Evaluating the Vermont State police's PEACE model training program: phase 1

7 authors, including:

Laura Fallon
Memorial University of Newfoundland
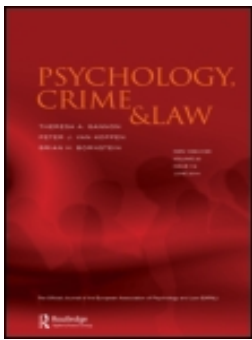**9** PUBLICATIONS **24** CITATIONS

SEE PROFILE

Brent Snook
Memorial University of Newfoundland
85 PUBLICATIONS **1,818** CITATIONS

SEE PROFILE

Evaluating the Vermont State police's PEACE model training program: phase 1

Laura Fallon , Brent Snook , Todd Barron , Angela Baker , Mike Notte , Jeff Stephenson & Dan Trottier

Published online: 05 Feb 2021.

Submit your article to this journal

View related articles

View Crossmark data

Routledge
Taylor & Francis Group

Check for updates

Evaluating the Vermont State police's PEACE model training program: phase 1

Laura Fallon[a], Brent Snook[a], Todd Barron[b], Angela Baker[c], Mike Notte[c], Jeff Stephenson[c] and Dan Trottier[c]

[a]Department of Psychology, Memorial University of Newfoundland, St. John's, Canada; [b]Department of Sociology, Memorial University of Newfoundland, St. John's, Canada; [c]Vermont State Police, Waterbury, VT, USA

ABSTRACT

The current study is the first phase of a larger evaluation of the Vermont State Police (VSP) PEACE Model interview training program. This evaluation measures the impact of the program on trainees' interviewing knowledge and attitudes, and opinions about the program. VSP Members participating in Tier 1 ($N = 72$) and Tier 2 ($N = 46$) PEACE interview training completed a questionnaire, both before and after training, about (1) their knowledge of interviewing and related concepts, (2) their attitudes about various interviewing practices, and (3) their beliefs about the training session. Results showed that training had a large effect on knowledge for participants in both training courses. As well, a majority of Tier 1 and Tier 2 trainees reported attitudes in line with scientific consensus after training. Specifically, most items related to eyewitness testimony, memory, and coercion were endorsed post-training by the majority of officers, while items covering deception detection and interrogation rights were endorsed less frequently. Additionally, trainees reported that their training experience was quite positive. Implications for the program and future directions for training are discussed.

The success of any criminal investigation and subsequent prosecution is tied directly to the quality of victim, witness, and suspect interviews; that is, interviews generate the information and evidence to propel legal processes (Kebbell & Milne, 1998; Milne & Bull, 2003; Sanders, 1986). Investigative interviewing researchers have measured interview quality, broadly, by the types and amount of ethical, science-based practices present in an interview, along with the types and amount of maladaptive tactics and the risk that those practices pose to the reliability of the information obtained (e.g. asking inappropriate and leading questions, suggestive interviewing, coercion; e.g. Alison et al., 2013; Davis & Loftus, 2017; Kassin et al., 2010). Moreover, the ability to minimize maladaptive and maximize constructive tactics depends partially on interviewers having conceptual knowledge of interview practices and also being instructed on the theories and research that explain

the effect such practices have on the quality of information elicited. Put simply, police officers need science-based training if they wish to be professional investigative interviewers.

The PEACE framework

One science-based interview training program promoted widely for interviewing adult witnesses, victims, and suspects is the PEACE framework. The PEACE framework draws heavily on scientific findings in the realm of police interviewing and related areas; specifically, the well-researched cognitive interview, conversation management, various psychological theories and taxonomies (e.g. how memory works, false confessions), and a plethora of empirical findings about investigative interviewing practices (see Alison et al., 2013; Milne & Bull, 2003). The name PEACE is simply an acronym that stands for the five stages of interviewing that comprise the model: Planning and Preparation, Engage and Explain, Account, Closure, and Evaluation. Each of the five stages contain specific behaviors that are thought to be required for a successful interview. In brief, the preparation and planning stage teaches interviewers to conduct a thorough review of all case materials, become familiar with the personal characteristics of the interviewee (e.g. vulnerabilities), and prepare questions to ask during the interview. During engage and explain, officers are instructed to build rapport with the interviewee by forming a good relationship, treating them with respect, speaking genuinely, and trying to identify something in common. The account phase involves eliciting information about an event in question from the interviewee through the use of open-ended (e.g. tell, explain, describe) and probing (e.g. 5W) questions, and allowing the interviewee to do the vast majority of the talking. The closure phase involves clarification and finalization of the interviewee's account and ensuring that all parties are on the same page about the next steps, while the evaluation phase allows officers to constantly assess incoming information and evaluate their interview skills (i.e. feedback).

The PEACE framework has a clear philosophical stance whereby officers are taught to be genuine, fair, respectful, and objective in their interviewing approach, as well as in their broader approach to conducting investigations (see Clarke et al., 2011). Officers trained in this model are instructed to collect evidence prior to making decisions, akin to the process of hypothesis testing in scientific disciplines (and in direct contrast to traditional accusatorial interview methods). PEACE-trained interviewers are also taught to be open-minded, to use appropriate questions for gathering information (e.g. open-ended), to avoid using coercion, manipulation, or lies, and to abstain from trying to detect deception through behavioral cues (Clarke et al., 2011).

The PEACE framework, as taught in North America, is also tiered (Tiers 1–5) so that police organizations can engage in investigative capacity building. Not only is PEACE based on scientific findings, but it also provides a structured process that helps professionals develop confidence in their investigative abilities, conduct interviews in a more consistent manner, and attain investigative consistency across an organization. For example, Tier 1 introduces investigators to the fundamental concepts associated with interviewing (e.g. effective communication skills, extracting information through free narratives, asking effective questions). When a good interviewing foundation is established, other more advanced skills can be added to the interviewing toolbox. In Tier 2,

investigators are given more in-depth instruction about witness and suspect interviewing, including introductions to cognitive interviewing and conversation management, which are tools that help to structure the witness and suspect interview processes, respectively. Tier 3 is aimed toward those who have excelled at Tier 1 and 2 skills, and pertains to advanced training on cognitive interviewing, child interviewing via the NICHD protocol, and conversation management. Tier 4 involves training advanced investigators to evaluate interviews conducted by PEACE-trained officers in their organization. Finally, Tier 5 pertains to the strategic advisement of the entire framework (see Griffiths & Milne, 2006 for more detailed information about each tier of the PEACE framework).

Does PEACE work?

There is a wide body of evidence that suggests that the practices and protocols that comprise the PEACE framework are effective. For example, meta-analyses assessing the effectiveness of the cognitive interview against a control interview produced large effect sizes for the number of correct details recalled (ds > 0.87), and small effect sizes for the number of incorrect details recalled (ds < 0.28; Köhnken et al., 1999; Memon et al., 2010). Other studies have found that training in the cognitive interview resulted in interviewers asking more appropriate questions and fewer inappropriate questions (Fisher et al., 1989; George & Clifford, 1992). Research looking at the effect of question type on information provision has shown that open-ended questions produce the most information from witnesses (Fisher et al., 1989; Snook et al., 2012) and also lead to more accurate responses (Lipton, 1977). Similarly, allowing an interviewee to provide an uninterrupted free narrative has been shown to produce a large amount of information (e.g. Lipton, 1977; Marquis et al., 1972). Studies have also shown that interviewers who build rapport with an interviewee are able to elicit more details than those who adopt a more neutral or abrupt tone during questioning (Collins et al., 2002; Roberts et al., 2004; Walsh & Bull, 2012; see Bull & Baker, 2019 for a review of the benefits of using rapport in interviews) and that a relaxed interview environment leads to increased memory performance (e.g. de Quervain et al., 2000). Walsh and Bull (2010), in their analysis of benefit fraud interviews, found that being skilled in many of the components of PEACE (e.g. rapport, communication, flexibility) contributed to an increased likelihood of good interview outcomes like comprehensive accounts and confessions.

There is also evidence demonstrating that the overall approach taken in PEACE interviews is associated with positive outcomes. For example, the removal of coercive practices decreases the likelihood of negative outcomes, including inadmissible confessions, civil liability for negligent investigations, resentment from those being interviewed by police, declining public confidence, and a disregard of legal rights (Gudjonsson, 2003; Marin, 2004). This decrease in negative outcomes is also thought to come without a decrease in the rate of confessions (Milne & Bull, 2003). A meta-analysis that compared accusatorial and information-gathering interview styles used against suspects found that the information-gathering approaches were much more diagnostic; that is, they produced more true confessions and fewer false confessions than the accusatorial interviews (Meissner et al., 2014; see Evans et al., 2013 for similar findings in the intelligence context). According to individuals convicted of crimes, approaches involving humane treatment and rapport building are much more conducive to eliciting information from suspects

than aggressive and confrontational approaches (e.g. Holmberg & Christianson, 2002; Kebbell et al., 2006; Snook et al., 2015). A convergence of evidence from outside the police interviewing literature (e.g. intelligence, counseling, crisis negotiation, learning psychology) further supports the use of ethical and non-confrontational interview methods, suggesting that methods aimed at developing trust, respect, and cooperation – as opposed to 'get tough', accusatorial methods – will be more productive and lead to more positive outcomes (e.g. more complete and accurate information provided; Cook & Roesch, 2012; Gershoff, 2002; Giebels & Taylor, 2009; Goodman-Delahunty & Martschuk, 2020; Miller & Rollnick, 2012). Granted this data, few studies exist with evidence directly supporting the effectiveness of PEACE interviewing at achieving positive interview outcomes in practice (e.g. MacDonald et al., 2017).

Evaluating PEACE training

Based on the evidence presented above, it is clear that certain aspects of the PEACE framework are effective at producing positive interview outcomes (i.e. increased quantity and quality of information provided, increased confession rate). What is yet to be comprehensively evaluated, however, is whether it is possible to effectively train interviewers using the PEACE framework. Practically every study that has previously assessed the effectiveness of PEACE training has looked at its impact on the quality of interviews. The majority of these studies, examining both witness and suspect interviews, have indicated that making major changes in officers' interview skills is difficult to accomplish, reporting only modest improvements to interview behaviors post-training (Clarke et al., 2011; Clifford & George, 1996; Dando et al., 2009; Griffiths & Milne, 2006; Kebbell & Milne, 1998; Walsh & Milne, 2008). Recently, however, MacDonald et al. (2017) found that PEACE trained interviewers used more engage and explain ($d = 1.65$), account ($d = 0.54$), and closure behaviors ($d = 0.90$) in witness interviews than untrained interviewers, and asked more appropriate questions.

The current evaluation

All of the studies mentioned above looked solely at one facet of change caused by PEACE training: a change in behavior during interviews. To determine the effectiveness of a training program, however, it is important to look at the success of the program from multiple angles. One way to do this is to conduct a comprehensive program evaluation. Program evaluation differs from traditional research primarily in its purpose; while research is conducted to add knowledge to a particular field, the primary purpose of an evaluation is to make a judgement about the value, worth, and/or merit of the object of evaluation (Fitzpatrick et al., 2011). In program evaluation, focus is not only placed on the ultimate goal of a program, but also on whether the program has achieved more immediate and short-term outcomes.

According to the classic framework outlined by Kirkpatrick (1959), when evaluating a *training* program, there are four levels on which the program should evaluated: reaction, learning, behavior, and results (but see Alliger et al., 1997 for concerns about the framework and suggestions for improvement). Evaluation of *reaction* involves measuring the affective reactions and utility judgements of program participants, and evaluation of

learning involves assessing immediate knowledge, knowledge retention, and demonstration of learned skills (Kirkpatrick & Kirkpatrick, 2006). Both of these levels correspond to the short-term outcomes of a program, which focus on changes in knowledge, skills, opinions, and attitudes (Fitzpatrick et al., 2011; Frechtling, 2007). To evaluate *behavior*, which is often considered an intermediate outcome of a program, one must measure program participants' use of their learned skills in the real world, or in other words, their transfer of their learned skills from training to practice. Finally, the measurement of *results* involves determining the overall impact a program has made to the organization in which it resides (Kirkpatrick & Kirkpatrick, 2006). The result criteria could be said to correspond to the long-term, or ultimate goal of a program.

In 2018, members of the Vermont State Police (VSP) were trained to be administrators of the PEACE framework, with the intention of implementing the model in their organization. As a result of the singular focus on behavior change in the interviewing literature thus far, a comprehensive program evaluation was chosen to assess the effectiveness of PEACE training from several angles. In consultation with members of the VSP, and using the Kirkpatrick four level framework, a logic model (i.e. a visual representation of a program's resources, activities, and expected outcomes that helps to illustrate the relationship between various components of a program; Frechtling, 2007) was created to help outline our primary objectives and guide the evaluation process (see Figure 1). The logic model highlights the resources that have been used for the program (i.e. inputs), the things that the administrators did to prepare, develop, and implement the program (i.e. activities), the direct products of the program (i.e. outputs), and the expected outcomes, or in other words, the goals the program was designed to achieve.

We proposed an evaluation that would provide evidence to help determine (1) whether or not PEACE training is being implemented as planned, and (2) whether or not PEACE training is beneficial, worth the resources, and should be continued. To outline our approach to the evaluation, we created an evaluation framework, which summarizes the main program objectives, the questions we need to answer to see if the
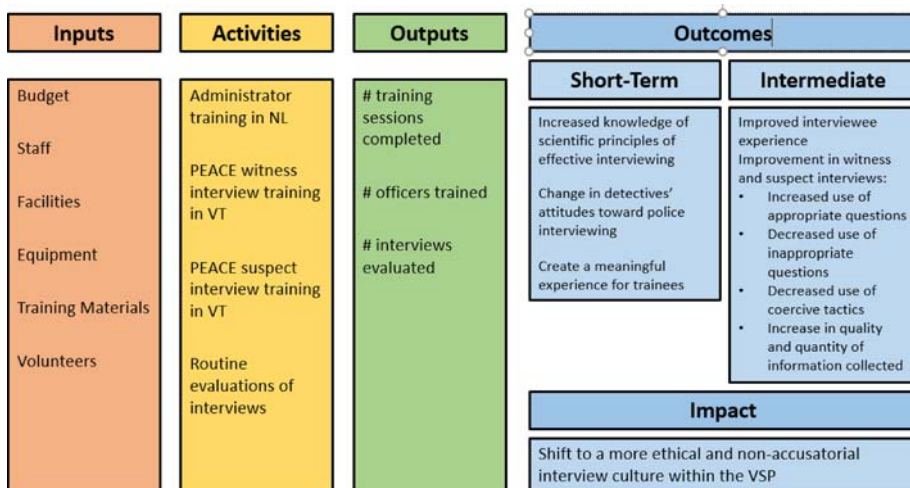
Figure 1. Logic model for the VSP PEACE training program.

objectives are being met (i.e. evaluation questions), the pieces of information we will need to answer these questions (i.e. indicators), and how/from whom we will get that information (i.e. methods and data sources; see Table 1).

The current study is the first phase of the evaluation of the VSP PEACE training program (see highlighted portion in Table 1). Before the lasting impact of PEACE training on

Table 1. Evaluation framework.

Program Objectives	Evaluation Questions	Indicators	Methods	Data Sources
1. Implement program effectively and as intended	Are program activities being implemented as planned?	• Output numbers • Views of program administrators	• Document review • Key informant interviews	• Administrative data • Program administrators (trainers)
	How is the program being perceived by trainees?	Views of trainees	Survey	Trainee detectives
2. Have a positive impact on detectives' knowledge, attitudes, and beliefs about police interviewing and related concepts (e.g. coercion, vulnerabilities, false confessions)	Does PEACE training succeed at teaching what it is meant to teach?	Survey results (knowledge questions)	Survey (pre-post)	Trainee detectives
	Have detectives' attitudes and beliefs toward police interviewing changed after completing PEACE training?	Survey results (attitude and belief questions)	Survey (pre-post)	Trainee detectives
3. Improve the experience of VSP interviewees	Do witnesses interviewed by a PEACE-trained VSP detective report more positive experiences than those interviewed by untrained detectives?	Views of interviewees	Survey	Witnesses interviewed by trained and untrained officers
4. Improve the quality of VSP witness and suspect interviews	Are the interviews conducted by trained VSP detectives better than those conducted by untrained detectives?	• Proportion of appropriate/ inappropriate questions • Amount of coercion • Quantity of information provided • Quality of details provided	Document review	Police interview transcripts (from trained and untrained officers)
5. Change in the overall culture of interviewing and investigation within the Vermont State Police	Has there been a shift in the interviewing culture within the VSP since the onset of the PEACE training program?	• Views of legal practitioners and policy makers • Views of VSP detectives • Views of the general public	Survey (1 year post-training)	• Prosecutors and judges • VSP detectives • Community members

interviews conducted by the VSP is assessed, it is important to ascertain whether PEACE training is successfully making the short-term changes it was intended to make (i.e. short term outcomes; see Figure 1). The first two objectives of this training program were to (1) increase trainees' knowledge about scientific principles related to interviewing, and (2) shift trainees' attitudes and beliefs about interview practices and the use of coercion to better align with scientific consensus on best practices for interviewing. We hypothesized that (1) knowledge about best-practice interviewing and associated concepts would increase from pre- to post-training, and (2) trainee attitudes about various aspects of interviewing would move closer to scientific consensus from pre- to post-training. We also wanted to measure the extent to which the training program was a positive and meaningful experience for the participating officers. Due to the exploratory nature of this objective, we refrained from making any predictions about officer experiences.

Method

Participant selection

All data collection activities for this project were conducted as part of a larger evaluation of the VSP PEACE training program. The researchers were not involved in the selection or recruitment processes; data was collected from individuals who were already scheduled to participate in the VSP's interview training sessions. Training administrators selected trainees for each individual session on the basis of logistics and convenience. When training is complete, all officers in the VSP will have participated in PEACE interview training.

PEACE training

Three senior VSP officers were trained in the administration of the PEACE framework by the second and third author of this paper, who are a psychology professor and former police officer/investigative interviewing advisor, respectively. Both individuals were trained previously in the PEACE model by senior PEACE trainers in the UK, have published extensively on investigative interviewing, and have been offering PEACE training to legal organizations across North America for the past decade.

Tier 1 training

Tier 1 training took place over three consecutive days and was taught by a senior officer who was trained in the administration of the PEACE model. Tier 1 training was primarily lecture-based; it introduced investigators to the fundamental concepts associated with interviewing and the benefits of interviewing protocols that are grounded in scientific findings. The training aimed to help investigators understand the importance of using effective communication skills (e.g. respect), being ethical, building rapport, relaxing interviewees to facilitate concentration on memory retrieval, avoiding distractions, extracting information through free narratives, asking effective questions (i.e. open-ended and probing questions), and effectively dissecting information (i.e. using People, Locations, Actions, and Times). Investigators were also taught about human memory (e.g. how people recall information, barriers to retrieving information, memory systems and how interviewers can unwittingly contaminate memory through bad question

practices), memory enhancement techniques (e.g. open invitations, report all/do not edit instructions), and how to avoid making common interviewing errors such as interrupting people while they are talking, over-talking, and using inappropriate types of questions (e.g. closed yes/no, leading).

Trainees. Individuals in the Tier 1 PEACE training session (*N* = 72) were trainee officers and road troopers from the VSP. Additional demographic information (e.g. gender, age) for each participant was not collected due to confidentiality concerns.

Tier 2 training

The Tier 2 training session was delivered over two non-consecutive weeks – usually about two months apart – by three senior members of the VSP trained in the administration of PEACE training; participants first took part in a week-long training session about witness interviewing, and then returned several weeks later for a week-long session on suspect interviewing.

   The first week of Tier 2 training focused on teaching participants the main aspects of PEACE and how to conduct a witness interview, while the second week focused on suspect interviewing. Each phase of training took place on a full-time basis (approximately seven hours per day) over the period of five consecutive weekdays. The training was co-administered by three VSP detectives, all of whom were trained previously in the PEACE model of interviewing by the second and third author. All participants 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 & Geiselman, 1992; Milne & Bull, 2003; NSLEC, 2004; Shepherd, 2007). The lectures on witness interviewing 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. The lectures on suspect interviewing discuss the model of conversation management, false confessions, the risk factors associated with false confessions, problems associated with deception detection, and research on the best way to administer Miranda rights.

   Interspersed with the lectures were practice interviews (ranging from 30 min to 1 h) with actual witnesses and mock suspects, where interviewers were provided with a checklist of behaviors that they could reference as needed and immediate verbal feedback was provided from the trainers. The practice sessions followed a scaffolding approach that developed interviewing skills though 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 during their practice interviews, and use a checklist of desirable behaviors during those interviews. At the end of training, each interviewer conducted a final mock interview (with a naïve, independent volunteer) encompassing all stages of the PEACE model. The non-interviewing investigators were required to watch all practice interviews and provide peer feedback using the checklist.

Trainees. All officers participating in Tier 2 of PEACE were detectives from the VSP ($N =$ 46). Additional demographic information was not collected due to confidentiality concerns.

Design and materials

This evaluation employed a single factor within-subjects design (Pre-Training vs. Post-Training). On the first day of PEACE training, participants completed a questionnaire containing questions assessing their knowledge and attitudes about police interviewing and associated topics. On the final day of training, participants completed the same questionnaire, but with an added section eliciting feedback on the training session itself. Due to logistics and scheduling, the time between questionnaires varied by tier. On average, there was about a two-month gap between completion of the pre- and post-training questionnaires for Tier 2 participants. Tier 1 participants completed the pre-training questionnaire of the first day of training and completed the post-training questionnaire at the end of the third and final day of training.

Both questionnaires contained 30 questions assessing respondents' knowledge of police interviewing, 15 of which were in a multiple-choice format (e.g. 'what is the main objective of an interview with a suspect?') and 15 true/false (e.g. 'suggesting facts to interviewees is a useful interviewing practice'). The questions reflect key learning outcomes of the program. Both questionnaires also contained 16 questions for which respondents reported their attitudes towards various aspects of interviewing using a five-point scale (*1 = strongly disagree, 5 = strongly agree*; e.g. 'eyewitness testimony can be affected by how the questions asked to the witness are worded'). Some of the attitude items were adapted from several existing studies that asked participants to rate their attitudes about memory processes, witness testimony, confessions, and Miranda rights, among others (Kassin et al., 2001, 2018; Magnussen & Melinder, 2012; Rogers et al., 2010; Simons & Chabris, 2011), while others were adapted from the primary training materials (i.e. Fisher & Geiselman, 1992; Milne & Bull, 2003; NSLEC, 2004; Shepherd, 2007). Beyond the knowledge check questions and attitude items, the post-training questionnaire also contained eight items asking respondents to rate aspects of their experience during the training program on a five-point scale (*1 = strongly disagree, 5 = strongly agree*; e.g. 'the trainers were knowledgeable about the topics being taught'), one item asking respondents to rate their overall training experience on a five-point scale (*1 = poor, 5 = excellent*), and four open-ended items eliciting additional feedback about training (e.g. 'what did you like the most about the training program?').

Procedure

The trainers administered both questionnaires; participants were told to find a quiet space outside the training room to complete the questionnaires for confidentiality purposes. The questionnaire stated that participants were not required to respond to any of the questions and could skip questions that they did not wish to answer. Participants were instructed to return their completed surveys, sealed in an envelope, to a drop box located outside the training room. Badge numbers were used to match individual

participants' pre- and post-training responses, but the envelopes containing the questionnaires were not opened until they were received by the researchers, who were unaware of any officer's badge number and were unable to use this information to link responses to participants. Once the data were transferred to the computer, badge numbers were replaced with a unique identification code. Ethics clearance was granted by the Interdisciplinary Committee on Ethics in Human Research at Memorial University.

Table 2. Percent of Trainees who Answered Correctly on Pre- and Post-Training Questionnaire by Tier

	Multiple Choice Questions	Tier 1		Tier 2	
		Pre	Post	Pre	Post
1.	The approach used in PEACE interviewing is described as …	90.3%	98.6%	100.0%	97.4%
2.	What is the main objective of an interview with a suspect?	83.3%	92.9%	88.6%	94.7%
3.	Which of the following is an example of an open-ended question?	27.8%	67.1%	56.8%	73.7%
4.	The relationship between eyewitness accuracy and confidence is …	27.8	75.7%	54.5%	63.2%
5.	Which of the following is the most effective way to extract information from a witness?	23.6%	35.7%	52.3%	81.6%
6.	Which of the following is NOT a memory enhancement tool used in PEACE interviews with witnesses?	56.9%	84.3%	84.1%	92.1%
7.	Memories for one-time events that were previously witnessed by a person are known as memories	48.6%	88.6%	79.5%	92.1%
8.	Which of the following is NOT an effective interviewing technique?	83.3%	92.9%	97.7%	94.7%
9.	About % of wrongful convictions involved false confessions.	27.8%	22.9%	15.9%	42.1%
10.	The best way to deliver the Miranda warning to a suspect is …	56.9%	91.4%	68.2%	94.7%
11.	Which of the following would NOT be considered a vulnerability factor?	84.7%	85.7%	93.2%	94.7%
12.	A detail given by a witness that has not been checked or verified is classified as …	90.3%	97.1%	90.9%	92.1%
13.	A combination of points mentioned by the witness that have been checked and verified is/are known as …	16.7%	18.6%	6.8%	7.9%
14.	A 'route-map' is …	54.2%	51.4%	65.9%	55.3%
15.	Out of 100% of the time spent talking between an interviewee and interviewer, the interviewee should talk about % of the time.	72.2%	92.9%	86.4%	97.4%
	True/False Questions				
16.	No suspect will ever fully confess to a crime, thus, they all require persuasion.	90.0%	97.0%	91.0%	100.0%
17.	It is best to write down the questions you want to ask before starting an interview.	94.0%	96.0%	73.0%	95.0%
18.	Taking notes in an interview is part of professional interviewing.	88.0%	100.0%	86.0%	100.0%
19.	Summarizing everything the interviewee stated in an interview is a waste of time.	97.0%	100.0%	98.0%	100.0%
20.	A minimal introduction (1 min or so) is enough before starting to ask questions.	53.0%	33.0%	82.0%	79.0%
21.	Body language provides a strong indication of who is lying.	11.0%	57.0%	59.0%	100.0%
22.	Suggesting facts to interviewees is a useful interviewing practice.	47.0%	90.0%	86.0%	84.0%
23.	It is important to write verbatim notes of everything the interviewee says.	69.0%	90.0%	96.0%	84.0%
24.	It is best to directly confront a suspect with false evidence of guilt at the start of an interview.	100.0%	99.0%	98.0%	100.0%
25.	Ideally, the Miranda warning should be given at the start of a suspect interview.	90.0%	67.0%	80.0%	84.0%
26.	It is okay to interrupt interviewee responses to clarify issues.	72.0%	90.0%	82.0%	95.0%
27.	Rapidly switching from one topic to another while obtaining information from an interviewee is a good way to ensure complete and accurate information.	94.0%	99.0%	100.0%	100.0%
28.	It is important to come into an interview without any judgements about the suspect/witness.	100.0%	99.0%	100.0%	97.0%
29.	A thorough review of a file is always required before conducting an interview.	78.0%	71.0%	84.0%	61.0%
30.	It is best to introduce yourself to a witness using your first name.	83.0%	97.0%	96.0%	97.0%

Results

Knowledge check questions

Responses to each of the 30 questions were coded as either correct or incorrect (see Table 2 for the percentage of respondents who answered each question correctly pre- and post-training). Correct responses were summed and converted to the percentage of questions answered correctly.

Tier 1

The average pre-training score for Tier 1 trainees was 67% ($SD = 9.42$) and the average post-training score was 79% ($SD = 7.25$). A paired-samples t-test revealed a significant difference between these scores, $t_{(63)} = -10.20$, $p < .001$, $d = 1.28$, indicating that respondents had more knowledge about interviewing after training than they did before training.

Tier 2

The average score on the pre-training questionnaire for Tier 2 trainees was 77% ($SD = 9.02$), and the average post-training score was 85% ($SD = 7.10$). A paired-samples t-test revealed a significant difference between the two scores, $t_{(35)} = 4.68$, $p < .001$, $d = 0.78$, indicating that participants knowledge about interviewing increased from pre- to post-training.

Attitudes toward interviewing: descriptive statistics

Since some attitude items were worded positively and some were worded negatively (i.e. some reflected attitudes that would be beneficial for officers to have, while others reflected attitudes they ought not to have), the negative items were reverse coded so that agreement (i.e. 4 or 5 on the scale) with all items would reflect an attitude that is in line with scientific consensus.

Tier 1

The percentage of Tier 1 trainees who endorsed attitudes in agreement with scientific consensus before and after training for each item are shown in Figure 2. As can be seen, there were eight items for which a vast majority of respondents (i.e. > 75%) agreed with scientific consensus after undergoing training (i.e. Q1, Q3, Q4, Q5, Q8, Q9, Q11, Q14); note however that, for four of these eight items (i.e. Q3, Q4, Q11, Q14), a vast majority of respondents already reported perceptions in line with science before training. There were another four items on which a majority (i.e. > 50%) of respondents agreed with science after their interview training (i.e. Q7, Q10, Q12, Q15); there was a large increase in respondents agreeing with science after training for one question (Q12), a small increase for two others (Q10, Q15), and no increase for the final item (Q7). For the remaining four items (i.e. Q2, Q6, Q13, Q16), less than half of respondents reported an attitude in line with scientific consensus after training. Notwithstanding the low level of agreement with science after training for two of these four items (Q2, Q16), there was still a large increase in the percentage of respondents who held
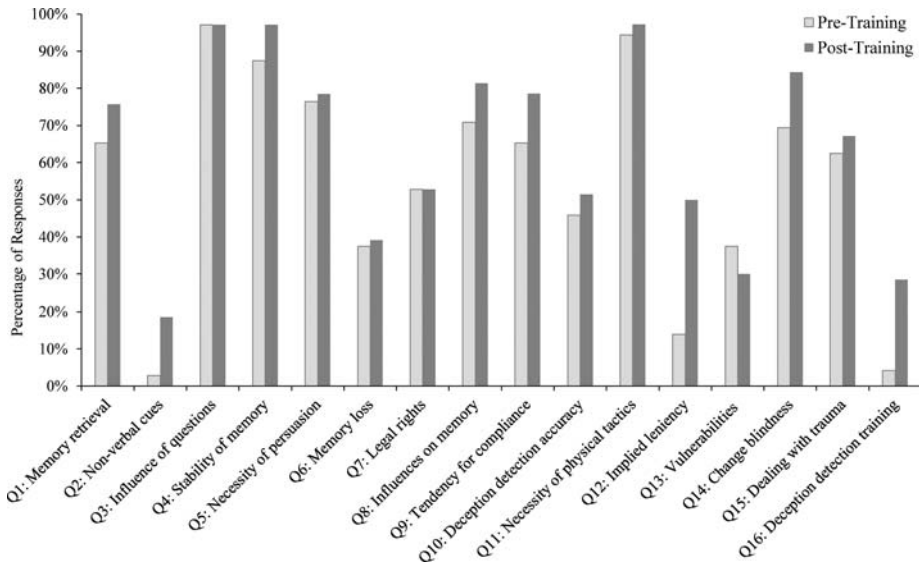
Figure 2. Percentage of respondents reporting attitudes agreeing with scientific consensus on pre- and post-training questionnaire items (Tier 1)

science-based attitudes from pre- to post training. In sum, a majority of respondents reported attitudes in line with science after training for 12 of the 16 items (75%).

Tier 2

The percentage of Tier 2 trainees who endorsed attitudes in agreement with science before and after training for each item is shown in Figure 3. As can be seen, there were
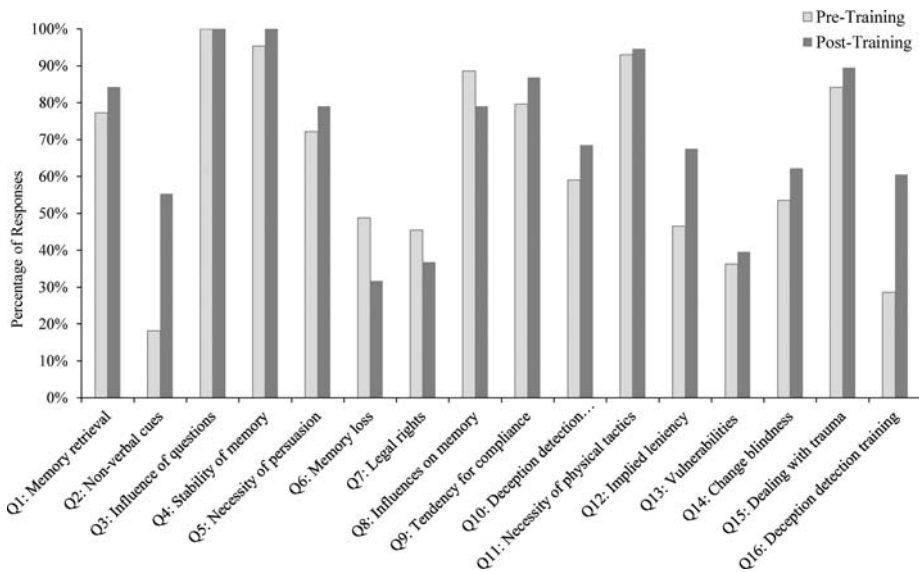

Figure 3 . Percentage of respondents reporting attitudes agreeing with scientific consensus on pre- and post-training questionnaire items (Tier 2)

eight items for which a vast majority of respondents (i.e. > 75%) agreed with scientific consensus after undergoing training (i.e. Q1, Q3, Q4, Q5, Q8, Q9, Q11, Q15). For all but one of these items (Q5), a vast majority of respondents already reported perceptions in line with science before training. There were five other items on which a majority (i.e. > 50%) of respondents agreed with science after their interview training (i.e. Q2, Q10, Q12, Q14, Q16); there was a large increase in respondents agreeing with science after training for three items (Q2, Q12, Q16), and a medium increase for the other two items (Q10, Q14). Lastly, there were three items for which less than half of respondents reported an attitude in line with scientific consensus after training (i.e. Q6, Q7, Q13). There was a small increase in the percentage of respondents who held science-based attitudes from pre- to posttraining for one item (Q13). For the remaining two items (Q6, Q7), however, there was a slight decrease in agreement with science after training. In sum, a majority of respondents reported attitudes in line with science after training for 13 of the 16 items (81%).

Attitudes toward interviewing: inferential statistics

For all attitude items, participants reported their perceptions toward various aspects of interviewing using 5-point scales. Questions were worded both positively and negatively (i.e. attitudes in line with scientific consensus and attitudes opposing scientific consensus). Average ratings on the pre- and post-training questionnaires are displayed in Table 3. Since we performed inferential statistics on each individual item in this section, a Bonferroni correction was applied to our interpretations regarding significance (0.05/16 = 0.003). In the following sections, the items are categorized by the magnitude and direction of the difference in responses from pre- to post-training. Interpretations of effect sizes were guided by Cohen (1988).

Tier 1
Items with a significant shift in attitudes. There were six items for which Tier 1 respondents showed a statistically significant attitude shift. In all cases, the direction of the attitude shift was closer to what is recommended in the literature – in other words, the officers' attitudes shifted closer to scientific consensus. The effect of training on attitude ratings for the item 'training in detecting deception can vastly improve one's ability to tell whether or not someone is lying' (Q16; $t_{(69)} = 7.63$, $p < .001$, $r = .61$, $d = 0.91$) was large. A medium effect of training on attitude ratings was seen for the items 'there are certain nonverbal cues that determine whether or not someone is being deceptive' (Q2; $t_{(69)} = 6.00$, $p < .001$, $r = .41$, $d = 0.72$) and 'implying leniency to a suspect is fine, as long as you don't explicitly promise them something' (Q12; $t_{(69)} = 5.43$, $p < .001$, $r = .06$, $d = 0.65$). The change in attitude ratings for the items 'once you form a memory of an event, that memory does not change' (Q4; $t_{(69)} = 4.19$, $p < .001$, $r = .03$, $d = 0.50$), 'memory is like a video recorder – events are recorded accurately into memory and can be played back later' (Q1; $t_{(69)} = 3.96$, $p < .001$, $r = .61$, $d = 0.47$), and 'people don't just go along with what other people tell them to do' (Q9; $t_{(69)} = 2.93$, $p = .005$, $r = .40$, $d = 0.35$), was small. Across all six items, an average of 18.2% more respondents agreed with scientific consensus after training (Figure 3).

Table 3. Mean Scores for Attitude Items on Pre- and Post-Training Questionnaire by Tier.

Question	Tier 1		Tier 2	
	Pre	Post	Pre	Post
Q1: Memory is like a video recorder – events are recorded accurately into memory and can be played back later	2.37 (1.04)	1.91 (0.96)	2.11 (0.97)	1.79 (0.91)
Q2: There are certain non-verbal cues that determine whether or not someone is being deceptive	4.10 (0.61)	3.49 (0.90)	3.59 (1.00)	2.42 (0.92)
Q3: Eyewitness testimony can be affected by how the questions asked to the witness are worded	4.06 (0.47)	4.31 (0.58)	4.16 (0.43)	4.26 (0.45)
Q4: Once you form a memory of an event, that memory does not change	2.06 (0.77)	1.57 (0.55)	1.77 (0.61)	1.68 (0.47)
Q5: Most suspects will not confess unless they are pressured into doing so by a police officer	2.14 (0.66)	2.07 (0.80)	2.19 (0.88)	1.95 (0.84)
Q6: The rate of memory loss for an event is greatest right after the event and then levels off over time	3.04 (0.98)	2.96 (1.09)	3.12 (0.93)	2.92 (1.05)
Q7: Most people do not fully understand their legal rights, even though most people tend to waive them	3.44 (0.90)	3.33 (0.99)	3.09 (1.03)	3.03 (0.92)
Q8: Eyewitness testimony reflects both information about the witnessed event and information obtained from outside sources (e.g. media, police, etc.).	3.56 (0.89)	3.81 (0.77)	3.84 (0.65)	3.71 (0.87)
Q9: People don't just go along with what other people tell them to do	2.43 (0.80)	2.13 (0.76)	2.18 (0.69)	2.00 (0.70)
Q10: The ability of people to tell whether or not someone is being deceptive is higher than 50%.	2.78 (0.95)	2.64 (0.89)	2.55 (0.85)	2.16 (0.82)
Q11: No one would admit to doing something they did not do unless they were beaten or tortured	1.71 (0.62)	1.66 (0.59)	1.98 (0.51)	1.68 (0.58)
Q12: Implying leniency to a suspect is fine, as long as you don't explicitly promise them something	3.54 (0.79)	2.74 (1.03)	2.88 (1.10)	2.27 (1.05)
Q13: All interviewees should be treated the same, regardless of any characteristics they may have (e.g. age, intelligence, crime committed)	3.28 (1.26)	3.57 (1.23)	3.32 (1.39)	3.34 (1.46)
Q14: When paying attention to something else, people generally do not notice when something unexpected comes into their view	3.61 (0.83)	4.04 (0.69)	3.28 (0.93)	3.46 (0.99)
Q15: A witness who has experienced a traumatic event should be interviewed immediately after that event to ensure their account is accurate	2.49 (0.93)	2.31 (1.03)	2.02 (0.98)	1.79 (0.62)
Q16: Training in detecting deception can vastly improve one's ability to tell whether or not someone is lying	4.03 (0.67)	3.31 (0.99)	3.26 (0.91)	2.34 (0.94)

NOTE: Items in italics are those that are not in line with scientific consensus. Also note that in our scale, 1 = *Strongly Disagree*, and 5 = *Strongly Agree*.

Items with no significant attitude shift: attitudes already in line with science

There were three items for which a vast majority (i.e. > 75%) of Tier 1 trainees reported attitudes on the pretest that were already consistent with the available science, and remained that way on the post-test. These items were 'eyewitness testimony can be affected by how the questions asked to the witness are worded' (Q3; $d = 0.35$), 'no one would admit to doing something they did not do unless they were beaten or tortured' (Q11; $d = 0.08$), and 'most suspects will not confess unless they are pressured into doing so by a police officer' (Q5; $d = 0.07$). The changes in attitude from pre- to post-training were not significant, but in all cases, scores trended in the direction of being even more in line with scientific knowledge.

Items with no significant attitude shift: Unclear Interpretation of attitudes. For Tier 1 trainees, seven attitude items had responses that were difficult to interpret. The following items had responses that trended toward being more in line with science from pre- to post-training: 'eyewitness testimony reflects both information about the witnessed

event and information obtained from outside sources' (Q8; $d = 0.25$), 'the ability of people to tell whether or not someone is being deceptive is higher than 50%' (Q10; $d = 0.10$), 'when paying attention to something else, people generally do not notice when something unexpected comes into their view' (Q14; $d = 0.48$), and 'a witness who has experienced a traumatic event should be interviewed immediately after that event to ensure their account is accurate' (Q15; $d = 0.14$). Conversely, three items had attitude ratings that trended further away from being in line with science from pre- to post-training: 'the rate of memory loss for an event is greatest right after the event and then levels off over time' (Q6; $d = 0.04$), 'most people do not fully understand their legal rights, even though most people tend to waive them' (Q7; $d = 0.14$), and 'all interviewees should be treated the same, regardless of any characteristics they may have (e.g. age, intelligence, crime committed)' (Q13; $d = 0.22$). However, in all of these cases, scores were close to the midpoint of the scale and differences were non-significant, which make it difficult to interpret much from these results.

Tier 2

Items with a significant shift in attitudes. There were three items for which Tier 2 respondents showed a significant attitude shift; in all cases, the officers' attitudes shifted closer to scientific knowledge. The effect of training on attitude ratings for the items 'there are certain non-verbal cues that determine whether or not someone is being deceptive' (Q2; $t_{(35)} = 8.47$, $p < .001$, $r = .50$, $d = 1.41$) and 'training in detecting deception can vastly improve one's ability to tell whether or not someone is lying' (Q16; $t_{(34)} = 4.70$, $p < .001$, $r = .31$, $d = 0.80$) were large, while the effect of training on the item 'implying leniency to a suspect is fine, as long as you don't explicitly promise them something' (Q12; $t_{(34)} = 3.40$, $p = .002$, $r = .25$, $d = 0.58$) was medium-sized. Across all three items with a significant change, an average of 30% more respondents agreed with scientific consensus after training (Figure 3).

Items with no significant attitude shift: attitudes already in line with science. There were seven items for which a vast majority (i.e. > 75%) of Tier 2 trainees reported attitudes on the pretest that were already consistent with the available science, and remained that way on the post-test. These items were 'memory is like a video recorder – events are recorded accurately into memory and can be played back later' (Q1; $d = 0.42$), 'eyewitness testimony can be affected by how the questions asked to the witness are worded' (Q3; $d = 0.26$), 'once you form a memory of an event, that memory does not change' (Q4; $d = 0.19$), 'eyewitness testimony reflects both information about the witnessed event and information obtained from outside sources' (Q8; $d = 0.10$), 'people don't just go along with what other people tell them to do' (Q9; $d = 0.24$), 'no one would admit to doing something they did not do unless they were beaten or tortured' (Q11; $d = 0.50$), and 'a witness who has experienced a traumatic event should be interviewed immediately after that event to ensure their account is accurate' (Q15; $d = 0.31$).

Items with no significant attitude shift: Unclear Interpretation of attitudes. For Tier 2 trainees, six items had responses that were difficult to interpret. The following four items had responses that trended toward being more in line with science from pre- to post-training: 'most suspects will not confess unless they are pressured into doing so by a

police officer' (Q5; $d = 0.35$), 'the ability of people to tell whether or not someone is being deceptive is higher than 50%' (Q10; $d = 0.38$), 'all interviewees should be treated the same, regardless of any characteristics they may have (e.g. age, intelligence, crime committed)' (Q13; $d = 0.04$), and 'when paying attention to something else, people generally do not notice when something unexpected comes into their view' (Q14; $d = 0.08$). Conversely, two items had attitude ratings that trended further away from being in line with science from pre- to post-training: 'the rate of memory loss for an event is greatest right after the event and then levels off over time' (Q6; $d = 0.22$), and 'most people do not fully understand their legal rights, even though most people tend to waive them' (Q7; $d = 0.11$).

Feedback about training

Generally, responses to questions asking trainees about their experience during the three-day PEACE training were positive. For questions 1-8, participants were asked to rate their level of agreement with various statements about their time in training (*1 = strongly disagree, 5 = strongly agree*). Question 9 asked participants to rate the overall quality of their experience in PEACE training (*1 = poor, 5 = excellent*). A content analysis approach was used to organize responses to four open-ended questions (only responses provided by five or more trainees are reported below).

Tier 1

We found that at least 88% of Tier 1 trainees agreed with each aspect of training quality (selected 4 or 5 on the scale). The highest average score for Tier 1 trainees was for Question 6 ($M = 4.75$, $SD = 0.47$), which asked participants to rate their agreement with the statement that the trainers were knowledgeable. The lowest score was for Question 3 ($M = 4.18$, $SD = 0.77$), which asked participants to rate their agreement with the statement that the material was engaging and kept their attention. The overall mean score for questions 1–8 was 4.51 ($SD = 0.42$), indicating a positive attitude toward the training program. Question 9, which asked participants to rate the quality of their experience in PEACE training, had a mean score of 4.45 ($SD = 0.63$).

In response to the first open-ended question, which asked what they liked most about the training program, most Tier 1 respondents said that they appreciated learning a new and effective method of conducting interviews ($n = 24$). Other aspects of the program that Tier 1 participants liked included the practical exercises ($n = 15$), the note-taking approach taught during training ($n = 8$), the psychology and research behind the PEACE model ($n = 6$), and the instructors ($n = 5$). The most common response to the second question asking what trainees liked the least about the program was that the material was dry ($n = 11$). The next most common responses were that there was nothing they did not like ($n = 6$) and that there was a lack of real-world interview examples and hands-on training ($n = 5$). For the third question, which solicited suggestions for improvement, the most popular response was that there should be more time spent on practical exercises ($n = 14$). Another popular response was the addition of examples of real PEACE interviews ($n = 9$). In response to the fourth question asking respondents to provide additional comments, trainees commented that the instructors were knowledgeable ($n = 6$), and that they enjoyed the training and found it useful *($n = 5$)*.

Tier 2

At least 86% of Tier 2 trainees agreed with each aspect of training quality (selected 4 or 5 on the scale). The highest average score for Tier 2 trainees was for Question 2 ($M = 4.59$, $SD = 0.60$), which asked participants to rate their agreement with the statement that the topics covered were relevant to them. The lowest score was for Question 4 ($M = 4.24$, $SD = 0.90$), which asked participants to rate their agreement with the statement that the content was well-organized and easy to follow. The overall mean score for questions 1–8 was 4.44 ($SD = 0.63$), indicating a positive attitude toward the training program. Question 9, which asked participants to rate the quality of their experience in PEACE training, had a mean score of 4.41 ($SD = 0.80$).

When asked what they liked most about the training program, the most common response from Tier 2 trainees was that they enjoyed the practical exercises ($n = 15$). Other aspects of the program that Tier 2 trainees liked included learning a new and effective method of conducting interviews ($n = 7$), the structure of PEACE interviews ($n = 8$), and the approach to asking questions taught during training ($n = 3$). The most common response to the second question asking what trainees liked the least about the program was the length of training ($n = 9$). The next most common responses were complaints about PEACE being a replacement for old methods ($n = 5$) and that there was nothing they disliked ($n = 5$). In terms of the third question, which asked for suggestions for improvement, the most popular responses were that improvements should be made to the mock suspect interview component of training ($n = 6$) and that the training session should be shorter ($n = 6$). Some officers also mentioned that they wished there were more practical exercises ($n = 5$). Lastly, when asked to provide any additional comments, several trainees commented that they enjoyed the training ($n = 6$)

Discussion

The current program evaluation was conducted to measure the impact of Tier 1 and Tier 2 PEACE interview training on the knowledge and attitudes of police officers from the VSP. Overall, the results demonstrated that trainees showed improvements in their *knowledge*, and some of their *attitudes,* about police interviewing and other related concepts. Trainee perceptions of the program were also overwhelmingly positive. Collectively, these results suggest that the program is being implemented as intended; that is, relevant concepts are being taught in a way that is meaningful to the trainees. The results also indicate that the short-term outcomes of the program are being achieved. Trainees are learning about important concepts related to evidence-based interviewing and are also shifting their view of interviewing to be more in line with scientific consensus.

Trainees in both Tier 1 and Tier 2 were more knowledgeable about evidence-based interviewing after training than they were before training. This result indicates that VSP PEACE training is teaching what it was intended to teach, and doing so in an effective manner. However, not all knowledge items were more likely to be answered correctly post-training. In Tier 1, two items from the multiple-choice section (i.e. Q9, Q14) and five from the true/false section (i.e. Q20, Q24, Q25, Q28, Q29) were answered correctly by fewer officers after the completion of training than before training. However, the decrease in correct responses was only greater than 5% for three of these questions (i.e. Q20, 25, 29). In Tier 2, this was the case for three items from the multiple-choice

section (i.e. Q1, Q8, Q14) and five from the true/false section (i.e. Q20, Q22, Q23, Q28, Q29), but only three had a decrease in correct responses greater than 5% (i.e. Q14, Q23, Q29). The lack of knowledge in these areas may be a function of the manner in which specific material was taught during training (e.g. mismatch in wording between training and questionnaire). It also could be that the particular material covered by those questions is complex and difficult to understand, even with instruction. The aforementioned findings will be used to improve and strengthen the material and teaching resources for the program.

The effect of training on officer attitudes was somewhat inconsistent for both training groups. Importantly, scores for all of the items that showed significant changes from pre- to post-training shifted closer to scientific consensus for both Tier 1 and Tier 2 trainees. As well, there were some items on which officers from both groups had already reported perceptions in agreement with science before training; this cannot be attributed to the training program of course, but it is nonetheless a positive finding. The percentage of officers who reported perceptions in line with science after training varied greatly depending on the subject matter of the item. For the most part, items related to eyewitness testimony, memory processes, and contamination of memory were endorsed by the vast majority of officers in both tiers on the post-training questionnaire, suggesting that the training program did well at teaching trainees about the role of memory in interviewing (see Kassin et al., 2001; Magnussen & Melinder, 2012; Simons & Chabris, 2011 for data on the attitudes of psychologists, legal professionals, and laypeople toward memory and eyewitness issues). Similarly, the results suggest that issues regarding false confessions, coercion, and persuasion were covered well by the training administrators, as each of the items covering those concepts were endorsed by more than half of respondents in both tiers (see Kassin et al., 2018 for views of experts on similar issues). Conversely, most items about deception detection were only endorsed by a small minority of Tier 1 officers after training; this is likely because detecting deception was not a major component of Tier 1 training. Tier 2 officers had a higher level of agreement with these items post-training (> 50%), which makes sense given that deception detection was a major topic covered in their suspect interview training. Another surprising finding was that Q7 (i.e. 'people are not aware of their legal rights') was endorsed by more respondents before training than after training by officers in both tiers, with Tier 2 officers showing even less agreement than Tier 1. This finding reflects the lack of knowledge regarding Miranda rights reported among criminal defendants and university students (Rogers et al., 2010). This suggests that the training administrators should perhaps consider a different approach to their instruction in this area.

Trainee reactions to the PEACE training program were positive. For both Tier 1 and Tier 2, mean scores for every specific item fell between 'agree' and 'strongly agree' on the scale, and the mean overall experience rating was between 'good' and 'excellent'. The responses to open-ended items revealed some of the strengths and weaknesses of the program. For example, trainees seemed to particularly enjoy the practical aspects of the training program, and suggested that program administrators add more hands-on components. They also praised the instructors for their knowledge and competence, and enjoyed learning about a new, research-based interviewing method. Complaints about the length of sessions and dry material were relatively common, as well as suggestions to add examples of PEACE interviews in the real world (e.g. show videos of real

interviews). Overall, most of these responses were positive or constructive, and will undoubtedly help to inform improvements to the next iteration of the program.

The aforementioned findings must be considered in light of some limitations. First, we would like to point out concerns that come with the use of reaction data: namely, that such data does not add value to the overall findings and is not indicative of a program's success. However, since we are conducting a comprehensive evaluation of the PEACE training program, as opposed to a more basic research endeavor, we believe there is value in reporting participants' reactions to training. Conducting an evaluation of a program requires assessing its success from various angles. While it is critical to collect data that speaks to the improvement in learning, knowledge, attitudes, and behavior that result from the program, information about how participants felt about the program can add another layer of evidence that speaks to the overall success of the program. Moreover, a study by Alliger et al. (1997) reported there is a moderate relationship between utility judgements (a type of reaction measure) and both learning and behavior/transfer criteria, suggesting that this type of reaction measure is useful to an overall evaluation.

As well, our sample size was small, especially for Tier 2 trainees. As we were collecting data in partnership with the VSP and in concert with their training efforts, we were limited in our ability to access participants (i.e. dependent on their training capacity and could only recruit individuals participating in their training program). Due to the in-depth nature of Tier 2 training in particular, the administrators were only able to train approximately eight officers at a time, over a two month period, which naturally limited the amount of available data. A larger sample would have mitigated potential issues such as the impact of participant guessing, however given that we saw a substantial increase in knowledge from pre- to post-training, we believe it is unlikely that our findings were caused by an improvement in trainees' guessing abilities as opposed to an improvement in their knowledge of interviewing. Nevertheless, future studies of larger-scale training operations are warranted to ensure generalizability of our findings.

Additionally, there may be some concerns about testing effects. Whereas Tier 2 participants completed their pre- and post-training surveys roughly two months apart, Tier 1 trainees only had three days in between completing each survey. While this was not an ideal situation, it was necessary due to the logistical requirements for Tier 1 training. For instance, the training administrators had to book classroom space for the training sessions, while working with police barracks from across the state to find dates that would work for all participating officers (~25-30 per session). After the training session was completed, it would have been difficult, if not impossible, to get the trainees back together again to complete the post-training survey. Since we wanted to keep the environments in which trainees completed the pre- and post-training surveys as consistent as possible to avoid the influence of extraneous variables, the lack of time between the two tests was unavoidable. Researchers conducting similar research in the future may consider adopting an entirely online survey administration process to mitigate this issue (i.e. administer the pre-test prior to training and the post-test after its completion).

We ideally would have liked to pilot the questionnaire prior to beginning the evaluation, but due to the tight timeline of our evaluation (i.e. we were not in control of the schedule for training and had to have all materials ready by the time training began), we were unable to conduct a pilot study. Without a pilot, we cannot be sure whether

or not there were any problematic questions included in the questionnaire (e.g. complicated wording), and thus this must be mentioned as a limitation. Future evaluations of the PEACE framework should strongly consider the use of pilot studies to validate study materials. Nevertheless, since our questions were derived directly from training materials and validated research instruments, we are confident that there were no major problems that would have impacted our results. Particularly, our attitude questions were analyzed individually, so any issues with a particular question would not impact the results for any other question, while our knowledge check questions were derived directly from source materials that were used to develop the training program.

This study constituted the first phase of the overall evaluation of VSP's PEACE interview training program. In following with the four-level framework for training program evaluation (Kirkpatrick & Kirkpatrick, 2006), there are a number of outcomes that still need to be assessed. To assess knowledge retention, the officers who completed training will be assessed again one-year post-training. To determine whether the treatment of interviewees has changed as a result of training, we will conduct a survey of witnesses who were interviewed by both trained and untrained members of the VSP. To determine whether there is an improvement in interview behaviors post-training (i.e. more appropriate questions and fewer inappropriate questions, less coercion, higher quality and quantity of information gathered), another intermediate outcome of the program, transcripts of suspect and witness interviews conducted by trained and untrained officers will be coded and analyzed. Finally, we will survey the legal community in Vermont (e.g. prosecutors, defense attorneys, judges, advocates) to assess any perceived changes in interview culture as a way to determine the overall result of the program. Recruitment of legal community members is underway for an initial survey regarding thoughts about the current VSP interview culture; a second will be administered at least one year after all training is completed.

Based on the results of Phase I of the PEACE Model training program evaluation, it is clear that the program is meeting its short-term outcomes. Trainees knew more about interviewing and reported interviewing-related perceptions in line with scientific consensus after training, and reported that participating in the training program was a positive experience. These findings are a first step in fully understanding the effectiveness of PEACE training; the remaining evaluation phases will help uncover whether the program is capable of achieving its more long-term goals of behavior and culture change, which have been proven to be difficult to achieve in previous research (Clarke et al., 2011; Clifford & George, 1996; Dando et al., 2009; Griffiths & Milne, 2006; Kebbell & Milne, 1998; MacDonald et al., 2017; Walsh & Milne, 2008). Ultimately, this evaluation will provide concrete evidence as to whether or not the PEACE Model training program is effective. For now, it can be said that the program is an effective tool for educating police officers on science-based interviewing practices, and that by the end of their program, members of the VSP will have a strong knowledge base and a positive mindset that will help the organization on their quest to shift to a more ethical interviewing culture. What is more, this evaluation provides some promising preliminary evidence to suggest that the PEACE model may be a viable interviewing method for other police agencies to adopt, as it has clear short-term benefits. The next steps of the evaluation will provide more evidence as to the long-term benefits of the program and whether it would be worthwhile to implement.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Data availability statement

The data that support the findings of this study are available from the corresponding author, [redacted], upon reasonable request.

References

Alison, L., Alison, E., Noone, G., Elntib, S., & Christiansen, P. (2013). Why tough tactics fail and rapport gets results: Observing Rapport-Based Interpersonal Techniques (ORBIT) to generate useful information from terrorists. *Psychology, Public Policy and Law*, *19*(4), 411–431. https://doi.org/10.1037/a0034564

Alliger, G. M., Tannenbaum, S. I., Bennett, W., Jr., Traver, H., & Shotland, A. (1997). A meta-analysis of relations among training criteria. *Personnel Psychology*, *50*(2), 341–358. https://doi.org/10.1111/j.1744-6570.1997.tb00911.x

Bull, R., & Baker, B. (2019). Obtaining from suspects valid discourse "PEACE"-fully: What role for rapport and empathy? In M. Mason & F. Rock (Eds.), *The discourse of police interviews* (pp. 42–64). University of Chicago Press.

Clarke, C., Milne, R., & Bull, R. (2011). Interviewing suspects of crime: The impact of PEACE training, supervision, and the presence of a supervisor. *Journal of Investigative Psychology and Offender Profiling*, *8*(2), 149–162. https://doi.org/10.1002/jip.144

Clifford, B. R., & George, R. (1996). A field evaluation of training in three methods of witness/victim investigative interviewing. *Psychology, Crime and Law*, *2*(3), 231–248. https://doi.org/10.1080/10683169608409780

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Collins, R., Lincoln, R., & Frank, M. G. (2002). The effect of rapport in forensic interviewing. *Psychiatry, Psychology, and Law*, *9*(1), 69–78. https://doi.org/10.1375/pplt.2002.9.1.69

Cook, A. N., & Roesch, R. (2012). 'Tough on crime' reforms: What psychology has to say about the recent and proposed justice policy in Canada. *Canadian Psychology/Psychologie Canadienne*, *53*(3), 217–225. https://doi.org/10.1037/a0025045

Dando, C., Wilcock, R., & Milne, R. (2009). The cognitive interview: Novice police officers' witness/victim interviewing practices. *Psychology, Crime & Law*, *15*(8), 679–696. https://doi.org/10.1080/10683160802203963

Davis, D., & Loftus, E. F. (2017). Internal and external sources of misinformation in adult witness memory. In M. P. Toglia, J. D. Read, D. F. Ross, & R. C. L. Lindsay (Eds.), *The handbook of eyewitness psychology: Volume I: Memory for events* (pp. 195–238). Psychology Press.

de Quervain, D. J., Roozendaal, B., Nitsch, R. M., McGaugh, J. L., & Hock, C. (2000). Acute cortisone administration impairs retrieval of long-term declarative memory in humans. *Nature Neuroscience*, *3*(4), 313–314. https://doi.org/10.1038/73873

Evans, J. R., Meissner, C. A., Ross, A. B., Houston, K. A., Russano, M. B., & Horgan, A. J. (2013). Obtaining guilty knowledge in human intelligence interrogations: Comparing accusatorial and information-gathering approaches with a novel experimental paradigm. *Journal of Applied Research in Memory and Cognition*, *2*(2), 83–88. https://doi.org/10.1016/j.jarmac.2013.03.002

Fisher, R. P., & Geiselman, R. E. (1992). *Memory enhancing techniques for investigative interviewing: The cognitive interview*. Charles C Thomas Publisher.

Fisher, R. P., Geiselman, R. E., & Amador, M. (1989). Field test of the cognitive interview: Enhancing the recollection of actual victims and witnesses of crime. *Journal of Applied Psychology*, *74*(5), 722–727. https://doi.org/10.1037/0021-9010.74.5.722

Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2011). *Program evaluation: Alternative approaches and practical guidelines*. Pearson Education.

Frechtling, J. A. (2007). *Logic modeling methods in program evaluation* (Vol. 5).Wiley.

George, R., & Clifford, B. (1992). Making the most of witnesses. *Policing (bradford, England)*, *8*(3), 185–198.

Gershoff, E. T. (2002). Corporal punishment by parents and associated child behaviors and experiences: A meta-analytic and theoretical review. *Psychological Bulletin*, *128*(4), 539–579. https://doi.org/10.1037/0033-2909.128.4.539

Giebels, E., & Taylor, P. J. (2009). Interaction patterns in crisis negotiations: Persuasive arguments and cultural differences. *Journal of Applied Psychology*, *94*(1), 5–19. https://doi.org/10.1037/a0012953

Goodman-Delahunty, J., & Martschuk, N. (2020). Securing reliable information in investigative interviews: Coercive and noncoercive strategies preceding turning points. *Police Practice and Research*, *21*(2), 152–171. https://doi.org/10.1080/15614263.2018.1531752

Griffiths, A., & Milne, R. (2006). Will it all end in tiers? Police interviews with suspects in Britain. In T. Williamson (Ed.), *Investigative interviewing* (pp. 167–189). Routledge.

Gudjonsson, G. H. (2003). *The psychology of interrogations and confessions: A handbook*. Wiley.

Holmberg, U., & Christianson, SÅ. (2002). Murderers' and sexual offenders' experiences of police interviews and their inclination to admit or deny crimes. *Behavioral Sciences and the Law*, *20*(1-2), 31–45. https://doi.org/10.1002/bsl.470

Kassin, S. M., Drizin, S. A., Grisso, T., Gudjonsson, G. H., Leo, R. A., & Redlich, A. D. (2010). Police-induced confessions: Risk factors and recommendations. *Law and Human Behavior*, *34*(1), 3–38. https://doi.org/10.1007/s10979-009-9188-6

Kassin, S. M., Redlich, A. D., Alceste, F., & Luke, T. J. (2018). On the general acceptance of confessions research: Opinions of the scientific community. *American Psychologist*, *73*(1), 63–80. https://doi.org/10.1037/amp0000141

Kassin, S. M., Tubb, V. A., Hosch, H. M., & Memon, A. (2001). On the general acceptance of eyewitness testimony research: A new survey of the experts. *American Psychologist*, *56*(5), 405–416. https://doi.org/10.1037/0003-066X.56.5.405

Kebbell, M., Hurren, E. J., & Mazerolle, P. (2006). Sex offenders' perceptions of how they were interviewed. *Canadian Journal of Police & Security Services*, *4*, 67–75.

Kebbell, M. R., & Milne, R. (1998). Police officers' perceptions of eyewitness performance in forensic investigations. *The Journal of Social Psychology*, *138*(3), 323–330. https://doi.org/10.1080/00224549809600384

Kirkpatrick, D. L. (1959). Techniques for evaluating training programs. *Journal of the American Society of Training Directors*, *13*, 3–9.

Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating training programs: The four levels* (3rd ed.). Berrett-Koehler Publishers.

Köhnken, G., Milne, R., Memon, A., & Bull, R. (1999). The cognitive interview: A meta-analysis. *Psychology, Crime and Law*, *5*(1-2), 3–27. https://doi.org/10.1080/10683169908414991

Lipton, J. P. (1977). On the psychology of eyewitness testimony. *Journal of Applied Psychology*, *62*(1), 90–95. https://doi.org/10.1037/0021-9010.62.1.90

MacDonald, S., Snook, B., & Milne, R. (2017). Witness interview training: A field evaluation. *Journal of Police and Criminal Psychology*, *32*(1), 77–84. https://doi.org/10.1007/s11896-016-9197-6

Magnussen, S., & Melinder, A. (2012). What psychologists know and believe about memory: A survey of practitioners. *Applied Cognitive Psychology*, *26*(1), 54–60. https://doi.org/10.1002/acp.1795

Marin, R. J. (2004). *Admissibility of statements*. Canada Law Book.

Marquis, K. H., Marshall, J., & Oskamp, S. (1972). Testimony validity as a function of question form, atmosphere, and item difficulty. *Journal of Applied Social Psychology*, *2*(2), 167–186. https://doi.org/10.1111/j.1559-1816.1972.tb01270.x

Meissner, C. A., Redlich, A. D., Michael, S. W., Evans, J. R., Camilletti, C. R., Bhatt, S., & Brandon, S. (2014). Accusatorial and information-gathering interrogation methods and their effects on true and false confessions: A meta-analytic review. *Journal of Experimental Criminology*, *10*(4), 459–486. https://doi.org/10.1007/s11292-014-9207-6

Memon, A., Meissner, C. A., & Fraser, J. (2010). The cognitive interview: A meta-analytic review and study space analysis of the past 25 years. *Psychology, Public Policy, and Law*, *16*(4), 340–372. https://doi.org/10.1037/a0020518

Miller, W. R., & Rollnick, S. (2012). *Motivational interviewing: Helping people change*. Guilford Press.

Milne, R., & Bull, R. (2003). Interviewing by the police. In D. Carson, & R. Bull (Eds.), *Handbook of psychology in legal contexts* (pp. 111–125). Wiley.

NSLEC. (2004). *Practical guide to investigative interviewing*. National Centre for Policing Excellence.

Roberts, K. P., Lamb, M. E., & Sternberg, K. J. (2004). The effects of rapport-building style on children's reports of a staged event. *Applied Cognitive Psychology*, *18*(2), 189–202. https://doi.org/10.1002/acp.957

Rogers, R., Rogstad, J. E., Gillard, N. D., Blackwood, H. L., Drogin, E. Y., & Shuman, D. W. (2010). "Everyone knows their Miranda rights": Implicit assumptions and countervailing evidence. *Psychology, Public Policy, and Law*, *16*(3), 300–318. https://doi.org/10.1037/a0019316

Sanders, G. (1986). On increasing the usefulness of eyewitness research. *Law and Human Behaviour*, *10*(4), 333–335. https://doi.org/10.1007/BF01047345

Shepherd, E. (2007). *Investigative interviewing: The conversation management approach*. Oxford University Press.

Simons, D. J., & Chabris, C. F. (2011). What people believe about how memory works: A representative survey of the US population. *PLoS ONE*, *6*(8). https://doi.org/10.1371/journal.pone.0022757

Snook, B., Brooks, D., & Bull, R. (2015). A lesson on interrogations from detainees: Predicting self-reported confessions and cooperation. *Criminal Justice and Behavior*, *42*(12), 1243–1260. https://doi.org/10.1177/0093854815604179

Snook, B., Luther, K., Quinlan, H., & Milne, R. (2012). Let 'em talk! A field study of police questioning practices of suspects and accused persons. *Criminal Justice and Behavior*, *39*(10), 1328–1339. https://doi.org/10.1177/0093854812449216

Walsh, D., & Bull, R. (2010). Interviewing suspects of fraud: An in-depth analysis of interviewing skills. *The Journal of Psychiatry & Law*, *38*(1-2), 99–135. https://doi.org/10.1177/009318531003800106

Walsh, D., & Bull, R. (2012). Examining rapport in investigative interviews with suspects: Does its building and maintenance work? *Journal of Police and Criminal Psychology*, *27*(1), 73–84. https://doi.org/10.1007/s11896-011-9087-x

Walsh, D., & Milne, R. (2008). Keeping the PEACE? A study of investigative interviewing practices in the public sector. *Legal and Criminological Psychology*, *13*(1), 39–57. https://doi.org/10.1348/135532506X157179