–510. https://doi.org/10.1207/s15327906mbr2603_7

Hanna, K., Davies, E., Crothers, C., & Henderson, E. (2012). Questioning child witnesses in New Zealand's criminal justice system: Is cross-examination fair? *Psychiatry, Psychology and Law*, 19, 530–546. https://doi.org/10.1080/13218719.2011.615813

Hanna, K., & Henderson, E. (2018). '[expletive], that was confusing wasn't it?' Defence lawyers' and intermediaries' assessment of the language used to question a child witness. *The International Journal of Evidence and Proof*, 22, 411–427. https://doi.org/10.1177/1365712718796527

Henderson, E. (2002). Persuading and controlling: The theory of cross-examination in relation to children. In H. L. Westcott, G. M. Davies, & R. H. C. Bull (Eds.), *Children's testimony: A handbook of psychological research and forensic practice* (pp. 279–294). Chichester.

Henderson, H. M., Andrews, S. J., & Lamb, M. E. (2019). Examining children in English high courts with and without implementation of reforms authorized in section 28 of the youth justice and criminal evidence act. *Applied Cognitive Psychology*, 33, 252–264. https://doi.org/10.1002/acp.3472

Henderson, H. M., & Lamb, M. E. (2019). Does implementation of reforms authorized in section 28 of the youth justice and criminal evidence act affect the complexity of the questions asked of young alleged victims in court? *Applied Cognitive Psychology*, 33, 201–213. https://doi.org/10.1002/acp.3466

Henry, L. A., Crane, L., Nash, G., Hobson, Z., Kirke-Smith, M., & Wilcock, R. (2017). Verbal, visual, and intermediary support for child witnesses with autism during investigative interviews. *Journal of Autism and Developmental Disorders*, 47, 2348–2362. https://doi.org/10.1007/s10803-017-3142-0

Henry, L. A., Messer, D. J., Wilcock, R., Nash, G., Kirke-Smith, M., Hobson, Z., & Crane, L. (2017). Do measures of memory, language, and attention predict eyewitness memory in children with and without autism spectrum disorder? *Autism and Developmental Language Impairments*, 2, 1–17. https://doi.org/10.1177/2396941517722139

Home Office. (1989). *Report of the advisory group on video-recorded evidence*. Chairman His Honour Judge Thomas Pigot.

Irvine, B., Jack, F., & Zajac, R. (2016). Preparing children for cross-examination: Do the practice questions matter? *Psychology, Crime & Law*, 22, 858–878. https://doi.org/10.1080/1068316X.2016.1197224

Judicial College. (2013). Equal treatment bench book. Retrieved 20 May 2015 from https://www.sentencingcouncil.org.uk/wp-content/uploads/equal-treatment-bench-book-2013-with-2015-amendment.pdf.

Judicial College (2018). Equal treatment bench book. Retrieved 13 February 2020 from https://www.judiciary.uk/wp-content/uploads/2018/02/equal-treatment-bench-book-february-v6-2018.pdf.

Judicial College Bench Checklist: Young witness cases. (2012). Retrieved 20 May 2015 from https://www.judiciary.uk/publications/jc-bench-checklist-young-wit-cases/.

Judicial College (2020). Equal treatment bench book (March 2020 revision). Retrieved 13 April 2021 https://www.judiciary.uk/wp-content/uploads/2018/02/ETBB-February-2018-amended-March-2020.pdf.

Klemfuss, J. Z., Quas, J. A., & Lyon, T. (2014). Attorneys' questions and children's productivity in child sexual abuse criminal trials. *Applied Cognitive Psychology*, 28, 780–788. https://doi.org/10.1002/acp.3048

Krahenbuhl, S. (2019). Mock jurors' perceptions of a child witness: The impact of the presence and/or absence of a registered intermediary during cross-examination. *Psychology, Crime & Law*, 25(7), 713–728. https://doi.org/10.1080/1068316X.2018.1543421

Lamb, M. E., Malloy, L. C., & La Rooy, D. J. (2011). Setting realistic expectations: Developmental characteristics, capacities, and limitations. In M. E. Lamb, D. J. La Rooy, L. C. Malloy, & C. Katz (Eds.), *Children's testimony: A handbook of psychological research and forensic practice* (pp. 15–48). Wiley- Blackwell.

Magnusson, M., Ernberg, E., Landström, S., & Akehurst, L. (2020). Forensic interviewers' experiences of interviewing children of different ages. *Psychology, Crime & Law*, 26(10), 967–989. https://doi.org/10.1080/1068316X.2020.1742343

Manly, T., Robertson, I. H., Anderson, V., & Nimmo-Smith, I. (1999). *The test of everyday attention for children*. Pearson.

Ministry of Justice. (2011). *Achieving best evidence in criminal proceedings: Guidance on interviewing victims and witnesses, and using special measures*. Her Majesty's Stationary Office.

Ministry of Justice. (2015). The registered intermediary procedural guidance manual. Retrieved 20 May 2015 from https://www.theadvocatesgateway.org/images/procedures/registered-intermediary-procedural-guidance-manual.pdf.

Ministry of Justice. (2020a). The witness intermediary scheme: Annual report 2019–20. Retrieved 13 October 2020 from https://assets.

publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/919858/witness-intermediary-scheme-annual-report-2019-2020.pdf.

Ministry of Justice. (2020b). *Registered intermediary procedural guidance manual.* Retrieved 13 April 2021 from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/955316/registered-intermediary-procedural-guidance-manual.pdf.

O'Neill, S., & Zajac, R. (2013). Preparing children for cross-examination: How does intervention timing influence efficacy? *Psychology, Public Policy, and Law, 19,* 307–320. https://doi.org/10.1037/a0031538

Pichler, A. S., Powell, M., Sharman, S. J., Zydervelt, S., Westera, N., & Goodman-Delahunty, J. (2020). Inconsistencies in complainant's accounts of child sexual abuse arising in their cross-examination. *Psychology, Crime & Law., 27,* 341–356. https://doi.org/10.1080/1068316X.2020.1805743

Plotnikoff, J., & Woolfson, R. (2009). *Measuring up? Evaluating implementation of government commitments to young witnesses in criminal proceedings.* 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). Hart Publishing.

Plotnikoff, J., & Woolfson, R. (2015). *Intermediaries in the criminal justice system: Improving communication for vulnerable witnesses and defendants.* Policy Press.

Reynolds, C., & Voress, J. K. (2007). *Test of memory and learning* (2nd ed.). Pro-Ed.

Righarts, S., Jack, F., Zajac, R., & Hayne, H. (2015). Young children's responses to cross-examination style questioning: The effects of delay and subsequent questioning. *Psychology, Crime & Law, 21,* 274–296. https://doi.org/10.1080/1068316X.2014.951650

Righarts, S., O'Neill, S., & Zajac, R. (2013). Addressing the negative effect of cross-examination questioning on Children's accuracy: Can we intervene? *Law and Human Behavior, 37,* 354–365. https://doi.org/10.1037/lhb0000042

Semel, E., Wiig, E. H., & Secord, W. A. (2006). *Clinical evaluation of language fundamentals* (4th UK ed.). Harcourt Assessment.

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). Hart Publishing.

Taggart, J. (2021). 'I am not beholden to anyone. I consider myself to be an officer of the court': A comparison of the intermediary role in England and Wales and Northern Ireland. *The International Journal of Evidence & Proof, 25,* 1–22. https://doi.org/10.1177/13657127211002291

The Advocate's Gateway. (2015). Toolkit 6: Planning to question a child or young person. Retrieved 20 May 2015 from https://www.theadvocatesgateway.org/images/toolkits/6-planning-to-question-a-child-or-young-person-141215.pdf.

The Council of the Inns of Court. (2019). The 20 Principles of Questioning: A guide to the cross-examination of vulnerable witnesses. Retrieved 13 February 2020 from https://www.icca.ac.uk/wp-content/uploads/2019/05/20-Principles-of-Questioning.pdf.

Wechsler, D., & Zhou, X. (2011). *Wechsler abbreviated scale of intelligence* (2nd ed.). Pearson.

Wilcock, R., Crane, L., Hobson, Z., Nash, G., Kirke-Smith, M., & Henry, L. A. (2018). Supporting child witnesses during identification lineups: Exploring the effectiveness of registered intermediaries. *Applied Cognitive Psychology, 32,* 367–375. https://doi.org/10.1002/acp.3412

Wilcock, R., Crane, L., Hobson, Z., Nash, G., Kirke-Smith, M., & Henry, L. A. (2019). Brief report: Eyewitness identification in child witnesses on the autism spectrum. *Research in Autism Spectrum Disorders, 66,* 101407. https://doi.org/10.1016/j.rasd.2019.05.007

Youth Justice and Criminal Evidence Act. (1999). Retrieved 20 May 2015 from http://www.legislation.gov.uk/ukpga/1999/23/contents.

Zajac, R., & Cannan, P. (2009). Cross-examination of sexual assault complainants: A developmental comparison. *Psychiatry, Psychology & Law, 16,* S36–S54. https://doi.org/10.1080/13218710802620448

Zajac, R., Gross, J., & Hayne, H. (2003). Asked and answered: Questioning children in the courtroom. *Psychiatry, Psychology and Law, 10,* 199–210. https://doi.org/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. https://doi.org/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,* 3–16. https://doi.org/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. https://doi.org/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. https://doi.org/10.1016/j.dr.2012.06.006

Zajac, R., Westera, N., & Kaladelfos, A. (2018). The "good old days" of courtroom questioning: Changes in the format of child cross-examination questions over 60 years. *Child Maltreatment, 23,* 186–195. https://doi.org/10.1177/1077559517733815

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.