

Witness Interview Training: a Field Evaluation

Sarah MacDonald, Brent Snook & Rebecca Milne

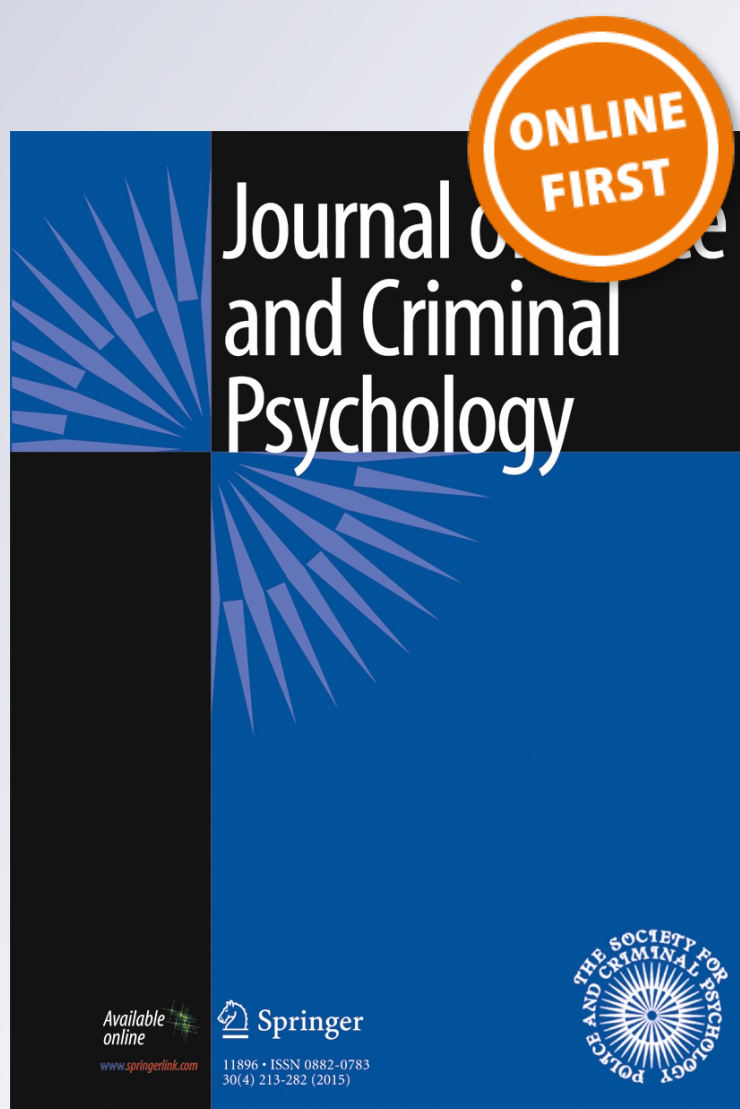
Journal of Police and Criminal Psychology

The Official Journal of the Society for Police and Criminal Psychology

ISSN 0882-0783

J Police Crim Psych

DOI 10.1007/s11896-016-9197-6



Your article is protected by copyright and all rights are held exclusively by Society for Police and Criminal Psychology. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Witness Interview Training: a Field Evaluation

Sarah MacDonald¹ · Brent Snook¹ · Rebecca Milne²

© Society for Police and Criminal Psychology 2016

Abstract The effect of a witness interviewing training program on interviewing performance in actual investigations was examined. Eighty interviews, conducted by police officers in one Canadian organization, were coded for the presence of 38 desirable practices. Results showed that, in general, trained interviewers outperformed their untrained counterparts. Specifically, there was a large improvement in *engage and explain* behaviors ($d = 1.65$), a moderate improvement in *account* behaviors ($d = 0.54$), and a large improvement in *closure* behaviors ($d = 0.90$). Trained interviewers also used more open-ended questions and fewer leading questions. The implications of the findings for transferring interviewing skills from the classroom to the field are discussed.

Keywords Witness interviewing · Training · Police · Evaluation

The quality of criminal investigations, and the legal proceedings that follow, are linked inexorably to how well interviews with adult witnesses and victims (hereafter referred to as witnesses) are conducted because interviews are a major conduit through which probative evidence is obtained (Kebbell and Milne 1998; Milne and Bull 2003; Sanders 1986). The fact that interviewers – especially untrained interviewers or trained interviewers who are not

afforded supervision and feedback – tend to use a number of inappropriate behaviors (e.g., asking leading questions, see Griffiths et al. 2011; Snook and Keating 2010) leads naturally to concerns about the quality of the information that forms the basis of legal decisions. Fortunately, evidence-based interviewing techniques are available to police organizations (e.g., cognitive interviewing; Geiselman et al. 1985), and there is evidence from laboratory settings that people can be trained to apply these techniques (Köhnken et al. 1999; Memon et al. 2010). However, relative to laboratory studies, only a handful of studies have examined how well adult witnesses interview training gets transferred into the field.

The few field studies that have assessed the effectiveness of adult witness training suggest that the transference of skills is difficult. For instance, Memon et al. (1994) investigated how well police officers' used various aspects of the CI after a four hour training course that taught officers about the importance of interviewing and general principles of communication (i.e., greet, explain, rapport, mutual activity, and closure). Memon and her colleagues found that post-training performance was inadequate. They found that officers had difficulty explaining what was required from witnesses and the officers did not use all available CI techniques (see also Dando et al. 2008; Kebbell and Milne 1998 for similar findings). Clifford and George (1996) also analyzed witness interviews from criminal investigations and discovered that none of the interviewers applied all four CI memory enhancement techniques after a four-day training session. Specifically, they found that the techniques of *context reinstatement*, *change perspective*, *change of temporal order*, *refrain from editing anything out*, and *working hard to recall the details*, were utilized in less than 10% of interviews after training.

Evaluations of suspect interview training have also reported difficulties in improving interviewing behavior. For

✉ Brent Snook
bsnook@mun.ca

¹ Department of Psychology, Science Building, Memorial University of Newfoundland, St. John's, NL A1B 3X9, Canada

² Centre of Forensic Interviewing Institute of Criminal Justice Studies, University of Portsmouth, Portsmouth, UK

instance, Griffiths and Milne (2006) found that officers who attended a three-week course on interviewing suspects used some practices (e.g., delivery of legal requirements) more than untrained officers, but that other evidence-based interviewing practices were not utilized more frequently (e.g., type of questions asked, sequence of questioning, and topic structure). Their *post hoc* classification of the interviewing behaviors determined that officers were able to implement simple skills (i.e., delivering legal rights to suspect) but struggled to employ the seemingly more complex skills (i.e., structuring the areas of the interview). Similarly, Walsh and Milne's (2008) analysis of a sample of suspect interviews found that, with the exception of following some legal and procedural steps (e.g., providing legal rights), there were modest increases in many desirable interviewing behaviors after training. More recently, Clarke et al. (2011) found that there were no meaningful improvements in interviewing behavior for a sample of officers who had received on-the-job PEACE interview training (PEACE being a mnemonic for the phases of an interview; planning and preparation; engage and explain; account; closure; a model of interviewing adopted all over the world by police officers). They found that the trained investigators administered only four out of 14 desirable interviewing behaviors (i.e., keeping on relative topic, deals with difficulty, encouraging an account, and appropriate structure/sequence).

Obtaining behavioral changes in child interviewing skills has also been notoriously difficult to achieve (Powell et al. 2005; Warren et al. 1999; Aldridge and Cameron 1999). For instance, evaluations of child interviewing practices in the UK – after revisions were made to national guidelines regarding best practices (e.g., Memorandum of Good Practice, 1999; Achieving Best Evidence, 2011) – have shown that the relatively simple techniques, such as building rapport and explaining ground rules, tend to be observed in police interviews (see Davies et al. 2016; Griffiths and Milne 2006; Walsh and Milne 2008 for an overview of those evaluations). These evaluations (e.g., Griffiths and Milne 2006) also found that providing written guidance on how interviews ought to be conducted (sometimes referred to as general principles training) does not tend to lead to complex skills, such as using open-ended questions, being transferred to the field. Despite the challenges, child interviewing research has shown that adherence to scripts, along with the provision of written feedback to interviewers, can produce large gains in the transference of desirable skills to the field (e.g., see NICHD protocol; Lamb et al. 2008; Price and Roberts 2011).

The Current Study

Evaluations of adult witness interview training courses have tended to examine a select few behaviors that are fundamental

to interviewing. Beyond an examination of question types and the implementation of CI mnemonics, limited research has assessed the effect that witness interview training has upon interviewers' performance in actual police interviews. As a consequence, the goal of this research was to assess the ability to transfer a wide range of desirable behaviors from training to actual investigations. Across different types of interview training (i.e., suspect, child, and witnesses), the take-home-messages are that transferring evidence-based practices to the field is challenging, but possible. We therefore predict desirable skills will be observed more frequently in interviews conducted with trained interviewers compared to those with untrained interviewers. Research also suggests that the magnitude of the improvements in complex behaviors is much smaller than it is for simpler skills. Therefore, we predict that the largest improvements will be observed for the seemingly simpler skills (i.e., those that are conducive to the use of scripts/checklists such as engage and explain behaviors) and, but to a lesser extent, more complex skills (i.e., dynamic skills such as asking questions). A wealth of child interviewing research has also made it clear that the failure to provide extensive feedback will result in a decline in learned skills over time. Given the lack of a formal feedback process for the participating police organization, we predict that there will be a negative relationship between the number of skills observed in interviews conducted by trained interviewers and time since their training ended.

Method

Design

It was requested that the following witness interviews be submitted from a Canadian police organization: 20 interviews conducted before PEACE training which commenced within the organization prior to 2008 (pre-experimental), 20 interviews conducted at least one week from the last day of any PEACE training course conducted after 2008 (post-experimental), 20 interviews conducted by individuals during the same general time period (2004–2008) as the pre-experimental interviews (pre-control), and 20 interviews conducted by individuals who did not participate in training during the same general time period (2008–2011) as the post-experimental interviews (post-control). The latter condition was included to control for instances where investigators may have learned some aspects of PEACE training by observing interviews conducted by trained colleagues.

Sample

A convenience sample of 80 police interviews (12 videotapes, 36 transcripts, 30 files, and 2 transcripts accompanied by

audio files) with adult witnesses was obtained from a Canadian police organization. The interviews were collected through requests made by an Inspector for police officers to submit a sample of their interviews.

Each transcript consisted of a verbatim written account of an audiotaped interview. Videotapes consisted of audio and video recording of the interview, and audio files only consisted of audio recording. Videos, transcripts, and audio files were extracted from the population of interviews conducted by the criminal investigation division of the organization. The interviews occurred between 2003 and 2010, with 1.3% occurring in 2003, 17.5% occurring in 2004, 17.5% occurring in 2006, 7.5% in 2007, 5.0% in 2008, 33.8% in 2009, and 17.5% in 2010. The interviews covered a wide range of offences; 45% of the interviews pertained to the investigation of assault, 19% pertained to sexual assault, 9% to uttering threats, 9% to homicide, 4% to armed robbery, and 1% to each of the following: possession of child pornography, trespassing, missing person, attempt to lure a child, and robbery. The offence under investigation was not made explicit in the remaining 9% of cases. The interviewer(s) and interviewee were the only people present in 49% of the interviews.

A total of 32 different officers conducted the interviews. All interviewers in each condition were Caucasian. In the pre-experimental condition, 80% of the interviewers were men, 95% were Constables, the mean age of the primary interviewer was 46 years ($SD = 3.70$), and had an average of 24 years ($SD = 4.26$) of policing experience. In the post-experimental condition, 80% of interviewers were men, 95% were Constables, the mean age of the primary interviewer was 41.25 years ($SD = 5.09$), and had 14.90 years ($SD = 8.25$) of policing experience. In the pre-control condition, 80% of the interviewers were men, 75% were Constables, the mean age of the primary interviewer was 43.45 years ($SD = 3.47$), and had 19.20 ($SD = 4.39$) years of policing experience. In the post-control condition, 65% of interviewers were men, 95% were Constables, the mean age of the primary interviewer was 41.60 ($SD = 3.47$), and had 19.20 years ($SD = 4.39$) of policing experience.

Materials and Procedure

Investigators who participated in PEACE training attended a two-week introductory training course that was designed to educate all interviewers working in a criminal investigations division about desirable interviewing practices (referred to in the UK as a Tier 2 PEACE course). The training took place on a full time basis (seven hours per day) over the period of ten consecutive weekdays. The training was co-administered by an investigative interviewing advisor and the second author, both of whom were trained previously on the PEACE model of interviewing in the UK. All investigators were provided with lectures (including discussions) over a minimum of a

four-day period on all aspects of best witness interviewing practices, which covered the content of several texts on investigative interviewing (e.g., Fisher and Geiselman 1992; Milne and Bull 2003; NSLEC 2004; Shepherd 2007). The lectures included information on the principles of memory and cognition, rapport building, active listening, communication fundamentals, exchanging expectations, transferring control of the interview, questioning skills, short-hand note-taking, overcoming witness limitations, controlling witness anxiety, increasing witness confidence, inducing detailed descriptions (four mnemonics), and witness compatible questioning.

Interspersed with the lectures were practice interviews (ranging from 30 minutes to 1 hour) with actual witnesses, where interviewers were provided with a checklist of behaviors that they could reference as needed during interviews and immediate verbal feedback was provided from the trainers. The practice sessions followed a scaffolding approach that developed interviewing skills through the following five discrete stages: Engage and Explain; Questioning Skills; Note-Taking; Account (applying mnemonics); and Closure. Each of the stages required the interviewers to apply the learned principles in mock interviews, and use a checklist of desirable behaviors during those interviews. The non-interviewing investigators were required to watch all practice interviews and provide peer feedback (using the checklist).

Coding Procedure

A total of 10 demographic and context variables were coded. They included: the date and time of the interview; the type of crime; the number of people present; the length of the interview; whether a KGB warning (i.e., a form outlining the consequences of a witness lying to the police) was administered; and the age, gender, rank, and years of experience of the primary interviewer were collected. The age and years of experience variables were provided by the participating police organization and not coded from the interviews.

The content of the training course was used to devise a 38-item guide for coding the interviews. A total of 14 “engage and explain” variables were coded. The behaviors that are recommended in this interview stage are designed to decrease the interviewee’s anxiety and uncertainty of the process by creating an environment where people feel comfortable providing information about witnessed events. The interviews were coded for whether or not the interviewer: greeted the interviewee in a polite and professional manner; established the interviewee’s preferred name; asked the interviewee to call him/her by their first name; identified others present in the room if there were any (and explained their role); built adequate rapport; encouraged questions; explained the outline of the interview structure; identified the time and date of the

interview; established the purpose of the interview; asked the interviewee why they think they are being interviewed; established the interviewee's needs; and explained the routines and expectations of the interview (e.g., audio/video recording, not interrupting one another).

A total of 18 "account" variables were collected. The variables coded included whether or not the interviewer: attempted to set up one of the four CI mnemonics and, if they did it properly (i.e., did the interviewer use the correct instructions), which technique was attempted, asked for a free narrative, summarized the free narrative, passed to the second interviewer (if applicable), avoided topic hopping, the percentage of talking time by the interviewer, whether they followed the 80/20 talking rule, avoided interrupting the interviewee, used appropriate pauses, and avoided using jargon.

Also included in the 18 account variables was the type of questions asked by interviewers. Specifically, the number of closed yes/no, leading, multiple, forced-choice, open-ended, and probing questions were coded (see Griffiths et al. 2011; Snook et al. 2012). Closed yes/no questions involve those that are restricted to having a *yes or no* response and are often asked at random points in the interview (e.g., Was her shirt blue?). Leading questions are those that actually suggest the answer to the witness (e.g., you witnessed the crime, right?). Multiple questions are simply instances when an interviewer asks more than one question at a time (e.g., where did you go, what did you do, and when did you go home?). Forced choice questions involve forcing a witness to choose between a limited number of possible options for answering (e.g., was the color of the car blue or black?). Lastly, open-ended questions tend to be those beginning with *tell, explain, or describe*, and probing questions are those that are asked when not enough information is obtained using open-ended questions (such as "who" or "what", etc.).

Six variables pertaining to the behaviors that comprise the "closure" phase of an interview were coded. In particular, variables coded included whether or not the interviewer: gave a summary of what was said, provided their contact information, explained what will happen after the interview, recorded the date and time of the interview, and provided a professional closure.

The behaviors comprising each interview phase (i.e., engage and explain, account, and closure) were combined to create a single score for that phase. Each score consisted of a percentage of behaviors observed in each interview. Percentage was used because the total number of behaviors that could have been exhibited by an interviewer varied depending on the situation that was presented to the interviewer. For example, in some instances, the ability of the primary interviewer to introduce the second interviewer was not available because a second interviewer was not present in the interview room. Due to the differences in interview format,

some variables were coded differently for each interview. For example, interviews in audio and video format were coded for witness talking time (and subsequent adherence to the 80/20 rule) by examining the number of minutes and seconds spoken by the witness and interviewer. Witness talking time was calculated for transcripts by tallying the number of words spoken by the witness. Interruptions were only coded in audio-recorded interviews.

Inter-Rater Reliability

Coding agreement of the variables was assessed by having an independent researcher code 20% of the sample ($n = 16$), which was selected randomly. The independent coder was provided with a 2-hr training session that consisted of the structure and content of the coding guide and dictionary as well as the practical aspects of coding the interviews. Additionally, the coder participated in a practice session that covered the coding of two interviews before beginning to code the actual interviews. Any confusions pertaining to the task were resolved before inter-rater reliability commenced. Reliability for question types and all coded variables was analyzed with Kappa, and any variables with low Kappa values ($< .60$) were deleted. The following five engage and explain variables were deleted: identification of others (and explained their role); built adequate rapport; identification of time; establish purpose of the interview. One account behavior was deleted (i.e., avoided topic hopping). For the remaining variables, Kappa ranged from 0.60 to 1.00 and the overall value was 0.75. The overall average correlation for eight continuous variables was $r = 0.77$. All subsequent results are based on the analyses of variables found to have high inter-rater reliability.

Results

Engage and Explain Phase

The interviews conducted by trained officers (i.e., those in the post-experimental condition) contained more of the 9 engage and explain behaviors than from interviews conducted by untrained interviewers. Trained interviewers tended to explain the routines and expectations ($n = 14$), outline the route map ($n = 12$), *ask* the witness about their understanding of the purpose of the interview ($n = 12$), and establish the witness' needs ($n = 10$) more than untrained officers. One behavior that did not change much as a result of the training was encouraging witnesses to ask questions.

On average, the percentage (with 95% CIs) of the engage and explain behaviors observed for interviews conducted by trained interviewers was 48.33% ($CI = 35.10, 61.65$), which is larger than the means for the pre-experimental 13.89%, $CI = 10.16, 17.61$), pre-control

(13.33%, $CI = 11.19, 15.46$), and post-control groups (15.56%, $CI = 11.64, 19.47$), respectively. Independent sample t -tests, and more importantly, effect sizes confirmed that the mean percentage of behaviors exhibited during the engage and explain phase was higher for interviews conducted by trained interviewers than those in the pre-experimental condition, $t(38) = -5.22, p < .001, d = 1.65$; post-control condition, $t(38) = 4.94, p < .001, d = 1.56$, as well as the pre-control condition, $t(38) = 5.43, p < .001, d = 1.72$. The average effect size for comparisons made between the trained group and three untrained groups was $d = 1.65$ ($CI = 1.55, 1.73$).

Account Phase

Interviews conducted by trained interviewers (the post-experimental group) exhibited more of the 11 account behaviors than in interviews in the other three conditions. The behaviors exhibited most often by trained interviewers were: attempting to set up a cognitive interview ($n = 8$) and avoidance of interruptions ($n = 13$). When interviewers attempted to set up a CI, those in the pre-experimental, pre-control, and post-control condition only attempted to use the report everything instruction, while interviewers in the post-experimental condition attempted to use the report everything instruction on three occasions and mental reinstatement of context on five occasions.

The mean percentage (with 95% CIs) of account behaviors for the post-experimental group was 72.50% ($CI = 64.09, 80.90$), which is larger than percentages from the pre-experimental (58.75%, $CI = 50.34, 67.16$), pre-control (60.00%, $CI = 51.59, 68.41$), and post-control conditions (67.50%, $CI = 59.09, 75.91$) groups, respectively. The effect sizes (and independent t -tests) demonstrated that trained interviews contained more required behaviors to obtain an account compared to those in the pre-experimental condition, $t(38) = -2.37, p < .05, d = .75$. Although there was no statistically significant difference in the behaviors observed in the account phase of the interview between the trained interviewers and either of the post-control, $t(38) = 0.92, p > .05$, and pre-control conditions, $t(38) = 1.82, p > .05$, there were small ($d = .22$) to moderate ($d = .57$) improvements in account behaviors, respectively. The effect size for comparisons made between the trained group and three untrained groups was $d = 0.54$.

Each of the four groups were also analyzed for the proportion of six question types: leading, multiple, forced-choice, open-ended, probing, and closed yes/no questions. The mean percentage (with 95% CIs) for each question type is contained in Table 1. Approximately half of questions comprising interviews in all conditions were closed yes/no questions. In general, there was an improvement in the sorts of questions asked in post-experimental interviews, as trained interviewers in this condition asked fewer inappropriate questions (e.g., leading,

Table 1 Means and 95% confidence intervals of the percentage of question types asked as a function of condition

Question type	Mean	95% Confidence interval	
		Lower limit	Upper limit
Leading questions			
Pre-control	1.89	0.68	3.12
Post-control	1.63	0.41	2.85
Pre-experimental	3.24	2.02	4.46
Post-experimental	0.45	0.00	1.67
Multiple questions			
Pre-control	2.88	1.35	4.40
Post-control	0.85	0.00	4.40
Pre-experimental	1.89	0.37	3.41
Post-experimental	1.5	0.00	3.02
Closed yes/no questions			
Pre-control	44.12	38.17	50.06
Post-control	43.09	37.15	49.04
Pre-experimental	49.93	43.99	55.87
Post-experimental	50.94	44.92	56.88
Forced-choice questions			
Pre-control	1.91	0.46	3.36
Post-control	4.53	3.07	5.98
Pre-experimental	3.77	2.32	5.22
Post-experimental	3.52	2.06	4.96
Open-ended questions			
Pre-control	2.19	0.64	3.76
Post-control	2.53	0.98	4.09
Pre-experimental	2.61	1.05	4.17
Post-experimental	5.05	3.49	6.61
Probing questions			
Pre-control	46.99	41.34	52.66
Post-control	47.36	41.70	53.02
Pre-experimental	38.56	32.90	44.22
Post-experimental	38.55	32.89	44.21

forced-choice, multiple), and a larger proportion of more appropriate question types (e.g., probing). Most notably, there was a large ($d = 0.96$) reduction in the percentage of leading questions, and a moderate ($d = 0.67$) increase in the percentage of open-ended questions asked by trained interviewers. Comparisons were also made between the trained group of interviewers and the three untrained groups (collapsed together). Trained interviewers asked significantly fewer leading questions than the untrained interviewers, $t(78) = 2.53, p < .05, d = 0.78$, and significantly more open questions than the untrained interviewers, $t(78) = -2.91, p < .05, d = 0.66$. There were no other statistically significant differences between trained and untrained interviewers for the remaining types of questions.

On average, the proportion of witness talking time for those in the pre-experimental, post-experimental, pre-control, and post-control was: 64.05% ($SD = 15.93$), 47.14% ($SD = 12.33$), 66.04% ($SD = 10.97$), and 53.58% ($SD = 15.61$), respectively. When witness talking time was also analyzed after subtracting the interviewers' contribution in the engage and explain phase, the proportion of witness talking time for those in the post experimental group ($n = 20$) was 49.95% ($SD = 12.66$). After taking out the time spent explaining what is going to happen during the interview, witness talking time was significantly higher in interviews conducted by trained interviewers compared to those interviews which included the interviewer's engage and explain phase, $t(19) = -4.66, p < 0.00$. In addition (based on sub-sample sizes indicating an interview in audio or video format), the average number of interruptions made by the interviewer was 4.2 ($SD = 8.84$), 1.79 ($SD = 3.39$), 0.00, and 0.68 ($SD = 2.31$), for the pre-experimental ($n = 5$), post-experimental ($n = 19$), pre-control ($n = 1$), and post-control ($n = 18$) conditions, respectively. Due to small and unequal sample sizes, inferential statistics that test for differences between conditions were not conducted.

Closure Phase

The interviews in the post-experimental condition exhibited more of the six closure behaviors. Specifically, *providing a professional closure* and *summarizing the interview* were observed most frequently ($n = 10$, and $n = 7$, respectively) in those interviews. The mean percentage (with 95% CIs in parentheses) of closure behaviors for interviews conducted in the pre-experimental, post-experimental, pre-control, and post-control conditions were 22.50% ($CI = 15.01, 30.00$), 39.17% ($CI = 31.67, 46.66$), 30.83% ($CI = 23.34, 38.33$), and 18.33% ($CI = 10.84, 25.83$), respectively. The mean proportion of closure behaviors exhibited was higher for trained interviewers than those in the pre-experimental condition, $t(38) = 2.68, p = .01, d = .84$, as well as the post-control condition, $t(38) = 4.16, p < .001, d = 1.40$. Although the difference between the post-experimental and pre-control conditions was non-significant, $t(38) = 1.42, p > .05$, the difference produced a medium effect size, $d = .45$. The average effect size for comparisons made between the trained group and three untrained groups was $d = 0.90$ (95% $CI = -0.29, 2.09$).

Time Analysis

The relationship between overall score (calculated as the average percentage of *engage and explain*, *account*, and *closure* scores) and delay (i.e., time between date of training and date of interview) was negatively correlated ($r = -.44, p > .05$).

Discussion

We examined the effect of a witness interviewing training program on the interviewing performance of a sample of actual Canadian police interviews. We found that trained interviewers used more desirable practices than their untrained counterparts. Trained interviewers used more engage and explain, closure, and, to a lesser extent, account behaviors. Although trained investigators also asked fewer leading questions and asked more open-ended questions, the proportion of most question types (i.e., probing, closed yes/no, forced choice, multiple) did not vary as a function of training. Our findings suggest that it is possible to transfer some of the interviewing skills comprising a witness interview course from the classroom to the field. Much like the strong message that has been emanating from child interviewing research for some time (see Lamb et al. 2008), our findings suggest that modifications to the delivery of the training protocol and post-training support services is needed to maximize the transference of skills to actual criminal investigations.

Inherent within interview training are basic communication skills where the interviewer is taught to enlighten the interviewee through an explanation of the interview process and ends the interview on a positive note (Fisher and Geiselman 1992). It is therefore reassuring in that the majority of trained interviewers used nearly half of the behaviors deemed essential for creating a positive and relaxing environment prior to asking questions – compared to 13% to 16% for untrained interviewers. Although the percentage of engage-and-explain and closure behaviors is not exceedingly high, the improvements for the trained interviewers are encouraging. Similarly, nearly 40% of behaviors deemed important for ending an interview were employed – compared to 18% to 31% for untrained interviewers. The aforementioned findings are certainly in line with past findings that such skills, which are viewed by practitioners and academics to be relatively simple to implement, tend to be transferred to actual interviews (see Clarke et al. 2011; Griffiths and Milne 2006; Walsh and Milne 2008 for similar results). Such behaviors were relatively easy to implement by the trained interviewers because they were provided with a bullet point checklist of the desirable behaviors to follow (i.e., referred to as a PEACE mat), and are more conducive mental scripts.

We also found that trained interviewers used more desirable practices during the account phase than untrained interviewers. For example, trained interviewers passed to the second interviewer more often than the untrained interviewers. Even though the training did not lead to an increase in the requests for free narratives, it was encouraging that most interviews (at least 70%) across all conditions contained a request for a free narrative and were devoid of jargon. Interviews with trained interviewers contained more attempts to set up a CI (i.e., use one of

the three mnemonics) than those with untrained interviewers, which is consistent with previous descriptive field studies (see Kebbell and Milne 1998).

Granted that the trained interviewers attempted to use one of the mnemonics in 40% of the interviews (only reinstate mental context or encourage the witness to report everything was used), only a quarter of those interviews contained a proper set-up of that memory enhancing technique. On the one hand, it was somewhat surprising that more CI mnemonics were not observed in the interviews conducted by trained interviewers because the trainees were provided with 'instruction scripts' that they could employ if they decided that the mnemonics would be useful to help their witness improve their memory. Having said this, research has shown that some CI mnemonics are not implemented consistently by practitioners because they view them as being difficult to use (if they do not have the experience or skill set to implement them without aids), too complicated to communicate to witnesses, and/or take too long to implement (Brown et al. 2008; Clarke et al. 2011; Kebbell et al. 1999; Memon et al. 1994; Schreiber Compo et al. 2012). Compounded with the previously mentioned barriers to implementation, it is also the case that the interviewers were given the flexibility of using the mnemonics as needed. Thus, it might have been the case that the officers did not view the mnemonics as being needed during the interviews we sampled. Perhaps a greater emphasis during training that at least one of the mnemonics should be used during all interviews would have been reflected in our observations.

Trained investigators asked more open-ended questions and asked fewer leading questions than untrained interviewers. Despite these improvements, the percentage of open-ended questions asked by trained investigators is comparable to the percentages reported in other field studies of untrained interviewers where only approximately 6% of all questions asked began with *tell*, *explain*, or *describe* (see Myklebust and Alison 2000; Snook and Keating 2010; Wright and Alison 2004). Unfortunately, we also found that closed yes/no questions and probing questions constituted the vast majority of interviews by trained and untrained interviewers. Although probing questions have their place in an interview for filling in details, the proportion of these questions should not be as high as observed in our study.

One intuitively appealing step to make improve the implementation of all desirable behaviors is to increase the intensity of the training (e.g., more practice interviews during training). However, research from the child interviewing research suggests that increasing the intensity of the training may not be the most effective way forward. The child interviewing research suggests that a two-pronged approach may be required whereby interviewers are provided with scripts to follow and are provided with feedback/monitoring of their interviewers in the field (e.g., Lamb et al. 2002; Price and Roberts 2011; Powell 2008; Powell et al. 2005; also see Baddeley and

Longman 1978; Rischke et al. 2011). For example, Powell et al. (2008) found the most effective kind of feedback involved an expert identifying problems with the interviewer's behavior during simulated interviews to prevent further errors being made and incorporated into the interviewer's interviewing style. Of course we are aware that implementing an intensive training regimen that includes ongoing supervision and constructive feedback by experts is not a straightforward task (Powell et al. 2014). In fact, quality training and evaluation requires skill and commitment on a number of levels, such as highly skilled interviewers who are able to provide instruction to officers and commitment to interviewing excellence on an organizational level (e.g., human resources). Nevertheless, the evidence is clear that a commitment to a quality interviewing training initiative is required if organizations are intent on improving their interviewing and investigative abilities.

A global assessment of our results suggests that it is possible to train investigators to conduct adequate interviews with witnesses. Notwithstanding the increase in desirable behaviors beyond what is exhibited by untrained interviewers, there is still much room for improvement in interviewing practices. We are encouraged by the fact that it is possible to transfer skills from the classroom to the field, but our results reiterate a growing concern about transferring complex skills into the field. We certainly look forward to future studies that define, *a priori*, interviewing skills as complex or simple and then proceed to test the extent to which complex skills are the more difficult ones to transfer into practice. We also look forward to research that isolates the aspects of training programs that impacts the transference of skills directly. Additional research on the effect of individual differences in trainers and interviewers – along with research on the effect of the content and structure of training programs – on training effectiveness will ultimately allow for improvements in interviewing training programs so that skills get transferred fully into practice.

Acknowledgments This research was supported by the Social Sciences and Humanities Research Council of Canada. We would like to thank the participating police organization for providing the data required for the study. We would also like to thank the anonymous reviewers for providing constructive ways to improve the quality of our manuscript.

References

- Aldridge J, Cameron S (1999) Interviewing child witnesses: questioning techniques and the role of training. *Appl Dev Sci* 3:136–147. doi:10.1207/s1532480xads0302_7
- Baddeley AD, Longman DJA (1978) The influence of length and frequency of training session on the rate of learning to type. *Ergonomics* 21:627–635
- Brown C, Lloyd-Jones TJ, Robinson M (2008) Eliciting person descriptions from eyewitnesses: a survey of police perceptions of

- eyewitness performance and reported use of interview techniques. *Eur J Cogn Psychol* 20(3):529–560. doi:10.1080/09541440701728474
- Clarke C, Milne R, Bull R (2011) Interviewing suspects of crime: the impact of PEACE training, supervision, and the presence of a supervisor. *J Investig Psychol Offender Profiling* 8:149–162. doi:10.1002/jip.144
- Clifford BR, George R (1996) A field evaluation of training in three methods of witness/victim investigative interviewing. *Psychol Crime Law* 2(3):231–248. doi:10.1080/10683169608409780
- Dando C, Wilcock R, Milne R (2008) Victims and witnesses of crime: police officers' perceptions of interviewing practices. *Leg Criminol Psychol* 13:59–70. doi:10.1348/135532506X162498
- Davies G, Bull R, Milne R (2016) Analysing and improving the testimony of vulnerable witnesses interviewed under the 'achieving best evidence' protocol. In: Radcliffe P, Heaton-Armstrong A, Gudjonsson G, Wolchover D (eds) *Witness testimony in sexual cases. investigation, law, and procedure*. Oxford University Press, Oxford
- Fisher RP, Geiselman RE (1992) *Memory-enhancing techniques for investigative interviewing: the cognitive interview*. Charles C Thomas, Springfield, IL England
- Geiselman RE, Fisher RP, MacKinnon D, Holland H (1985) Eyewitness memory enhancement in the police interview: cognitive retrieval mnemonics versus hypnosis. *J Appl Psychol* 70(2):401–412. doi:10.1037/0021-9010.70.2.401
- Griffiths A, Milne B (2006) Will it all end in tiers? In: Williamson T (ed) *Investigative interviewing* (167–189). Willan, Cullompton
- Griffiths A, Milne R, Cheryman J (2011) A question of control? The formulation of suspect and witness interview strategies by advanced interviewers. *Int J Police Sci Manag* 13(3):255–267. doi:10.1350/ijps.2011.13.3.21910.1350/ijps.2011.13.3.219
- Kebbell M, Milne R (1998) Police officers' perception of eyewitness factors in forensic investigations. *J Soc Psychol* 138:323–330. doi:10.1080/00224549809600384
- Kebbell MR, Milne R, Wagstaff GF (1999) The cognitive interview: a survey of its forensic effectiveness. *Psychol Crime Law* 5(1-2):101–115. doi:10.1080/10683169908414996
- Köhnken G, Milne R, Memon A, Bull R (1999) The cognitive interview: a meta-analysis. *Psychol Crime Law* 5(1-2):3–27. doi:10.1080/10683169908414991
- Lamb ME, Sternberg KJ, Orbach Y, Esplin PW, Mitchell S (2002) Is ongoing feedback necessary to maintain the quality of investigative interviews with allegedly abused children? *Appl Dev Sci* 6(1):35–41. doi:10.1207/S1532480XADS0601_04
- Lamb ME, Hershkowitz I, Orbach Y, Esplin PW (2008) *Tell me what happened: structured investigative interviews of child victims and witnesses*. Wiley, Chichester
- Memon A, Holley A, Milne R, Kohnken G, Bull R (1994) Towards understanding the effects of interviewer training in evaluating the cognitive interview. *Appl Cogn Psychol* 8:641–659. doi:10.1002/acp.2350080704
- Memon A, Meissner CA, Fraser J (2010) The cognitive interview: a meta-analytic review and study space analysis of the past 25 years. *Psychol Public Policy Law* 16:340–372. doi:10.1037/a0020518
- Milne R, Bull R (2003) Does the cognitive interview help children to resist the effects of suggestive questioning? *Leg Criminol Psychol* 8(1):21–38. doi:10.1348/135532503762871219
- Myklebust T, Alison L (2000) The current state of police interviews with children in Norway: how discrepant are they from models based on current issues in memory and communication? *Psychol Crime Law* 6(4):331–351. doi:10.1080/10683160008409810
- NSLEC (2004) *Practical guide to investigative interviewing*. National Centre for Policing Excellence, Wyboston
- Powell MB (2008) Designing effective training programs for investigative interviewers of children. *Curr Issues Crim Justice* 20:189–208
- Powell MB, Fisher R, Wright R (2005) *Investigative interviewing*. In: *Psychology and law: an empirical perspective*. New York, Guilford Press, pp 11–42
- Powell MB, Fisher R, Hughes-Scholes C (2008) The effect of intra-versus post-interview feedback during simulated practice interviews about child abuse. *Child Abuse Negl* 32:213–227. doi:10.1016/j.chiabu.2007.08.002
- Powell MB, Hughes-Scholes CH, Smith R, Sharman SJ (2014) The relationship between investigative interviewing experience and open-ended question usage. *Police Pract Res* 15:283–292. doi:10.1080/15614263.2012.704170
- Price HL, Roberts KP (2011) The effects of an intensive training and feedback program on police and social workers' investigative interviews of children. *Can J Behav Sci* 43:235–244. doi:10.1037/a0022541
- Rischke AE, Roberts KP, Price HL (2011) Using spaced learning principles to translate knowledge into behavior: evidence from investigative interviews of alleged child abuse victims. *J Police Crim Psychol* 26:58–67. doi:10.1007/s11896-010-9073-8
- Sanders G (1986) On increasing the usefulness of eyewitness research. *Law Hum Behav* 10(4):333–335. doi:10.1007/BF01047345
- Schreiber Compo N, Hyman Gregory A, Fisher R (2012) Interviewing behaviors in police investigators: a field study of a current US sample. *Psychol Crime Law* 18(4):359–375. doi:10.1080/1068316X.2010.494604
- Shepherd E (2007) *Investigative interviewing: the conversation management approach*. Oxford University Press, Oxford
- Snook B, Keating K (2010) A field study of adult witness interviewing practices in a Canadian police organization. *Leg Criminol Psychol*. doi:10.1348/135532510X497258, **Advance online publication**
- Snook B, Luther K, Quinlan H, Milne R (2012) Let 'em talk! A field study of police questioning practices of suspects and accused persons. *Crim Justice Behav* 39:1328–1339. doi:10.1177/0093854812449216
- Walsh D, Milne R (2008) Keeping the PEACE? A study of investigative interviewing practices in the public sector. *Leg Criminol Psychol* 13(1):39–57. doi:10.1348/135532506X157179
- Warren AR, Woodall CE, Thomas M, Nunno M, Keeney JM, Larson SM, Stadfeld JA (1999) Assessing the effectiveness of a training program for interviewing children. *Appl Dev Sci* 3:128–135. doi:10.1207/s1532480xads0302_6
- Wright A, Alison L (2004) Questioning sequences in Canadian police interviews: constructing and confirming the course of events? *Psychol Crime Law* 10(2):137–154. doi:10.1080/1068316031000099120