Journal of Forensic Psychiatry & Psychology

Catching liars: training mental health and legal professionals to detect high-stakes lies

Julia Shaw [a] , Stephen Porter [a] & Leanne ten Brinke [a]

[a] Department of Psychology, University of British
Columbia Okanagan, Kelowna, Canada

PLEASE SCROLL DOWN FOR ARTICLE

whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Routledge
Taylor & Francis Group

Catching liars: training mental health and legal professionals to detect high-stakes lies

Julia Shaw*, Stephen Porter and Leanne ten Brinke

Department of Psychology, University of British Columbia Okanagan, Kelowna, Canada

Although the ability to detect deception is critical in many professional contexts, most observers (including professional lie-catchers) are able to identify deceivers at the level of chance only. Further, almost all studies of deception detection have used low-stakes deception scenarios in determining deceptive behavior and training effectiveness. We evaluated the effectiveness of a comprehensive, empirically based full-day training workshop in improving the ability of 42 legal and mental health professionals to detect extremely high-stakes emotional lies. Their ability to discriminate sincere and insincere pleaders was measured at baseline and post-training. Overall, accuracy increased significantly from $M = 46.4$ to 80.9%. We cautiously suggest that training professionals to apply empirically validated methods to deception detection can increase their ability to correctly discriminate between liars and truth-tellers. Strategies to facilitate the detection of deception via the development of training programs are discussed.

Keywords: deception detection; training; high-stakes lies

Deception is a pervasive problem throughout the legal system (e.g. Rogers, Salekin, Sewell, Goldstein, & Leonard, 1998; Sullivan, Lange, & Dawes, 2007). Despite the judiciary's view that detecting lies is a straightforward matter best guided by simple common sense (e.g. Supreme Court of Canada in *R. v.* Marquard, 1993), empirical research suggests that it is a flawed process with errors occurring in nearly half of all assessments (e.g. Blair, Levine, & Shaw, 2010; Bond & DePaulo, 2006; Vrij, Granhag, & Porter, 2010). While one might predict that professionals who need to detect deception on a daily basis (e.g. judges, police officers, etc.) would outperform laypersons, they too typically perform at or below chance in judging the credibility of speakers (e.g. Hartwig, Granhag, Strömwall, & Vrij, 2004; Porter & ten Brinke, 2010; Vrij, 2008a).

---

*Corresponding author. Email: Julia.Shaw@ubc.ca

This low rate of deception detection accuracy among professionals with a vested interest in the task has been attributed to faulty assumptions about cues to deception that contradict the scientific literature (Porter & ten Brinke, 2010; Stromwall & Granhag, 2003; Stromwall, Granhag, & Hartwig, 2004; Vrij, 2008b; Vrij & Granhag, 2007; Vrij, Akehurst, & Knight, 2006). For example, people around the world rely heavily on gaze aversion, nervous behaviors, and other body language cues as signs of lying (The Global Deception Research Team, 2006; see Vrij, 2008a). Other evidence suggests that human biases (such as those relating to first impressions: see Porter & ten Brinke, 2009; Willis & Todorov, 2006), relating to automatic brain processes, compromise the ability to objectively evaluate credibility (e.g. Meissner & Kassin, 2002). Further, other than police, most professional lie-catchers never receive empirically based training in such biases or in credibility assessment (e.g. Porter & ten Brinke, 2010). In the Western world, deception detection training for the police most commonly is based on the Reid model of interrogation (Kassin, 2006). However, when Kassin and Fong (1999) randomly assigned participants to receive Reid-based training or no training, the trainees performed more poorly than controls but became more confident in their assessments.

Another issue, regarding whether individual differences contribute to performance in deception detection, also remains controversial. Some argue that individual differences are negligible (Aamodt & Custer, 2006; Bond & DePaulo, 2008), except that people differ in their biases regarding the veracity of statements, also known as a truth or deception biases (Masip, Garrido, & Herrero, 2009). Further, deception detection performance is unreliable over testing sessions (Leach et al., 2009), arguing against the role of stable individual differences. However, specific studies have indicated that individual differences play a role in deception detection, including age (e.g. O'Sullivan, 2008), profession (Ekman & O'Sullivan, 1991; O'Sullivan, Frank, Hurley, & Tiwana, 2009), and handedness/hemispheric dominance (Etcoff, Ekman, Magee, & Frank, 2000; Porter, Campbell, Stapleton, & Birt, 2002). As with research examining behaviors of deceivers, most research on observers/judges has used undergraduate samples (Porter & ten Brinke, 2010), and more research is needed on the abilities of professionals in legal settings to further examine potential individual differences. Some researchers have claimed that certain professional groups (e. g. see Ekman & O'Sullivan, 1991; Ekman, O'Sullivan, & Frank, 1999; Mann, Vrij, & Bull, 2002) and individuals, or 'wizards' (O'Sullivan & Ekman, 2004) are considerably better than chance at detecting lies (cf. Bond & Uysal, 2007). A study by Bond (2008) showed videotapes of paroled offenders either lying or telling the truth to law enforcement personnel and students. Two 'experts' were identified, both female Native American correctional officers who were over 80% accurate in the first assessment and 90% accurate in a second assessment. The results (including eye-tracking analyses) indicated that these experts relied primarily on body language/non-verbal cues to make accurate decisions. Other recent research suggests that the use of low-stakes lie scenarios may have led to

an underestimate of deception detection performance by criminal justice staff (O'Sullivan et al., 2009). Thus, a small number of studies with criminal deceivers as targets (as in the present study) suggests the possibility of individual differences perhaps overlooked in research with lies by non-forensic samples (e.g. Leach et al., 2009). On the other hand, whether or not we have underestimated the ability of professional lie-catchers, they certainly miss many of the lies confronting them; for example, psychopaths, who have a high recidivism rate and lie prodigiously, are more than twice as likely as their non-psychopathic counterparts to be granted parole after the parole interview (Porter, ten Brinke, & Wilson, 2009), likely attributable to convincing acting jobs.

In addition to the potential importance of the type of observer, the type of lie being communicated also may play a pivotal role in deception detection accuracy. Motivation level and personal relevance of the lie have been shown to affect the behavior of deceivers (DePaulo et al., 2003; DePaulo, Lanier, & Davis, 1983; Gustrafson & Orne, 1965; Vrij & Heaven, 1999). High-stakes lies have personal relevance and involve significant consequences to the deceiver and deception recipient, while low-stakes lies carry no or minimal consequences (O'Sullivan et al., 2009). Porter and ten Brinke (2010) argued that there likely are major qualitative differences in the deception cues exhibited during these two types of lies. Lies of consequence are generally more difficult to tell than other lies and should be accompanied by more salient behavioral signs, or 'leakage' (the 'motivational impairment effect'; DePaulo & Kirkendol, 1989), discernable to the informed observer. Lying about one's knowledge of a committed or planned crime can be a complex undertaking and enormously stressful for most people. In telling such lies, the liar must concurrently keep the details of a narrative consistent and appear 'credible' to a potentially apprehensive listener whose conclusion about the deceiver's veracity may carry tremendous consequences. Many high-stakes lies are accompanied by powerful emotions, including fear, remorse, anger, or even excitement, that must be inhibited or convincingly faked. Consider the would-be terrorist smiling and chatting politely with airport staff, while covertly feeling intense hatred and contempt towards his intended targets, and perhaps fear of discovery and/or death; or the mother publicly pleading for the safe return of a child who, in reality, she has murdered. Each of these liars must concurrently monitor his/her body language, facial expressions, and stories while dealing with an acute awareness of the potent consequences of getting caught. Unfortunately, most studies of deception detection have involved low- or no-stakes lies as stimuli, and very few have used truly high-stakes lies (Mann, Vrij, & Bull, 2002; Vrij, Mann, Robbins, & Robinson, 2006). As high-stakes lies more appropriately reflect real-world conditions faced by psychologists and legal professionals than low-stakes lies, these differences warrant further investigation.

The forensic relevance and societal consequences of high-stakes deception make the development, dissemination, and evaluation of deception detection training strategies critical goals. Over the past few decades, attempts to train

professionals in deception detection have had mixed results (Bull, 2004), with some being modestly successful (e.g. deTurck, Roman, & Feeley, 1997; Porter, Juodis, ten Brinke, Klein, & Wilson, 2010; Porter, Woodworth, & Birt, 2000; Vrij, 1994; Zuckerman, Koestner, & Colella, 1985), some having no effect (e.g. Köhnken, 1987), and others being detrimental (e.g. Kassin & Fong, 1999). Collectively, however, the literature shows small but reliable success with an average training effect size of $r = .20$ (Frank & Feeley, 2003).

What constitutes the 'best' training approaches in deception detection? Porter and ten Brinke (2010) emphasized the need for a holistic assessment strategy, including training in how to avoid the major pitfalls of deception detection (also see Vrij, 2008), and training on each of verbal, body language, and facial expressions associated with lying. Frank and Freely (2003) summarized six major criteria that contribute to the efficacy of training programs. First, (1) they suggested the importance of the relevance of the materials being trained to the unique types of situations faced by various professional groups, including (2) the importance of training to detect high-stakes lies. They also suggested that (3) training must be understandable and transmissible to professionals who are unfamiliar with the literature on deception. Fourth, (4) they suggested the need to test the effectiveness of each deception training program by giving both a pre- and post-measures of deception detection ability. Finally, (5) they recommended that the need to demonstrate that the material covered generalizes across deception situations and (6) that training has an impact on the decision-making of professionals post-training.

The purpose of the present study was to implement and evaluate the effectiveness of a one-day, empirically based deception detection workshop for legal and mental health professionals using extremely high-stakes lies and considering the recommendations for state-of-the-art training outlined above. Using a within-subjects design, the study examined changes in performance on a high-stakes deception detection task from pre-training to post-training. We were asked to provide training to a large group of primarily forensic psychologists and psychiatrists, but also some legal professionals such as lawyers. The potential impact of training and individual differences on deception detection ability was investigated.

Method

Control study

Participants

Twenty-five undergraduate students (17 females and 8 males) in a third-year psychology class voluntarily participated in the control study. The control study was conducted to ensure that there was no bias in the stimuli (videos) used for the main study. The sample was predominantly Caucasian (17 Caucasian and 8 non-Caucasian), with an average age of 22.1 (range: 20–28). Three participants indicated that they had previously completed training in deception detection.

Materials

The stimuli used for the control study were eight videos depicting family members publically pleading for the safe return of a missing relative (see Mann et al., 2002). In half of the videos, the pleader is actually the killer of the missing person (refer to the description of the main study for more detailed information on the videos).

Procedure

Control group testing occurred in a group format in a small lecture hall room. Prior to viewing the videos, participants were asked to complete a demographics questionnaire and indicate the confidence in their ability to 'tell whether another person is lying' and how adept they felt at telling lies ('not often get caught'). A large projection screen with a professional sound system was used to present the test stimuli. Participants were informed of the nature of the task and told that some of the stimuli were truthful and some deceptive. Stimuli were viewed in the same order by the controls as by the main study participants. Participants indicated that whether in their opinion each speaker was genuine or deceptive and the reasons for their choice, with a one-minute break between the clips. Following the test, feedback was provided to the participants.

Main study

Participants

Forty-two professionals (24 females, 12 males, and 6 undisclosed) attending a 2010 forensic psychiatry conference voluntarily participated in/registered for the credibility training seminar. They included: nineteen forensic psychologists (including one criminal profiler), eight forensic psychiatrists, two correctional service workers, two lawyers, one forensic science expert, and one nurse. Nine participants did not disclose their profession. The number of years of experience in their fields ranged from 1 to 50 years, with a mean of 15.94 (SD = 10.7). The sample was predominantly Caucasian (36 Caucasian, 2 non-Caucasian, and 6 undisclosed), with an average age of 42.3 (range: 24–76). Nine participants indicated that they had previously completed training in deception detection. This sample was considered an excellent representation of individuals working in the field who need to detect deception in a professional context on a regular basis.

Materials

The training program was a full-day (6.5 h) workshop. The workshop followed the general (but updated) approach outlined in Porter et al. (2000, 2010). Further, it considered the major recommendations for optimal training practices in deception detection outlined by Frank and Feeley (2003) and Porter and ten

Brinke (2010), and involved three main components, such as (Part 1) Avoiding the Major Pitfalls in Detecting Deception, (Part 2) How Do Liars Behave?, and (Part 3) An Integrated Approach to Detecting High-Stakes Lies. Throughout the interactive training, video demonstrations and transcripts were used and reviewed. Further, participants practiced detecting deception and feedback was provided throughout the day.

The stimuli used to evaluate pre- and post-deception detection ability were randomly selected from a large database of videos our team has collected from around the English-speaking world over the past several years. These videos show family members publically pleading for the safe return of a missing relative (see Mann et al., 2002). While the authors were familiar with the videos used in the final sample, no clips from these videos were used in the training session. In about half of the videos, the pleader is actually the killer of the missing person. By selecting closed cases in which the fate of the missing person has been established with certainty (e.g. the person is found safe; DNA evidence establishes the guilt or innocence of the pleader), we have been able to establish ground and categorize sincere individuals vs. deceptive killers. In other words, these are truly high-stakes lies with comparable sincere targets. Eight videos were randomly selected and assigned for use in pre-training (four videos: two genuine and two deceptive) and post-training (four videos: two genuine and two deceptive) tests of the ability of participants to detect high-stakes lies. All of the videos show clear, close-up shots of the pleaders faces.

Part (1) *Avoiding the Major Pitfalls in Detecting Deception*: This training segment included a baseline test of the participants' ability to detect high-stakes lies where four videos were shown, two truth-tellers and two deceivers pleading for the return of a missing relative. It also included an overview of deception throughout society, two prevalent assumptions about deception in forensic settings (common sense, and that demeanor evidence is important), and relevant legal cases. Finally, it addressed to what extent and how people are duped successfully with data on the performance of various professional groups, deception detection biases, the psychology of successful deceivers, how humans automatically form judgments of trustworthiness/relevant neuroscience, research on structural elements of facial appearance that influence our impressions of honesty, and the dangerous decisions theory (Porter & ten Brinke, 2009), including relevant real false conviction cases.

Part (2) *How Do Liars Behave?* The second component of the training program was an empirically based information session in which trainees were presented with a review of the latest scientific knowledge on the detection of deception in applied settings (based on Porter and ten Brinke's (2010) and Vrij's (2008a) major conclusions). In other words, whereas Part 1 focused on what not to do to avoid pitfalls, Part 2 was more directive in how to evaluate credibility. The training emphasized a baseline method (see Porter & ten Brinke, 2010) and addressed body language/non-verbal behavior (e.g. illustrators), emotional deception/facial expressions (including detailed analyses of

each universal emotion/underlying musculature and research on how deceptive and genuine facial expressions can be discriminated with videotaped high-stakes lies as examples), and verbal cues/statement analysis (criterion-based content analysis, reality monitoring, with emphasis on utilizing the most reliable criteria according to research).

Part (3) *An Integrated Approach to Detecting High-Stakes Lies.* The final component integrated the material from Part 2, with the use of several real-life high-stakes lies (and truthful stories), including cases in which the first author had been consulted by police in serious crime investigations. An analysis of videotaped and transcribed narratives by (high-stakes) liars and truth-tellers that considered body language, facial expressions, and their statements was pro-vided. Further, participants practiced extensively and were given feedback throughout Part 3 to enhance comprehension and application of the empirical information provided. Finally, the participants were given a post-training test, paralleling the pre-training test. For more information on the deception detec-tion workshop, please refer to Porter et al. (2010).

Procedure

The training (and pre- and post-tests) occurred in a group format in a large training room. Prior to the start of the training session, before videos and train-ing were administered, participants were asked to complete a demographics questionnaire and indicate the confidence in their ability to 'tell whether another person is lying' and how adept they felt at telling lies ('not often get caught'). Two large auditorium projection screens with a professional sound system were used to present the training and test stimuli. For the pre-test, prac-tice videos throughout training, and post-test, participants were informed of the nature of the three types of tasks and told that some of the stimuli were truth-ful and some deceptive. Participants indicated that whether in their opinion each speaker was genuine or deceptive and the reasons for their choice, with a one-minute break between the clips. Following the pre-training test, feedback was provided to the participants. Then, the main training components were delivered, and the post-training test occurred in the same manner as the pre-test and, again, was followed by feedback and discussion.

Results

Control study

Primary analysis

Performance of control study participants was evaluated for set 1 (equivalent to the 'pre-test' stimuli in the main study) and set 2 (equivalent to the 'post-test' stimuli in the main study) together and separately. Overall, it was possible to get a maximum score of 8 (i.e. classifying all eight videos correctly). As

expected, the average overall score was 4 out of 8 possible points, or $M = 50\%$ accuracy (95% CI = 3.45–4.55). The standard deviation and standard error for the set one scores were 1.32 and 0.26, respectively. For each set individually, it was possible to get a maximum score of 4. The average set one score was 1.6 out of 4 possible points, or $M = 40\%$ accuracy (95% CI = 1.26–1.94). The standard deviation and standard error for the set one scores were 0.82 and 0.16, respectively. The average set two score was 2.4/4, or $M = 60\%$ accuracy (95% CI = 2.00–2.80). The standard deviation and error for the set two scores were 0.96 and 0.19, respectively. The range for the set one was 3 while the set two range was 4, indicating that some individuals scored 0/4 on both sets. A paired samples *t*-test was conducted to compare the means of the two sets by computing difference scores for each case and evaluating whether the average difference is significantly different from zero. Results of the paired samples *t*-test indicate that the difference between set one and set two scores are significant at the .05 level (two-tailed $t = 3.36$; 95% CI: 0.31–1.29).

Signal detection analysis

Overall, the hit rate (true positives or cases in which deception is correctly identified, calculated by dividing the total number of hits by the total number of deceptive cases) and false alarm rate (false positives or cases in which truth-telling is incorrectly labeled as deception, calculated by dividing the total number of false alarms by the total number of truthful cases) were both 51%. This indicates that over all eight videos, participants did not differ from the level of chance in their ability to discriminate between truth-tellers and liars.

Further investigation into the nature of change for each set individually indicates that participants' hit rates increased from set one to set two, with a non-significant increase in accuracy from 44 to 58% (paired $t = 1.57$, $p = 0.13$; 95% CI: (−)0.32–0.044). Additionally, false alarm rates decreased from set one to set two. Participants significantly decreased from a false alarm rate of 64–38% ($t = 3.98$, $p < .05$; 95% CI: 0.13–0.40). It thus appears that participants' guilt-bias diminished for set two, explaining the improved performance compared to set one.

Individual differences

Participants also completed two Likert-scale questions on their ability to detect deceit and their ability to deceive others. On average, participants rated themselves as 4.0 out of a possible 7 (95% CI: 3.36–4.64; SD = 1.55), indicating modest confidence in their ability to detect deception. A similar but slightly elevated pattern was found for confidence in lie-telling abilities ($M = 4.32$; 95% CI: 3.59–5.05; SD = 1.77). Neither of these ratings was significantly correlated with deception detection accuracy.

Main study

Primary analysis

For the baseline pre-test, it was possible to get a maximum score of 4 (i.e. classifying all four speakers correctly). The average pre-seminar score was 1.86 out of 4 possible points, or $M = 46\%$ accuracy (95% CI = 1.48–2.24). The standard deviation and standard error for the pre-test scores were 1.22 and 0.19, respectively. The average post-seminar score was 3.24/4, or $M = 81\%$ accuracy (95% CI = 2.97–3.50). The standard deviation and error for the post-seminar scores were 0.85 and 0.13, respectively. The range for the pre-test was 4 while the post-test range was 3, indicating that some individuals scored 0/4 on the pre-test, but no one scored 0/4 on the post-test.

The null hypothesis was evaluated with a Wilcoxon signed-ranks test and a paired samples *t*-test. The Wilcoxon test compares the median of a single column of numbers against a hypothetical median. It is a non-parametric method of examining statistical significance in repeated-measures designs and is often preferred over parametric tests because it does not require assumptions about the form of the distribution of measurements (Corder & Foreman, 2009). Application of the Wilcoxon test to the data resulted in a significant difference ($z = 4.72$, $p < .001$) between pre- and post-test in the accurate classification of sincere and deceptive speakers. According to the analysis, 2 participants' scores decreased from pre- to post-test, 7 scores remained the same, and 33 increased.

The paired samples *t*-test compares the means of two variables by computing difference scores for each case and evaluating whether the average difference is significantly different from zero. As *t*-tests are more familiar to most readers, this test was performed to reinforce the results found by the Wilcoxon
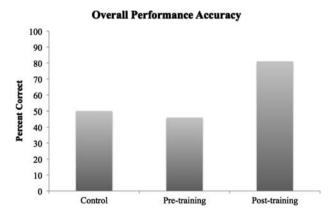

Figure 1.   This figure indicates that participants performed at the level of chance in overall deception detection accuracy in the control condition and pre-training, and were significantly better than chance post-training.

test. Results of the paired samples *t*-test indicate that the difference between pre- and post-test scores was significant at the .05 level (two-tailed $t = 6.85$; 95% CI: 0.97–1.79) (Figure 1).

Signal detection analysis

Further investigation into the nature of the improvement in deception detection ability indicated that participants' hit rates (cases in which deception is correctly identified) increased from pre- to post-test, with a significant increase in accuracy from 56 to 93% (paired $t = 5.78$, $p < .001$; 95% CI: 0.30–0.84). Additionally, false alarm rates (cases in which truth-telling is incorrectly labeled as deception) significantly decreased from pre- to post-test. Participants dropped from a false alarm rate of 57–29% ($t = 4.30$, $p < .001$; 95% CI: 0.48–1.00). These changes indicate that participants were significantly better at detecting deception post-training, and they were less likely to incorrectly attribute deception to truth-tellers (Figure 2).

Individual differences

Participants also completed two Likert-scale questions on their ability to detect deceit and their ability to deceive others. On average, participants rated themselves as 3.77 out of a possible 7 (95% CI: 3.41–4.14; SD = 1.06), indicating modest confidence in their ability to detect deception. An almost identical pattern was found for confidence in lie-telling abilities ($M = 3.84$; 95% CI: 3.40–4.29; SD = 1.28).

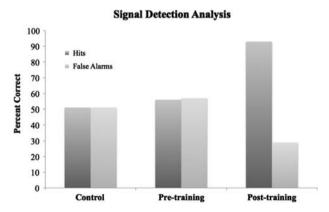Figure 2.   Signal detection analysis of participants' hit-rates (correct identification of deception) and false alarms (incorrect attribution of deception to truth-telling). This figure indicates that participants performed at the level of chance in discriminating between honesty and deception both in the control condition and pre-training, and were significantly better than chance post-training.

Difference scores (between the pre- and post-tests) were run through correlational analysis with the following variables that were collected in a demographics questionnaire, but were found to be non-significant: gender, age, years of professional experience, ethnicity, previous deception detection training, self-evaluation of ability to detect deception, self-evaluation of ability to deceive others, and profession (psychologist/non-psychologist).

Discussion

This study examined the effectiveness of a comprehensive deception detection training program targeting forensic psychiatrists, forensic psychologists, and various legal professionals. The results suggest that the training resulted in a substantial improvement in ability to detect extremely high-stakes deception, from chance to 81% accuracy. Further, participants improved in their ability to identify deception, while decreasing in attributions of deceptiveness towards sincere speakers. The availability of a group of forensic professionals and the use of high-stakes lies lends this research particular external validity, as these are the types of observers and targets who, respectively, assess and communicate high-stakes lies with great relevance in society. The training strategies also adhered to the 'state-of-the-art' guidelines surrounding empirically based deception detection offered by Porter and ten Brinke (2010) and Vrij, Granhag, et al. (2010) and the recommendations of Frank and Freely (2003) for implementing training. The results improve upon those of earlier (e.g. Porter et al., 2000) and briefer (Porter et al., 2010) programs, and are similar to recent findings by Vrij, Leal et al. (2010) who found that up to 80% of deceivers and truth-tellers could be correctly classified using a different approach (asking unanticipated questions) in a more active deception detection scenario.

While we are enthusiastic about the apparent success of this training, there are some limitations that should be considered in interpreting our results and their applicability. Most importantly, our sample sizes were relatively small (due to a lack of access). Although, based on the previous literature, we would expect the control group to approach the level of chance for all components and for the main study results to maintain their current pattern as sample size increases, it cannot be ruled out that a larger sample may show different results. Additionally, we were only able to use eight pleader videos, as time was a limiting factor for both the students and professionals we included in this study, and they would have likely declined participation had we made the study length longer. As with all published evaluations of deception detection training programs to date, we also cannot be sure that the effectiveness of training would be maintained over time or whether any skills gained would transfer into real professional lie-catching situations. The type of high-stakes lies used in the pre- and post-tests here were highly idiosyncratic 'pleader' lies, and we do not know whether the knowledge gained would generalize to other types of lies. On the other hand, a strength of this training was the use of

real-life, high-stakes lies, unlike the trivial lies that are typically studied in experimental deception detection research. Another consideration in interpreting these results is that the gains could be related to a day of heavy practice, rather than the information provided per se. Although this issue does require further research, it seems reasonable to assume that the training components provided including an integration of body language, facial expressions, and statement analysis incorporated in the current training helped participants make better decisions about truthfulness at post-test. But, we also cannot ascertain which elements of the training contributed most to the gains witnessed. Other issues that require investigation is whether other instructors using the same program would obtain similar results, whether other professional groups would benefit to this extent, and whether participants would improve at making in-person credibility assessments (since we used videotapes and a passive judgment task exclusively). It also is important to note that the current state of the literature on deception detection is far from complete, and as new relevant data are published they should be incorporated into deception detection workshops. These issues provide great potential for future research to advance our knowledge of improving deception detection training.

Conclusions

Based on the results obtained in this study, and similar promising results found previously, it appears that professionals can learn how to better discriminate between truthful speakers and liars relating extremely high-stakes lies. While these are promising results, it is important we continue to investigate how, why, and for whom training is most effective. Professionals in psychology and the legal system need to make decisions regarding the veracity of high-stakes statements on a regular basis, with errors potentially leading to tremendous consequences. We suggest that this warrants a call for further research to develop and evaluate empirical deception detection training for professionals. We cautiously suggest that training can be effective despite some skepticism in the field (e.g. Masip et al., 2009), although the long-term consequences of such training remain to be examined.

References

Aamodt, M., & Custer, H. (2006). Who can best catch a liar?: A meta-analysis of individual differences in detecting deception. *The Forensic Examiner, 15*, 6–11. Retrieved from PsycINFO database.

Blair, J., Levine, T., & Shaw, A. (2010). Content in context improves deception detection accuracy. *Human Communication Research, 36*, 423–442. doi: 10.1111/j.1468-2958.2010.01382.x

Bond, C., & Uysal, A. (2007). On lie detection 'Wizards'. *Law and Human Behavior, 31*, 109–115. doi: 10.1007/s10979-006-9016-1

Bond, C., & DePaulo, B. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*, 214–234. doi: 10.1207/s15327957pspr1003_2

Bond, C., & DePaulo, B. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin, 134*, 477–492. doi: 10.1037/0033-2909.134.4.477

Bond, G.D. (2008). Deception detection expertise. *Law And Human Behavior, 32*, 339–351. doi: 10.1007/s10979-007-9110-z

Bull, R. (2004). Training to detect deception from behavioural cues: Attempts and problems. The detection of deception in forensic contexts (pp. 251–268). New York, NY: Cambridge University Press. doi: 10.1017/CBO9780511490071.011

Corder, G., & Foreman, D. (2009). *Nonparametric statistics for non-statisticians: A step-by-step approach*. New York, NY: John Wiley.

DePaulo, B., & Kirkendol, S. (1989). The motivational impairment effect in the communication of deception. Credibility assessment (pp. 51–70). New York, NY: Kluwer Academic/Plenum.

DePaulo, B., Lanier, K., & Davis, T. (1983). Detecting the deceit of the motivated liar. *Journal of Personality and Social Psychology, 45*, 1096–1103. doi: 10.1037/0022-3514.45.5.1096

DePaulo, B., Lindsay, J., Malone, B., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*, 74–118. doi: 10.1037/0033-2909.129.1.74

deTurck, M., Roman, L., & Feeley, T. (1997). Vocal and Visual Cue Training in Behavioral Lie Detection. *Communication Research Reports*, *14*, 249–259. Retrieved from Communication & Mass Media Complete database.

Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist, 46*, 913–920. doi: 10.1037/0003-066X.46.9.913

Ekman, P., O'Sullivan, M., & Frank, M. (1999). A few can catch a liar. *Psychological Science, 10*, 263–266. doi: 10.1111/1467-9280.00147

Etcoff, N.L., Ekman, P., Magee, J.J., & Frank, M.G. (2000). Lie detection and language comprehension. *Nature, 405*. doi: 10.1038/35012129

Frank, M., & Feeley, T. (2003). To catch a liar: Challenges for research in lie detection training. *Journal of Applied Communication Research*, *31*, 58. Retrieved from Communication & Mass Media Complete database.

Gustrafson, L., & Orne, M. (1965). The effects of verbal responses on the laboratory detection of deception. *Psychophysiology, 2*, 10–13. doi: 10.1111/j.1469-8986.1965.tb02629.x

Hartwig, M., Granhag, P., Strömwall, L., & Vrij, A. (2004). Police officers' lie detection accuracy: Interrogating freely versus observing video. *Police Quarterly, 7*, 429–456. doi: 10.1177/1098611104264748

Kassin, S.M. (2006). A critical appraisal of modern police interrogations. In T. Williamson (Ed.), *Investigative interviewing: Rights, research, regulation* (pp. 207–228). Devon: Willan Publishing.

Kassin, S., & Fong, C. (1999). 'I'm innocent!': Effects of training on judgments of truth and deception in the interrogation room. *Law and Human Behavior, 23*, 499–516. doi: 10.1023/A:1022330011811

Köhnken, G. (1987). Training police officers to detect deceptive eyewitness statements: Does it work? *Social Behaviour*, *2*, 1–17. Retrieved from PsycINFO database.

Leach, A., Lindsay, R.L., Koehler, R., Beaudry, J.L., Bala, N.C., Lee, K., & Talwar, V. (2009). The reliability of lie detection performance. *Law And Human Behavior, 33*(1), 96–109. doi: 10.1007/s10979-008-9137-9

Mann, S., Vrij, A., & Bull, R. (2002). Suspects, lies, and videotape: An analysis of authentic high-stake liars. *Law and Human Behavior, 26*, 365–376. doi: 10.1023/A:1015332606792

Masip, J., Garrido, E., & Herrero, C. (2009). Heuristic versus systematic processing of information in detecting deception: Questioning the truth bias. *Psychological Reports, 105*, 11–36. Retrieved from PsycINFO database.

Meissner, C., & Kassin, S. (2002). 'He's guilty!': Investigator bias in judgments of truth and deception. *Law and Human Behavior, 26*, 469–480. doi: 10.1023/A:1020278620751

O'Sullivan, M. (2008). Home runs and humbugs: Comment on Bond and DePaulo (2008). *Psychological Bulletin, 134*, 493–497. doi: 10.1037/0033-2909.134.4.493.

O'Sullivan, M., Frank, M., Hurley, C., & Tiwana, J. (2009). Police lie detection accuracy: The effect of lie scenario. *Law and Human Behavior, 33*, 530–538. doi: 10.1007/s10979-008-9166-4

O'Sullivan, M., & Ekman, P. (2004). The wizards of deception detection. In P. Granhag & L. Strömwall (Eds.), *The detection of deception in forensic contexts* (pp. 269–286). New York, NY: Cambridge University Press. doi:10.1017/CBO9780511490071.012

Porter, S., Campbell, M., Stapleton, J., & Birt, A.R. (2002). The influence of judge, target, and stimulus characteristics on the accuracy of detecting deceit. *Canadian Journal of Behavioural Science, 34*, 172–185. doi: 10.1037/h0087170

Porter, S., Juodis, M., ten Brinke, L., Klein, R., & Wilson, K. (2010). Evaluation of the effectiveness of a brief deception detection training program. *Journal of Forensic Psychiatry & Psychology, 21*, 66–76. doi: 10.1080/14789940903174246

Porter, S., & ten Brinke, L. (2009). Dangerous decisions: A theoretical framework for understanding how judges assess credibility in the courtroom. *Legal and Criminological Psychology, 14*, 119–134. doi: 10.1348/135532508X281520

Porter, S., & ten Brinke, L. (2010). The truth about lies: What works in detecting high-stakes deception? *Legal and Criminological Psychology, 15*, 57–75. doi: 10.1348/135532509X433151

Porter, S., ten Brinke, L., & Wilson, K. (2009). Crime profiles and conditional release performance of psychopathic and non-psychopathic sexual offenders. *Legal and Criminological Psychology, 14*, 109–118. doi: 10.1348/135532508X284310

Porter, S., Woodworth, M., & Birt, A. (2000). Truth, lies, and videotape: An investigation of the ability of federal parole officers to detect deception. *Law and Human Behavior, 24*, 643–658. doi: 10.1023/A:1005500219657

R. v. Marquard (1993) 4 S.C.R. 223. Retrieved from http://canlii.ca/t/1frx2

Rogers, R., Salekin, R., Sewell, K., Goldstein, A., & Leonard, K. (1998). A comparison of forensic and nonforensic malingerers: A prototypical analysis of explanatory models. *Law and human behavior, 22*, 353–367. doi: 10.1023/A:1025714808591

Strömwall, L.A., Granhag, P., & Hartwig, M. (2004). Practitioners' beliefs about deception. In P. Granhag & L. Strömwall (Eds.), *The detection of deception in forensic contexts* (pp. 229–250). New York, NY: Cambridge University Press. doi: 10.1017/CBO9780511490071.010

Strömwall, L., & Granhag, P. (2003). How to detect deception? Arresting the beliefs of police officers, prosecutors and judges *Psychology, Crime & Law, 9*, 19–36. doi: 10.1080/10683160308138

Sullivan, K., Lange, R., & Dawes, S. (2007). Methods of detecting malingering and estimated symptom exaggeration base rates in Australia. *Journal of Forensic Neuropsychology, 4*, 49–70. doi: 10.1300/J151v04n04_04

The Global Deception Research Team. (2006). A world of lies. *Journal of Cross-Cultural Psychology, 37*, 60–74. doi: 10.1177/0022022105282295

Vrij, A. (1994). The impact of information and setting on detection of deception by police detectives. *Journal of Nonverbal Behavior, 18*, 117–136. doi: 10.1007/BF02170074

Vrij, A. (2005). Criteria-based content analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy, and Law, 11*, 3–41. doi: 10.1037/1076-8971.11.1.3

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities* (2nd ed.). Chichester: Wiley.

Vrij, A. (2008). Nonverbal dominance versus verbal accuracy in lie detection: A plea to change police practice. *Criminal Justice and Behavior, 35*, 1323–1336. doi: 10.1177/0093854808321530

Vrij, A., Akehurst, L., & Knight, S. (2006). Police officers', social workers', teachers' and the general public's beliefs about deception in children, adolescents and adults. *Legal and Criminological Psychology, 11*, 297–312. doi: 10.1348/135532505X608.16

Vrij, A., & Granhag, P. (2007). Interviewing to detect deception. In S.Å. Christianson (Ed.), *Offenders' memories of violent crimes* (pp. 279–304). New York, NY: John Wiley.

Vrij, A., Granhag, P.A., & Porter, S. (2010). Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychological Science in the Public Interest, 11*, 89–121. doi:10.1177/1529100610390861

Vrij, A., & Heaven, S. (1999). Vocal and verbal indicators of deception as a function of lie complexity. *Psychology, Crime & Law, 5*, 203–215. doi: 10.1080/10683169908401767

Vrij, A., Mann, S., Robbins, E., & Robinson, M. (2006). Police officers ability to detect deception in high stakes situations and in repeated lie detection tests. *Applied Cognitive Psychology, 20*, 741–755. doi: 10.1002/acp. 1200

Vrij, A., Leal, S., Mann, S., Warmelink, L., Granhag, P., & Fisher, R.P. (2010). Drawings as an innovative and successful lie detection tool. *Applied Cognitive Psychology, 24*, 587–594.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-Ms exposure to a face. *Psychological Science, 17*, 592–598. doi: 10.1111/j.1467-9280.2006.01750.x

Zuckerman, M., Koestner, R., & Colella, M. (1985). Learning to detect deception from three communication channels. *Journal of Nonverbal Behavior, 9*, 188–194. doi: 10.1007/BF01000739