

Interviewers' and intermediaries' perceptions of problematic interview questions and their proposed solutions

The International Journal of
Evidence & Proof
2026, Vol. 30(2) 113–131
© The Author(s) 2025



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/13657127251361174
journals.sagepub.com/home/epj



Sonja P. Brubacher¹ , Sarah L. Deck¹ , David Plater² ,
Michael E. Lamb³  and Martine B. Powell¹ 

Abstract

Witness intermediaries facilitate communication between legal actors and children or other vulnerable populations, and sometimes intermediaries assist in investigative interviews. This research explored the skills of investigative interviewers and intermediaries (who either had experience acting in that role or were eligible to do so) in identifying problematic questions during interviews. Forty-one participants (18 interviewers, 23 eligible or experienced intermediaries) listened to excerpts from interviews with three alleged victims of sexual or physical abuse which contained a mix of appropriate and inappropriate questions. Participants paused the audio whenever they wished to make a comment and also completed a knowledge test on questioning best practices. Interviewers identified more problematic questions and scored higher on the knowledge test than intermediaries. Whereas interviewers were primarily concerned with the extent to which the interview involved open-ended questioning, intermediaries focused on the extent to which questioning involved supports and clear scaffolding. The findings suggest that well-trained interviewers are equipped to handle most questioning issues with children independently; however, intermediaries' attentiveness to question scaffolding suggests they might aid communication with interviewees facing more pronounced communication challenges (e.g., developmental delay).

Keywords

intermediaries, child witnesses, investigative interviews, forensic interviews, vulnerability

¹Centre for Investigative Interviewing, Griffith Criminology Institute, Griffith University, Brisbane, Australia

²Adelaide Law School, The University of Adelaide, Adelaide, Australia

³Department of Psychology, University of Cambridge, Cambridge, UK

Corresponding Author:

Sonja P. Brubacher, Centre for Investigative Interviewing, Griffith Criminology Institute, Griffith University, Brisbane, Australia.
Email: s.brubacher@griffith.edu.au

The amount and quality of information communicated by child witnesses is influenced by the adult information gatherer and plays a significant role in determining case outcomes (Brown and Lamb, 2015; Pipe et al., 2013). The memory and language skills required can be difficult for young children and other vulnerable populations for whom these skills are developing (Lamb et al., 2015). To successfully provide clear and detailed reports, some children may benefit from intermediary support. Most frequently, intermediaries are employed in courts, but some jurisdictions employ them in police interviews with vulnerable witnesses (e.g., Northern Ireland, South Australia: Cashmore and Shackel, 2018; Department of Justice Northern Ireland, 2015; South Australia, 2016, rr 23). The present research explored their use in that context.

Witness intermediaries, also called communication assistants in some jurisdictions, are professionals whose intended role is to assess the communication capacities and special needs of children and other vulnerable witnesses (e.g., people with complex communication needs). In doing so, they facilitate witness communication with professionals in the justice system, ideally without undermining a defendant's right to a fair trial (*R v B*, 2010). Their role is to help a vulnerable witness to give their "best evidence," a concept that, while defined slightly differently across jurisdictions, typically involves reference to the most complete and accurate evidence a witness is can provide (e.g., Royal Commission into Institutional Responses to Child Sexual Abuse [Criminal Justice Report, August 2017] Part VII, 5–6, p. 5). The first official use of an intermediary in the criminal justice system occurred in England and Wales in 2004. There are now intermediary schemes in several countries including Australia, New Zealand, Ireland, Northern Ireland, and South Africa (Cashmore and Shackel, 2018; Cooper and Mattison, 2017, Cooper and Wurtzel, 2014; Howard et al., 2020; Kearns et al., 2024). The legislation for intermediaries in England and Wales, Northern Ireland, and Australia is similar (Cooper and Mattison, 2017). The majority of intermediaries have backgrounds in Speech and Language Therapy, with a smaller number from social work and clinical psychology (Cashmore and Shackel, 2018; Plotnikoff and Woolfson, 2015).

Intermediaries are matched with witnesses via referral services according to the intermediary's skillset, location, and availability. The intermediary then gathers information about the witness and the nature of the case, seeks third-party information about the witness's communication needs and abilities (where consent has been given), and conducts assessments in the presence of a third party (to avoid concerns about the intermediary coaching the witness or the witness disclosing abuse to the intermediary). Intermediaries work collaboratively with the legal professionals who will engage with the witness in numerous ways. This could include observation of the witness's and interviewer's communication during assessment (prior to the interview) or helping them practice answering questions (often about an unrelated event) via video link or in the courtroom (Collins and Krähenbühl, 2020; Henderson, 2015). Other activities include helping witnesses refresh their memories, reviewing their witness assessments with the interviewers or court personnel, monitoring the phrasing of questions in police interviews or court, and relaying answers from witnesses when witnesses cannot do so verbally (Cooper and Mattison, 2017).

Because the intermediary role was only recently created, relatively little is known about their skillsets, abilities, and knowledge, the kinds of questions they identify as problematic (*cf.* Hanna and Henderson, 2018), and the types of recommendations they make. Intermediaries are typically trained or informed about topics such as the ethical and procedural boundaries of their work, how to conduct assessments with vulnerable persons and write summary reports, and how to support witnesses in court (Collins and Krähenbühl, 2020). If intermediaries are not trained in best-practice interviewing, however, there is a risk that they may make recommendations that yield a contaminated memory report (Cooper and Wurtzel, 2014).

Although intermediaries are widely valued (Henderson, 2015; SALRI, 2021), we need to better understand the circumstances in which their services should be used. Scholars have noted a lack of clarity about the status and function of the intermediary role which hinders intermediaries from providing assistance

effectively and to those who need it most (Fairclough et al., 2023; Taggart, 2021). Various challenges associated with effective implementation of intermediaries' roles have been documented (Kearns et al., 2024), and these challenges contributed to the recent effective termination of an intermediary scheme in South Australia (Hoff et al., 2022; SALRI, 2021). One of these challenges is the process by which intermediaries are appointed to cases. For instance, legal professionals with experience working with intermediaries (for defendants), have expressed the perception that intermediaries always recommend their expertise is needed—a view that undermines both the credibility of the intermediaries' recommendations and the legitimacy of their role (Fairclough et al., 2023). Moreover, key stakeholders, including intermediaries and legal practitioners, have called for a standardized process for appointing intermediaries to cases, ensuring their skills are aligned with the case (Gallagher et al., 2024). We thus need to understand clearly how best to design intermediary schemes and define their scope of work (Hoff et al., 2022). Because financial and temporal resources are not infinite (Fairclough et al., 2023), attempts should be made to reserve intermediaries for those cases where their assistance is critical.

A recent review of 26 qualitative studies highlighted recurring stakeholder concerns about intermediary involvement (Kearns et al., 2024). Stakeholders included intermediaries, children and vulnerable people involved in legal matters, families of vulnerable people, law enforcement, legal professionals, and other members of the judiciary who had experience with intermediaries. Some stakeholders advocated for including intermediary services in police interviews; however, it is unclear how widely this perception extended beyond intermediary participants themselves because the review did not separate perceptions by stakeholder professional group. Henderson (2015) outlined how stakeholders' criticisms of intermediary usage tend to focus on how (vs whether) they were used and whether the witnesses who truly needed them were being identified. A third of the advocates and judges interviewed believed that witnesses who would have benefited from intermediaries were not assigned one or were assigned too late (Henderson, 2015). Additionally, around a quarter of the advocates and judges felt that intermediaries were sometimes used unnecessarily—a concern echoed by practitioners in a recent evaluation of the terminated South Australian scheme (Hoff et al., 2022). Furthermore, multiple studies have documented that lawyers and judges sometimes perceive intermediaries as offering little additional value in assisting vulnerable individuals beyond what the legal professionals themselves could provide (Fairclough et al., 2023; Henderson, 2015; Hoff et al., 2022; Taggart, 2021). In comparison, police working in Northern Ireland expressed positive feedback about their experiences working with intermediaries, though the basis for these positive views were generally not delineated (Department of Justice Northern Ireland, 2015). Hence, there is a need to identify where—and where not—additional communication assistance is warranted.

Contributions of intermediaries to evidence quality

Children's abilities to participate effectively in investigative interviews is heavily influenced by how they are questioned (Brown and Lamb, 2015). Yet children are often questioned in suboptimal ways, in both investigative interviews and in court (Lamb et al., 2018; Powell et al., 2022). Best-practice principles of interviewing are based on a foundation of human memory, child development, and social support. Intermediaries should have access to specialty knowledge about how the communication process can be achieved in a way that maximizes the quality of evidence from the complainant (Powell et al., 2015).

Several small-sample studies have suggested that intermediaries can help to improve the quality of questions asked (Hanna and Henderson, 2018) and the accuracy of children's accounts (Henry et al., 2017) but these same studies also demonstrated wide variability across intermediaries. For example, in Henry et al.'s study (2017), two intermediaries provided assessments of 38 typically developing children before the children were interviewed about a staged laboratory event about which ground truth was known. The accuracy of those children's reports, but not the quantity of details they provided, was superior to that of children who received a standard best-practice interview without an intermediary but there was a statistically significant difference between the amounts of information elicited by the two

intermediaries. In other words, the use of intermediaries did not guarantee that all children had equivalent experiences and opportunities to share their memories of the event. In a subsequent study, Henry et al. (2021) assessed whether intermediaries effectively reduced typically developing children's compliance with misleading statements made by lawyers during cross-examination. Acquiescence to misleading questions decreased when an intermediary was involved, demonstrating the value of their participation.

Especially relevant to the current study, Krähenbühl (2011) compared the abilities of lawyers and intermediaries to determine what questions they assessed as inappropriate for vulnerable witnesses and how they would rephrase them when presented with vignette scenarios. There was a great deal of inconsistency across lawyers and intermediaries regarding what questions were problematic, why they were problematic, and how they could be appropriately rephrased. Intermediaries identified twice as many instances of developmentally inappropriate communication as lawyers, suggesting a need for specialized communication support to aid lawyers in their conversations with vulnerable witnesses. However, no researchers have compared intermediaries' and well-trained investigative interviewers' skills in identifying and rephrasing questions. Since high-quality interview training focuses on developing abilities to identify and avoid problematic questions (Brubacher et al., 2022), interviewers may not require the same communicative assistance as lawyers, who are often criticized for lacking up-to-date knowledge in this area (e.g., Powell et al., 2022).

Evidence suggests that intermediary services are more critical in some contexts than in others. For example, intermediaries may be required when a witness has specialized communication needs that are likely to be beyond the bounds of training for legal professionals (e.g., cognitive impairment, trauma or mental health problem, complex communication needs; Cashmore and Shackel, 2018). Conversely, for mainstream witnesses—even very young ones—there may be little need for intermediaries when interviews are of very high quality (Powell et al., 2015). The use of intermediaries is costly given the breadth of their roles, their time spent gathering necessary information for assessment, and the need for extra meetings between intermediaries and legal professionals to relay recommendations—as well as the fact that they should never be alone with the witness, necessitating a third party to be present (Cooper and Mattison, 2017; Fairclough et al., 2023). As such, it is critical to determine the boundaries of intermediaries' roles so that resources can be directed to ensure they are supporting the vulnerable witnesses in the justice system who need them the most. This was the goal of the present research.

Current study

We aimed to 1) gain better insight, through controlled scenarios, into the skills and knowledge of trained interviewers relative to those of intermediaries and potential intermediaries; and 2) determine whether the nature of the interview—specifically the quality of questioning—affects the ability to identify problematic questions. Intermediaries assist in communication between parties, such as between vulnerable witnesses and police or lawyers in legal situations. Despite this definition, little is known about how they achieve this. For example, do they primarily help by flagging linguistically challenging questions as problematic (and to what extent can they substitute objectively better questions)? Do they also support the accuracy of witnesses' accounts by addressing questions that have a negative impact on memory (e.g., questions that are leading or suggestive)? Across two sessions incorporating an experimental design we addressed these questions. Our participant pool was a heterogeneous sample of professionals who were registered as intermediaries or who were potential intermediaries (i.e., had the qualifications to be intermediaries) in Australia and internationally, and professionals who were skilled investigative interviewers. Participants completed case recommendations based on background information about three children: a 12-year-old with developmental delay, a 13-year-old with language delay, and a typically developing 6-year-old. Subsequently, they listened to (mock) audio recordings of the interviews and were permitted to pause the recordings to describe any concerns they had. In a second, self-paced session, they identified question types or suggested a “next best question” in response to brief transcript sequences.

Hypotheses

This research was largely exploratory due to the limited understanding of intermediaries' contributions to assisting well-trained forensic interviewers; however, we made two hypotheses. Hypothesis 1: The ability to recognize problematic questions may be affected by the quality of the interview. When interviews were of high quality, we expected interviewers and intermediaries to be less attentive to problematic questions overall than when interviews were of low quality (main effect of interview type). Hypothesis 2: the questions that intermediaries identify as problematic may differ depending on the nature of the question. In Australia, intermediaries' backgrounds tend to be in speech pathology (Cashmore and Shackel, 2018), so we expected that they would focus on linguistic problems (e.g., double negatives) rather than questions that were simply worded but cognitively challenging or interfered with memory processes. For example, a simply worded question such as, "What color was her skirt?" may be misleading if the person in question wore trousers. The present study tested these hypotheses by looking at performance in real time using standardized materials while controlling statistically for variables such as interviewer performance, child developmental level, and disability status.

Method

Participants

A priori power analyses for $2 \times 2 \times 3$ mixed measures ANOVA (the first factor between-subjects) with 80% power, $\alpha = .05$, Cohens' $f = .25$ (moderate), with two groups, and six within-subjects measurements with $r = 0.3$ (medium) expected correlation among measures, indicated a desired sample size of 56. We thus aimed to recruit 30 in each group of interviewers and intermediaries. Participants were recruited via the authors' professional networks (all interviewers and six intermediaries), and the remaining intermediaries were obtained from a pool that signed up for training on effective questioning for children ($n = 17$; they participated in the current study before training began and participation was optional). All participants gave their informed consent.

Despite substantial recruitment efforts spanning three years that reached out to at least 120 professionals, only 41 professionals (18 interviewers and 23 eligible intermediaries) took part. Those who did not participate either did not respond or cited lack of time. Interviewers were all highly trained and came from Australia ($n = 6$), Canada ($n = 6$), New Zealand ($n = 2$) and the USA ($n = 4$), and they were either known either to a member of the research team directly, or to collaborators of the research team who vouched for the quality of their training. Intermediaries came from Australia ($n = 20$) and Canada ($n = 3$). Due to a computer error, demographic surveys were not sent to seven participants (six intermediaries). Of the remaining 34 participants, all were female, four were between the ages of 25 and 34, 15 were between 35 and 44, 11 were between 45 and 54, and four were age 55 +. Interviewers and intermediaries were represented in all age groups and there was no difference in age composition, $X^2(3, N = 34) < 1, p = .984$, Cramer's $V = .068$. Two participants indicated high school as their highest level of education, 3 indicated college, 13 an undergraduate degree, and 16 a graduate degree. There was no significant difference between groups, $X^2(3, N = 34) = 3.28, p = .351$, Cramer's $V = .310$.

Fifteen of the interviewers had interviewed an average of 785.73 children each ($SD = 1065.71$); range = 30–3500, $Mdn = 500$. The remaining three interviewers did not provide a specific number in response to this question. All interviewers responded "No" to questions about intermediary experience. Only one intermediary had any prior interviewing experience and had interviewed three children. Of the 12 intermediaries who indicated they had provided third-party support in police interviews, three did not specify the number of times. The remaining nine had provided such support an average of 10.89 times ($SD = 11.29$) for children and 6.78 ($SD = 7.92$) times for adults. In response to a question about providing intermediary support in court, 10 of the 13 intermediaries who indicated having done so specified the number

Table 1. Education and training summaries.

	n	Interviewer		Intermediary		X ²	p	V
		Yes	No	Yes	No			
Have you ever provided third-party communication support in a police interview?	34	0 (0%)	17 (100%)	12 (71%)	5 (29%)	18.545	<.001	.739
Have you ever provided third-party communication support in court?	34	0 (0%)	17 (100%)	12 (71%)	5 (29%)	18.545	<.001	.739
Have you been trained as an intermediary?	34	0 (0%)	17 (100%)	10 (67%)	7 (41%)	14.167	<.001	.645
Have you been trained to interview vulnerable witnesses?	34	16 (94%)	1 (6%)	5 (28%)	13 (72%)	16.032	<.001	.677
Have you been trained in child development?	34	15 (88%)	2 (12%)	16 (94%)	1 (6%)	.366 ^a	.545	.104
Have you received training in speech and language disorders?	34	6 (35%)	11 (65%)	12 (71%)	5 (29%)	4.250	.039	.354
Have you received training in responding to trauma?	34	16 (94%)	1 (6%)	13 (77%)	4 (24%)	2.110 ^a	.146	.249
Have you received mental health training?	34	13 (76%)	4 (24%)	14 (82%)	3 (18%)	.180 ^a	.671	.073
Have you received cognitive disability training?	34	11 (65%)	6 (35%)	13 (77%)	4 (24%)	.567	.452	.129

a Expected cell count less than 5.

of times: an average of 8.90 times ($SD=14.26$) for children and 3.50 times ($SD=5.46$) for adults. Participants were further asked about various types of training they had received (see Table 1).

Materials

Child profiles and scenarios. The research team developed profiles for three children, in consultation with experts in autism research, speech pathology, and neurodivergent development based at an Autism Research Centre at Griffith University. The profiles characterized three female children, aged 6 (Ani), 12 (Maya), and 13 (Lorelei). Explicit diagnoses were not provided; instead, the profiles described the children's teacher-observed behaviors, vision and hearing, and scores on the Vineland Adaptive Behavior Scales-3 (Sparrow et al., 2016) and the Peabody Picture Vocabulary Test (Dunn and Dunn, 2007).

To verify the suitability of the scenarios, five experts in learning disability (with minimum qualifications at the graduate certification level) unassociated with the research and blind to study design and hypotheses read the three scenarios and judged the child's description to be a) typically developing, b) language delay, or c) other developmental delay. Fourteen (93%) of the 15 classifications (five experts \times three scenarios) matched the intended classification. One expert categorized the typically developing description as "developmental delay," because we described the child as scoring just outside the typical range for internalizing and externalizing behaviors. Based on the results of the pilot testing, we decided that the scenarios were suitable for our purposes.

The research team then developed three scenarios describing the child's living arrangements and information about a potential outcry or other evidence suggesting potential sexual or physical abuse (e.g., Case A: "...said Alex does rude things with her in the bath", Case B: "...she has been talking about 'taking naughty pictures' on her neighbor's phone", Case C: "...her clothing and hair are frequently dirty and recently the teacher identified some red flags for abuse"). The three profiles and three scenarios were

combined to produce nine possible vignettes (i.e., pairing each child profile with each abuse scenario). Each participant was offered three of these vignettes, depicting all three children in all three abuse scenarios. All supplementary materials, including these materials and a counterbalance table are available on the Open Science Framework (OSF).

Case recommendation sheets. The research team created a single-page document for participants to complete with any recommendations they had for interviewing each child. Participants were asked, “What recommendations/considerations/special measures would you suggest are needed in the police interview with this child? Please make a brief assessment (point form is fine) based on the scenario information you were given. We know that you have less information to go on than you typically would.” The page was otherwise blank.

Interview audio files. Three interview transcripts from actual cases (two containing sexual abuse allegations and one physical abuse) were heavily edited to create high-quality, moderate-quality, and low-quality versions of each one, and to embed 10 *problematic questions* within each. In high-quality interviews, 60% of the questions were open-ended, 20% were wh-type, and 20% were option-posing questions, whereas the respective percentages were 20, 60, and 20 for moderate-quality interviews, and 20, 20, and 60 for low-quality interviews. Importantly, children’s responses within each case were the same regardless of the question type, so as not to confound interviewer questioning behavior with child response.

The 10 problematic questions comprised five questions that could create memory errors (e.g., leading questions) and five that were linguistically problematic (e.g., involving ambiguous pronoun use). Within each abuse scenario (Case A, B, or C), the 10 problematic questions were the same. The audio files each contained 10 minutes of dialogue between an interviewer and child interviewee, read by actors. The actors were the same across levels of interview quality but differed across abuse cases (i.e., six actors played interviewer and interviewee across the two sexual abuse and one physical abuse case). Interview quality was counterbalanced across case scenarios and delivered in random order. Thus, in some order and combination, every participant was assigned to listen to a predominantly open-ended interview, a predominantly wh- interview, and a predominantly option-posing interview; with a typically developing 6-year-old child, a 12-year-old child with language delay, and a 13-year-old child with developmental delay; alleging sexual abuse by a step parent (A), sexual abuse by an acquaintance (B), and physical abuse by the mother’s boyfriend (C).

Demographic and knowledge surveys. Using LimeSurvey, the research team created a set of demographic questions that asked about the participants’ ages, gender identities, education, experience interviewing children and/or working as an intermediary, and training (see Table 1). The second page of the survey contained six questions assessing knowledge of interviewing and question types. The first four questions were in multiple-choice format with four or five options. The last two questions each contained a unique set of 10 prompts or statements. Question 5 asked participants to identify problematic questions (of any type) from a series and rephrase them as better questions, and question 6 asked participants to identify which of multiple prompts were open-ended. The first four questions had been used by the research team and had been answered by several hundred interview trainees. Item analysis of these questions suggests that the difficulty index was good, with questions passed by roughly 50% of trainees on average. All questions are provided on OSF.

Procedure

After participants gave consent, a research assistant scheduled their online videochat session and sent the three written vignettes to which the participant had been randomly assigned. Prior to the online session,

participants were asked to complete case recommendations for all three children. These were included to align with the preparation intermediaries would typically experience in the field. They were not scored or otherwise analyzed; however, all participants sent their recommendations to the researchers to confirm they had completed them. When the participants met individually with a researcher, they listened to the cases (each presented once, in one of the three possible interview qualities). As the audio played, the participant could stop the recording anytime they wanted to make a comment about a question. They verbally explained their reasons for stopping and any recommendations for rephrasing the question or other procedure (e.g., using a prop). At the end of each interview, before moving to the next, participants were asked a) "What were the most important concerns about this interview that you would have flagged as problematic, if any?" and b) "Has the witness in this case been able to give her best evidence? Please explain." The participants were not given a definition for this term because we were interested in observing natural variability in how it was interpreted. This session took about 1 h and was audio-recorded. After the conclusion of the videochat session, participants received a link to complete the demographic survey and knowledge quiz, which took about 15 min.

Coding

Responses to audio interview questions. Participants' responses during the audio interviews were transcribed. To ground their comments in the context of the interview, the interview prompt on which they were commenting was also denoted. Comments concerning problematic questions were scored as correct when the participant stopped the audio recording and explained the issue, suggested an alternative question which corrected the issue or simply identified the question as poor. Problematic questions were not scored as correct if the audio recording was not stopped or the participant stopped the audio and criticized a problem that was wholly unrelated to the central issue (e.g., tone of voice). Sometimes participants correctly identified the problem, but their rephrasing did not correct the fundamental problem with the question or was inappropriate (e.g., the suggestion involved using anatomical terms rather than the child's label or the rephrased prompt was leading).

The appropriateness of comments made by participants to any other questions were also coded. Comments were coded as appropriate or inappropriate based on their alignment with best-practice interviewing principles. For instance, when a participant's recommendation involved rephrasing a specific question to be non-leading and open-ended, comments were generally coded as appropriate (considered in the context of the interview). However, recommendations to rephrase a non-leading open-ended question as a specific narrow question were generally coded as inappropriate. Comments that reflected stylistic preferences or more subjective comments (e.g., criticisms to the interviewer's tone of voice) were not coded.

To identify the nature of inappropriate comments, coding criteria were created by one researcher who had experience training forensic interviewers, in consultation with other team members. The criteria were established based on the content of participants' responses and best-practice interviewing principles. Comments could be assigned more than one category code.

Coders identified 11 categories of inappropriate comments, listed alphabetically. *Clarify:* Recommends that the interviewer interrupts the narrative to clarify the child's labels (e.g., terms for genitalia). *Complex:* Recommends a question that involves complex concepts and/or language, such as long and multipart questions or jargon. *Leading:* The rephrasing of a prompt is leading when the initial question was not. *More detail:* Recommends eliciting additional descriptive detail (e.g., about the offender) at an inappropriate point in the narrative (i.e., the request would cause the child to deviate from their narrative about what happened) or involves eliciting an unnecessary level of detail about an object or aspect of the event (e.g., room layout) that is not relevant or central to the offense. *Open-ended question:* Appropriate open-ended questions are criticized as vague, or for some other reason. *Other best-practice principles:* Recommendation conflicts with other best-practice principles. For instance, the participant

recommends repeating a specific question when the child did not answer it initially. *Specific question*: Recommends a narrow specific question in place of an appropriate open-ended one. *Temporal specificity*: Recommends clarifying event frequency by asking a question that is unlikely to yield an accurate answer (e.g., “How many times did [x] happen?”) or criticizes best-practice approaches to inquiring about event frequency and facilitating recall of specific event occurrences. *Support*: Recommends contingent forms of support or suggests a break in the middle of the child’s narrative when there is no strong indication a break is needed. *Visual aids*: Recommends aids at an unsuitable point in an interview (i.e., during a productive narrative account), or for an unsuitable purpose (e.g., using the visual aid in an interpretive way or suggesting a timeline approach, which at least one study has found unhelpful for children; Zhang et al., 2019). See OSF for the coding manual.

Post-interview follow-up questions. The post-interview questions regarding participants’ main concerns with the interview and whether the child was able to give her best evidence were coded with similar categories to the above, with a few exceptions. First, the response to the “best evidence” question was coded as “No”, “Yes”, or “Mixed” when the participant did not explicitly give a yes or no response and their rationale included both positive and negative features of the interview. The remainder of the participants’ responses to these questions were coded as *Strengths* (positive interview features), *Problems* (negative interview features), and *Recommendations* (when a clearly actionable suggestion was made about what the interviewer should have done). Frequently a problem was reported along with an actionable remedy to that problem (e.g., “the pace was too quick. If the interviewer paused for longer after the child’s responses, she might have had more time to think”), in which case both a problem and recommendation were coded.

Knowledge quiz. Responses to the multiple-choice questions on the knowledge quiz were scored for accuracy automatically. The first four questions received one point each. Question 5 and 6 each had 10 sub-items. Thus, the maximum possible score was 24. For question 5, the question re-phrasings were coded as acceptable (1) or not (0). For question 6, each subitem received a code of (0) if inaccurate (either the participant did not identify an open-ended question as open-ended, or mis-identified a specific question as open-ended) and (1) if accurate.

Reliability

Responses to the problematic questions and all other comments made during the audio file were coded by Author 2 and a second researcher (postdoctoral fellow). Responses to the interview follow-up questions were coded by the first author and Author 2. Responses to the short answer questions on the knowledge quiz were coded by the first author and a research assistant (undergraduate student). All coders were blind to experimental conditions during coding, but only the postdoctoral and student coder were blind to study hypotheses. First, coders trained together on data from 2–5 participants. Following completion of the training phase, a random set of new participants’ data were selected for reliability coding.

For the problematic questions and all other comments made while listening to the audio files, the postdoctoral fellow coded a sample of 10 participants (8 of whom commented on three interviews and 2 of whom were only able to comment on two interviews within the session timeframe). Thus 28 transcripts (25% of the total number) were double-coded. Interrater reliability was calculated using percent agreement. Cohen’s kappa was not suitable because the categories were not mutually exclusive. Initially, agreement was not satisfactory (68.9%); however, there were some systematic differences in coding, particularly identifying the following categories: clarity, leading, specific questions, temporal specificity. The postdoctoral reliability coder did not have background in research related to the interviewing of children, so many of the concepts were new. After resolving differences via discussion and retraining, agreement was high (94%). Author 2 coded the remainder of the sample.

For the post-interview follow-up questions, 30 responses (27% of the total) were double-coded (10 per interview type). For the best evidence question, there was only one disagreement (93.6%). Agreement regarding the remaining categories was 86.4% *Strengths*, 85.3% *Problems*, and 86.9% *Recommendations*. Disagreements were resolved through discussion. Frequent sources of disagreement were simply that one coder had missed an item, or whether both a problem and a recommendation should be coded for a similar topic. Regarding the latter, although an early decision was made that both should be coded, they were ultimately collapsed after coding and reliability was complete. This occurred because participants frequently received a code for both problems and recommendations associated with the same concept. Author 2 coded the remainder of the sample.

Cohen's Kappa was used to determine whether the sub-items from questions 5 and 6 of the knowledge quiz were acceptable/accurate or not. Kappa was very good ($k = .86$). These were resolved through discussion and Author 1 coded the remainder of the sample.

Results

Due to time limitations, six participants (three interviewers, three intermediaries) were unable to complete the last interview they were assigned. For one participant it was the open-ended interview, for three participants the wh- interview, and for two participants the option-posing interview. The mean is used to replace their data only in the analyses where comparisons are made across interview types, to allow as much data as possible to be included.

In advance of analyzing the data, we checked that statistical assumptions were met and to ensure that there were no unanticipated effects of the specific study materials. Participants' scores on the 10 problematic interview questions did not differ by order of presentation of interview type, $F(2, 114) < 1$, $p = .686$, $n^2 = .007$, description of the child, $F(2, 114) < 1$, $p = .796$, $n^2 = .004$, or case features, $F(2, 114) = 2.63$, $p = .076$, $n^2 = .044$. When checking any influence of these variables on answers to the best evidence question (9 chi-squared analyses; corrected alpha = .005), there were no significant effects at corrected alpha levels. At a conventional alpha level, there was an effect of order in the option-posing interview only ($p = .014$; all other $ps \geq .287$). The option-posing interview received more "yes" responses to the best evidence question when it was played third than would be expected by chance. As this was the only significant difference, and interview order, child, and case were counterbalanced across the levels of each other, these variables are not considered further.

Responses to the problematic questions

We analyzed responses to the problematic questions in a 2 (group: interviewer, intermediary) \times 2 (problematic question type: language, memory) \times 3 (interview type: open, wh-, option-posing) mixed Analysis of Variance (ANOVA), the latter two factors within subjects. A full table of means can be found on OSF. There was a main effect of professional group, $F(1, 39) = 6.80$, $p = .013$, $n^2 = .149$: out of the maximum 30 questions (five of each type per three interviews), interviewers ($M = 15.34$, $SD = 6.61$) identified more problematic questions than intermediaries ($M = 10.55$, $SD = 5.14$). There was also a main effect of problematic question type [$F(1, 39) = 59.87$, $p < .001$, $n^2 = .606$]: all professionals identified questions that could cause problems due to language or communication issues ($M = 7.72$, $SD = 3.30$, max 15) more frequently than questions that could lead to memory errors ($M = 4.94$, $SD = 3.14$). Finally, there was a significant interaction between problematic question type and interview quality, $F(1, 78) = 6.06$, $p = .004$, $n^2 = .135$. Post-hoc tests (LSD) showed that questions that could lead to memory errors were identified more frequently in the option-posing ($p = .010$) and open-ended interview ($p < .001$) than in the wh- interview. No other post-hoc comparisons ($ps \geq .241$) or other effects in the ANOVA were significant ($Fs \geq .44$, $ps \geq .310$, $n^2 \geq .011$). Thus the hypotheses that participants would notice more problematic questions in the option-posing

interview than in the open-ended interview, and that interviewers and intermediaries would recognize different types of questions as problematic, were not confirmed.

Responses to other interview questions

Next, we looked at our coding of all possible responses (omitting responses to the 10 problematic questions). We counted the number of times the participants stopped the audio recording to make a comment. We classified what they said as appropriate or inappropriate. We entered tabulations of these into a 2 (group)×3 (interview)×2 (appropriateness) mixed ANOVA, the latter two factors within-subjects. All main effects and interactions were significant, $F_s \geq 10.41$, $p_s \leq .003$, $\eta^2_s \geq .211$; where necessary, we unpacked these effects with follow-up tests (LSD or simple effects analyses where applicable), as described below.¹ See Figure 1 for group means. Levene's tests indicated there were some violations of the assumption of homogeneity of variance, especially for the distribution of appropriate comments. Because of these violations, we ran non-parametric tests for analyses comparing the numbers of inappropriate and appropriate comments. The conclusions were identical with one exception, which we note below (for a description of these analyses and output, see OSF).

Overall, interviewers made more comments than intermediaries, though the frequency of comments differed across interview type: while both groups made a similar number of comments about the wh- and open interview ($p_s \geq .063$), interviewers made significantly more comments about the option-posing interview ($p < .001$). All participants made significantly more appropriate than inappropriate comments across all interview types ($p_s < .001$)², though this difference was especially pronounced in the option-posing interview. Finally, the three-way interaction indicated that the higher number of appropriate comments (compared to inappropriate comments) made about the option-posing interview, compared to the other interview types, was especially pronounced among interviewers. This analysis provided some partial support for the hypothesis that interview quality would affect participants' recognition of problematic questions.

Inappropriate comments

We explored the frequency with which each category of inappropriate comment was recognized by members of each of the professional groups. The number of participants who made at least one comment in each category was tallied (see OSF for details). We collapsed across interview type since the patterns were similar in each. The most common categories of unhelpful comments made by intermediaries were to use visual aids (65.1% of participants) and clarify terms (65.1%) in early stages of the interview before the child could give a complete account, followed by recommending specific narrow questions (60.9%), and then suggesting leading questions (47.8%). For interviewers, the most common categories were suggesting a leading (22.2%) or complex question (22.2%), following by clarifying terms (16.7%) and recommending a specific question early in the interview (16.7%).

Knowledge quiz

The maximum score for the knowledge quiz was 24. Due to a computer error, six intermediaries and one interviewer did not receive the quiz. Their data were not replaced. Interviewers scored higher ($M = 18.35$, $SD = 3.43$) than did intermediaries ($M = 14.09$, $SD = 2.53$), $t(32) = 4.13$, $p < .001$, Cohen's $d = 1.417$.

¹ When the test of sphericity was violated, the Greenhouse–Geisser correction was applied.

² A Friedman test (non-parametric) indicated that the number of appropriate and inappropriate comments was similar in the open interview, $\chi^2(2) = 3.270$, $p = .071$.

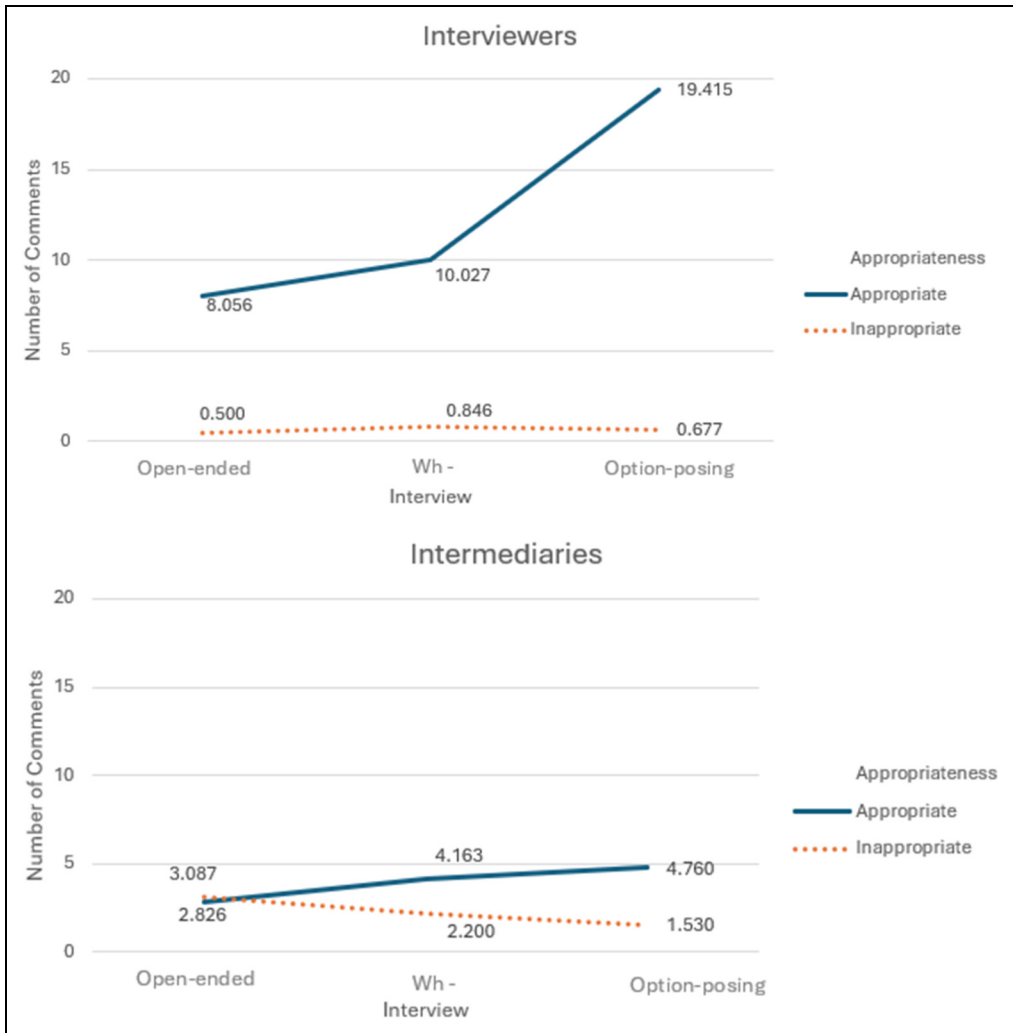


Figure 1. Appropriateness of rephrased questions by interviewers and intermediaries.

Scores on the knowledge quiz were positively correlated with participants’ scores on the 10 problematic questions in the open-ended interview $r(32) = .514, p = .002$, the wh- interview, $r(32) = .604, p < .001$, and the option-posing interview, $r(32) = .492, p = .003$.

Responses to best evidence question

Interviewers and intermediaries did not differ in response to the question as to whether the child gave her “best evidence” for the open-ended interview, $X^2(2, N=40) < 1, p = .739, V = .123$. Overall, participants did not perceive that the child’s best evidence was given: 40% said “no”, 45% were mixed, and only 15% said “yes”. These percentages were similar across respondent groups. They did differ, however, in their answers to the best evidence question for the wh- interview, $X^2(2, N=38) = 6.05, p = .049, V = .399$. Interviewers were more likely (and intermediaries less likely) than would be expected by chance to disagree that the child was able to give her best evidence. Over half the interviewers (56%) answered “No”

Table 2. Interview strengths and problems and associated recommendations by group and interview type.

	Open		Wh-		Option-posing	
	InV (18)	InM (22)	InV (16)	InM (22)	InV (17)	InM (21)
Strengths						
Detail (child)	33.3%	31.8%	12.5%	40.9%	17.7%	9.1%
Questioning-Open (interviewer)	44.4%	18.2%	12.5%	13.6%	5.9%	4.5%
Clarity (child)	11.1%	4.5%	12.5%	9.1%	17.7%	4.5%
Questions-Simple (interviewer)	–	13.6%	6.3%	9.1%	–	18.2%
Questions-Specific (interviewer)	–	–	–	4.5%	17.7%	13.6%
Rapport (either/both)	–	4.5%	6.3%	9.1%	5.9%	9.1%
Temporal specificity (either/both)	11.1%	9.1%	18.8%	9.1%	–	–
Problems/Associated Recommendations						
Temporal specificity lacking/Obtain	50.0%	63.6%	50.0%	40.9%	23.5%	63.6%
Questioning-Too complex/Simplify	27.8%	68.2%	37.5%	36.4%	23.5%	50.0%
Clarity lacking/Clarify	50.0%	45.5%	18.8%	59.1%	11.8%	40.9%
Questioning-Too specific/Ask more open-ended	66.7%	–	62.5%	18.2%	70.6%	22.7%
Questioning-Leading/Avoid	27.8%	27.3%	4.183.8%	13.6%	58.8%	36.4%
Narrative not exhausted before specific questions asked	44.4%	13.6%	62.5%	13.6%	64.7%	18.2%
Visual aids or props needed/Should be used	5.6%	31.8%	–	45.5%	–	36.4%
Child emotions/Should be managed	5.6%	18.2%	12.5%	13.6%	11.8%	27.3%
Interviewer should “check in” with child	–	13.6%	–	9.1%	–	9.1%
Interview lacks credibility	11.1%	9.1%	–	–	17.6%	–
Pace too quick/Slow pace	5.6%	13.6%	–	4.5%	5.9%	4.5%
Pre-planning of interview needed	–	4.5%	6.3%	9.1%	–	9.1%
Questioning-Not specific enough/Ask more specific	–	9.1%	–	4.5%	–	13.6%
Rapport lacking/Build rapport	–	–	6.3%	9.1%	5.9%	4.5%

Note: InV = interviewer, InM = intermediary. Number of participants who provided responses in parentheses.

to this question, while 25% had a mixed opinion and 19% said “Yes”. In contrast, only 18% of intermediaries said “No” to the question and the remainder were mixed (41%) or said yes (41%). The groups did not differ for the option-posing interview, $\chi^2 (2, N = 39) = 3.81, p = .149, V = .313$. Most participants answered “No” (59%) or were mixed (31%), with only 10% of participants saying “Yes” when asked whether the child gave her best evidence in the option-posing interview. Of interest is whether the responses differed across interview types. A 3 (interview type) \times 3 (response) chi-squared analysis indicated they did not, $\chi^2 (4, N = 117) = 9.23, p = .056, V = .199$.

Interview strengths and recommendations

Participant-identified interview strengths and recommendations are presented in Table 2. Owing to the large number of possible comparisons, we elected not to analyze these data statistically. Instead, we provide a general overview of important trends.

The most frequently mentioned strengths of the open-ended interviews were that the child was able to provide detail and the questions were open-ended. Interviewers were particularly likely to comment on the open-ended nature of the questions. A small number of intermediaries (but no interviewers) commented on the simple wording of the questions and the likelihood that rapport was established. Intermediaries’ and interviewers’ greatest concerns for the open-ended interview related to question complexity and lack of clarity (in general and specifically related to concepts of time and sequencing). Interviewers were

descriptively more likely than intermediaries to complain that the questioning was too specific (12 interviewers and 0 intermediaries); in contrast, two intermediaries felt it was not specific enough. Interviewers also tended to comment that the child's narrative account had not been exhausted before specific questioning commenced. Intermediaries were descriptively more likely to recommend visual aids (this was true across all levels of interview quality), that the interviewers slow their pace, to "check in" with the child to determine whether a break was needed and, relatedly, to manage the child's emotions more carefully.

In the wh- interviews, the percentage of participants identifying strengths was noticeably lower than for the open-ended interview. Yet intermediaries found that the child was able to give detail, and some interviewers perceived the temporal specificity of the child's account to be reasonably good. Interviewers and intermediaries differed noticeably in the recommendations they gave for the wh- interview. Nearly three-quarters of interviewers (but less than a fifth of intermediaries) suggested that the questions were too specific and should be more open-ended, and that the child's narrative was not exhausted. Intermediaries found that clarity of the child's evidence (including temporal detail) was lacking, and frequently recommended visual aids.

Few participants identified strengths in the option-posing interview. Three interviewers and three intermediaries lauded the specific nature of the questions (i.e., perceived that the interviewer got more information with narrow questions, even though the child's responses were constant across interview types), and four intermediaries appreciated the simple wording of the questions. Conversely, most interviewers felt that the questions were too specific (compared to less than a quarter of the intermediaries). Most intermediaries felt that the questions were too complex, and clarity was lacking. Both groups were descriptively more likely to perceive that the interview involved leading questions compared to the open-ended and wh- interviews.

Discussion

The present study aimed to compare the concerns about interviewer questioning identified by intermediaries and trained forensic interviewers during investigative interviews with children. Using a controlled experimental design, we observed the types of problems intermediaries and (trained) investigative interviewers raised while monitoring mock forensic interviews, to determine how the use of intermediaries might affect the quality of investigative interviews. That is, when interviewers have received adequate training in interviewing children, do intermediaries still provide added benefits to identifying problematic interview questions?

The findings of this study highlighted the divergent specialties of interviewers and intermediaries, with each focusing on different characteristics of the interview. Interviewers scored higher on the knowledge test than intermediaries and identified more problematic questions during the interviews. We anticipated that members of both groups would identify varying numbers of problems depending on the quality of the interviews being evaluated, but this hypothesis was only partially supported. Contrary to our second hypothesis, the two groups did not differ with respect to the types of questions (i.e., language vs memory issues) they identified as problematic.

In accordance with best-practice guidelines (e.g., Lamb et al., 2018), interviewers tended to focus on the degree to which open-ended questions helped the child give a narrative account. Intermediaries focused on whether the questioning involved clear scaffolding. For example, intermediaries tended to perceive the concrete structuring afforded in the wh- interview (e.g., "Where was Johnny?" "Where were you?") more positively than interviewers did; they were more likely than interviewers to report that the child had provided their best evidence in this interview. In all interview types, intermediaries' concerns and recommendations tended to focus on temporal specificity, clarity, and the use of visual aids. These differing emphases likely stem from the different goals of forensic interviewers and intermediaries. Forensic interviewers aim to enhance the accuracy and completeness of interviewees'

evidence. High-quality interview training involves a specialized focus on how different question types affect the accuracy and detail of interviewees' recall, given the high accuracy standards required in forensic and legal contexts (e.g., Brown and Lamb, 2015). Conversely, intermediaries (e.g., speech-language pathologists) seek to enhance communication clarity, with a comparatively reduced focus on recall accuracy, which is generally not emphasized in their professional training. While these differing aims are complementary in some respects, at other times they may conflict. For instance, questions that are clearly communicated and simple may be closed or leading, undermining children's recall accuracy and productivity (Powell and Snow, 2007). Overall, the intermediaries' responses regarding strengths and recommendations for interviewer questioning were very diverse. These findings, coupled with existing research describing differences among intermediaries (Henry et al., 2017), indicate that standardization in intermediary training and certification is needed.

The present research yields insight about when intermediaries might provide added benefit in forensic interviews. While in some jurisdictions children under a certain age are automatically granted an intermediary in court (Henderson, 2015), and some stakeholders have urged equal access to intermediaries for all vulnerable individuals (Gallagher et al., 2024), this may not be the best use of limited public resources (Fairclough et al., 2023). Rather, intermediaries should be engaged when they offer a distinct contribution that extends beyond the skillset of trained interviewers. The present findings suggest that when interviewers are well-trained, they can handle questioning issues with most children independently, so intermediaries are not routinely needed.

Reducing the demands on intermediaries in jurisdictions where they are indiscriminately employed when young children are interviewed would allow them to be utilized more selectively. This could include cases involving witnesses with communication needs that exceed the basic skillset of most interviewers. In the present study, the intermediaries' attention to scaffolding suggested that they could be particularly helpful in cases involving interviewees with more pronounced communication difficulties, such as severe autism or language delay (e.g., Bearman et al., 2019; Hoff et al., 2022; Maras et al., 2020). Moreover, intermediaries have noted that police interviewers sometimes do not understand the need to modify the way they question certain respondents, indicating the potential for intermediaries to aid communication in these cases (e.g., someone with severe autism; Henderson, 2015).

Intermediaries' contributions in specific cases could be enhanced through training in best-practice questioning. The intermediaries in the current study were equipped to identify key issues in forensic interviews: like the interviewers, they made significantly more appropriate than inappropriate comments and tended to agree with interviewers about whether the child had provided their best evidence following the open and closed interviews. The ability of both intermediaries and interviewers to identify problematic questions differed minimally across interview quality, and while intermediaries identified fewer problematic questions than interviewers overall, their relative ability to identify language and cognitive issues with questions was similar to that of interviewers. However, intermediaries also made suggestions that conflicted with best-practice recommendations for interviewing children. For instance, around 50% of intermediaries (as opposed to 20% of interviewers) rephrased a non-leading interviewer prompt in a leading way at least once. Additionally, a high proportion (60%) of intermediaries suggested replacing at least one effective open-ended prompt with a more specific question.

The training of intermediaries is heterogenous. In the current study, not all participants in the intermediary group had prior experience in that role. However, all were eligible to work as intermediaries in their jurisdictions, reflecting real-world settings where intermediaries often lack standardized training, and where even any mandated training is minimal (e.g., Cashmore and Shackel, 2018; Department of Justice Northern Ireland, 2015; Hoff et al., 2022). One relatively extensive form of training is offered in England and Wales, where intermediaries receive seven days of training over a two-week period, and they must complete five assessments before receiving accreditation from the Ministry of Justice. One of those assessments involves "amending cross examination questions as appropriate to a selected case" (Collins and Krähenbühl, 2020, p. 376) but it is not clear whether they are trained in best-practice

questioning strategies. Anecdotally, during a post-participation conversation, one Canadian intermediary participant opined how helpful it would be for intermediaries to receive the training that *interviewers* receive.

Because intermediaries' abilities to rephrase questions appropriately enhances witnesses' abilities to provide the best possible evidence, basic training in best-practice interviewing could expand the value that intermediaries provide. Indeed, intermediaries have stated that they sometimes feel inadequately equipped to meet the technical demands of their role (Kearns et al., 2024), and legal practitioners have noted that intermediaries did not offer advice that exceeded the practitioners' own questioning skills (Hoff et al., 2022). To make meaningful contributions, intermediaries need specialized expertise in adapting and rephrasing questions to address the unique needs of vulnerable witnesses (Hoff et al., 2022), and this skill could be refined through training in best-practice interviewing. Such skills would be helpful not only in forensic interview contexts but in the courtroom as well (Powell et al., 2022).

Limitations of the present research

We assessed what intermediaries knew about questioning in forensic interviews, but intermediaries serve a variety of functions outside police interviews (see Collins and Krähenbühl, 2020). The emphasis of the present research was not on other potential ways in which intermediaries support vulnerable individuals in the criminal justice system. Further, this research was conducted using best-practice questioning guidelines as a framework, focusing on techniques most likely to elicit reliable evidence. We classified participant recommendations (as appropriate or inappropriate) in accordance with this framework. Had we prioritized what would help the child understand questions and communicate most clearly, the classification of recommendations might have differed (e.g., well-structured closed questions might have been considered appropriate).

Of note, the study was underpowered due to the inherent difficulty in recruiting and retaining skilled professionals for research. This means that the results should be interpreted with caution. Further, practicing interviewers and intermediaries have varying levels of skill, experience, training, and effectiveness when working with vulnerable populations. In our sample, the interviewers had a great deal more experience than the intermediaries. Interviewers had, on average, conducted over 700 interviews while intermediaries had provided support, on average, 10 times. Practice and experience may enhance the ability of both interviewers and intermediaries to identify problematic questions when working with vulnerable populations. The sample size coupled with large differences in experience documented between intermediaries and interviewers thus limit the generalizability of the findings but provide important insight into how intermediaries and highly skilled interviewers identify problematic questions.

Conclusion

The use of intermediaries has clear benefits but is not a "silver bullet" (SALRI, 2021, xxi, xli, 20). The interviewers in the current study demonstrated strengths in implementing best-practice principles for interviewing children, whereas the intermediaries appeared to identify issues more closely related to question scaffolding. These differing specialties indicate that, while well-trained interviewers are equipped to avoid problematic questioning without additional assistance, intermediaries could provide valuable assistance in specialized cases where interviewees require additional communication guidance (e.g., Bearman et al., 2019).

The preliminary recommendations stemming from the present research are two-fold: First, we suggest that it may not be necessary to employ intermediaries in all interviews with children as a matter of routine. Secondly, interviewers should receive high-quality training about best-practice questioning strategies to enhance their abilities to identify problematic questions and rephrase them effectively. This knowledge is valuable for maximizing resources dedicated to supporting vulnerable individuals in the criminal justice

system. Reserving the use of intermediaries for cases in which they are most needed for communicative assistance, and ensuring that interviewers themselves are well-trained, will help to prevent miscommunication with children in forensic interviews.

Acknowledgement

We would like to thank Emily Denne for her assistance with the literature review, and members of the Autism Centre of Excellent (Griffith University) for their assistance with the child profiles.

Author note

Note. The data and materials that supported this research are available on the Open Science Framework: https://osf.io/5tibr/?view_only=d529d22d911d47549f0a6ae6d0d50e95.


Funding


The authors disclose receipt of the following financial support for the research, authorship, and/or publication of this article: The research was funded by the Law Foundation of South Australia Inc as part of a project by the South Australia Law Reform Institute (SALRI) based at the University of Adelaide to examine the “Communication Partner” (or “intermediary” role) in South Australia.


Declaration of conflicting interests


The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.


ORCID iDs

Sonja P. Brubacher  <https://orcid.org/0000-0002-2422-9913>

Sarah L. Deck  <https://orcid.org/0000-0001-9213-0352>

David Plater  <https://orcid.org/0000-0002-9373-4980>

Michael E. Lamb  <https://orcid.org/0000-0002-6792-3526>

Martine B. Powell  <https://orcid.org/0000-0001-5092-1308>

References

- Bearman M, Brubacher SP, Timms L, et al. (2019) Trial of three investigative interview techniques with minimally verbal adults reporting about occurrences of a staged repeated event. *Psychology, Public Policy, and Law* 25(4): 239–252.
- Brown DA and Lamb ME (2015) Can children be useful witnesses? It depends how they are questioned. *Child Development Perspectives* 9(4): 250–255.
- Brubacher SP, Shulman EP, Bearman MJ, et al. (2022) Teaching child investigative interviewing skills: Long-term retention requires cumulative training. *Psychology, Public Policy, and Law* 28: 123–136.
- Cashmore J and Shackel R (2018) *Evaluation of the Child Sexual Offence Evidence Pilot: Final Outcome Evaluation Report*. Sydney, Australia: University of New South Wales.
- Collins K and Krähenbühl S (2020) Registered intermediaries’ assessment of children’s communication: An exploration of aims and processes. *The International Journal of Evidence and Proof* 24: 374–395.
- Cooper P and 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(4): 351–370.
- Cooper P and Wurtzel D (2014) Better the second time around? Department of Justice registered intermediaries schemes and lessons from England and Wales. *Northern Ireland Legal Quarterly* 65(1): 39–61.

- Department of Justice Northern Ireland (2015) *Registered intermediaries schemes pilot project: Post-project review* (January 2015). Available at: www.justice-ni.gov.uk/publications/northern-ireland-registered-intermediaries-schemes-pilot-project (accessed 17 July 2025).
- Dunn LM and Dunn DM (2007) *Peabody Picture Vocabulary Test, Fourth Edition*. San Antonio, TX: Pearson.
- Fairclough S, Taggart J and Backen P (2023) Lost in translation? Perceptions of defendant intermediaries in the legal profession. *Criminal Law Review* 7: 440–461. Available at: https://pureadmin.qub.ac.uk/ws/portalfiles/portal/464378628/CLR_Final_submission_SF_JT_PB_2023_ACCEPTED_VERSION_.pdf (accessed 17 July 2025).
- Gallagher AL, Cusack A, Kearns Á, et al. (2024) Establishing stakeholder priorities for advancing the role of the intermediary in the justice system: Findings from a consensus study. *International Journal of Evidence and Proof*. Epub ahead of print 27 December 2024. DOI:10.1177/13657127241306173.
- Hanna K and 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 & Proof* 22(4): 411–427.
- Henderson E (2015) “A very valuable tool”: Judges, advocates and intermediaries discuss the intermediary system in England and Wales. *The International Journal of Evidence and Proof* 19: 154–171.
- Henry LA, Crane L, Millmore A, et al. (2021) Intermediaries and cross-examination resilience in children: The development of a novel experimental methodology. *Applied Cognitive Psychology* 35(6): 1387–1404.
- Henry LA, Crane L, Nash G, et al. (2017) Verbal, visual, and intermediary support for child witnesses with autism during investigative interviews. *Journal of Autism and Developmental Disorders* 47: 2348–2362.
- Hoff S, Powell M and Plater D (2022) When good intentions are not enough: Professionals’ perceptions of the South Australian communication partner scheme. *Australian Journal of Social Issues* 57: 970–986.
- Howard K, McCann C and Dudley M (2020) What is communication assistance? Describing a new and emerging profession in the New Zealand youth justice system. *Psychiatry, Psychology and Law* 27(2): 300–314.
- Kearns Á, Clarke D, Cusack A, et al. (2024) Intermediaries in the justice system for people with communication disability: Enacting sustainable development goal 16 in Ireland, Northern Ireland, and New Zealand. *International Journal of Speech-Language Pathology* 25(1): 141–146.
- Krähenbühl S (2011) Effective and appropriate communication with children in legal proceedings according to lawyers and intermediaries. *Child Abuse Review* 20(6): 407–420.
- Lamb ME, Brown DA, Hershkowitz I, et al. (2018) *Tell Me What Happened: Questioning Children About Abuse*. 2nd ed. Hoboken, NJ: Wiley-Blackwell.
- Lamb ME, Malloy LC, Hershkowitz I, et al. (2015) Children and the law. In: ME L and Lerner RM (eds) *Handbook of Child Psychology and Developmental Science: Socioemotional Processes*. Hoboken, NJ: John Wiley & Sons, Inc, 464–512.
- Maras K, Dando C, Stephenson H, et al. (2020) The witness-aimed first account (Wafa): A new technique for interviewing autistic witnesses and victims. *Autism* 24(6): 1449–1467.
- Pipe M-E, Orbach Y, Lamb ME, et al. (2013) Do case outcomes change when investigative interviewing practices change? *Psychology, Public Policy, and Law* 19(2): 179–190.
- Plotnikoff J and Woolfson R (2015) *Intermediaries in the Criminal Justice System: Improving Communication for Vulnerable Witnesses and Defendants*. Chicago, IL: Policy Press.
- Powell MB, Bowden P and Mattison M (2015) Stakeholders’ perceptions of the benefit of introducing an Australian intermediary system for vulnerable witnesses. *Australian & New Zealand Journal of Criminology* 48(4): 498–512.
- Powell MB, Goodman-Delahunty J, Deck SL, et al. (2022) An evaluation of the question types used by criminal justice professionals with complainants in child sexual assault trials. *Journal of Criminology* 55(1): 106–124.
- Powell MB and Snow PC (2007) Guide to questioning children during the free-narrative phase of an investigative interview. *Australian Psychologist* 42: 57–65.
- R v B (2010) EWCA Crim 4, para 42. Lord Judge, the Lord Chief Justice of England and Wales.
- SALRI (2021) *Providing a Voice to the Vulnerable: A Study of Communication Assistance in South Australia*. Adelaide: South Australian Law Reform Institute.

- South Australia (2016). Summary offences regulations, Version 6.11.2023. Available at www.legislation.sa.gov.au (accessed 17 July 2025).
- Sparrow SS, Cicchetti DV and Saulnier CA (2016) *Vineland Adaptive Behavior Scales, third edition (Vineland-3)*. San Antonio, TX: Pearson.
- Taggart J (2021) Intermediaries in criminal proceedings - a role in need of clarification? *Archbold Review* 1: 6–9.
- Zhang HH, Roberts KP and Teoh Y-S (2019) Children's recall and source monitoring of a repeated event using a timeline as an interview aid. *Applied Cognitive Psychology* 33: 176–187.