

2011-01-01

Examination of Investigative Interviewing Techniques for Use in Secondary Screening for Malintent

Allyson Judith Horgan

University of Texas at El Paso, ajhboston42@hotmail.com

Follow this and additional works at: https://digitalcommons.utep.edu/open_etd



Part of the [Psychology Commons](#)

Recommended Citation

Horgan, Allyson Judith, "Examination of Investigative Interviewing Techniques for Use in Secondary Screening for Malintent" (2011).
Open Access Theses & Dissertations. 2506.

https://digitalcommons.utep.edu/open_etd/2506

This is brought to you for free and open access by DigitalCommons@UTEP. It has been accepted for inclusion in Open Access Theses & Dissertations by an authorized administrator of DigitalCommons@UTEP. For more information, please contact lweber@utep.edu.

EXAMINATION OF INVESTIGATIVE INTERVIEWING TECHNIQUES
FOR USE IN SECONDARY SCREENING FOR MALINTENT

ALLYSON J. HORGAN

Department of Psychology

APPROVED:

Christian A. Meissner, Ph.D., Chair

Daniel J. Martin, Ph.D., ABPP

Harmon M. Hosch, Ph.D.

Steve L. Crites, Jr., Ph.D.

Matthew H. Scullin, Ph.D.

Patricia D. Witherspoon, Ph.D.
Dean of the Graduate School

Copyright ©

By

Allyson J. Horgan

2011

Dedication

To my brother, Brett, who taught me to always keep smiling.

EXAMINATION OF INVESTIGATIVE INTERVIEWING TECHNIQUES

FOR USE IN SECONDARY SCREENING FOR MALINTENT

by

ALLYSON J. HORGAN, M.A.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

THE UNIVERSITY OF TEXAS AT EL PASO

MAY 2011

Author Note

This research was funded in part by the Science and Technology Directorate of the Department of Homeland Security.

Acknowledgements

There are many people that made this project possible. First, I would like to thank Dr. Chris Meissner for his support and guidance over the last five years. I would also like to thank Dr. Hosch, Dr. Crites, and Dr. Scullin for their insights into this project. Additionally, I would like to acknowledge the members of the IILab, especially Steve Michael and Dr. Jacki Evans for their assistance with data collection and coordination.

I would also like to express my appreciation for my colleagues at MRAC and the FAST team, especially those that helped with data collection. Special thanks to Drs. Dan and Jen Martin for their encouragement and guidance over the past year. This project would not have become a reality without their assistance. Also, the Science and Technology Directorate of the Department of Homeland Security for allowing me to complete this project, and especially Bob Burns, Bob Middleton, and Chuck Ryn for their continued support.

I would like to thank my family and friends for standing by me during this process. Thank you for reminding me that there is a light at the end of the tunnel. And finally, a special thanks to my mother for her wisdom, patience, and encouragement throughout this journey.

Abstract

There has been a renewed focus on the development of innovative screening procedures and technologies since the September 11th attacks. While attention has been placed on the need for more effective screening methods, there has been a lack of research on screening for future intent, which is remarkable given its importance to national security (Granhag, 2010; Martin, Martin, & Coskren, 2007). The current project draws on previous research on deception detection and investigative interviewing, as well as recent research on false intent in order to investigate the utility of several interviewing methods in screening for intent. Participants ($N = 144$) went through a security screening module. After passing through the primary screening process, participants entered secondary screening where they were interviewed using either an anxiety-based, cognitive-based, or direct questioning approach. Results indicated that interviewers and independent observers demonstrated a truth bias when making determinations about false intent, and that overall accuracy rates were no better than chance performance. Some support for the use of cognitive-based interview methods was obtained when reviewers used a coding tool based on empirically validated cues to deception.

Table of Contents

	Page
Author Note.....	v
Acknowledgments.....	vi
Abstract.....	vii
Table of Contents.....	viii
List of Tables.....	x
List of Figures.....	xi
Introduction.....	1
Deception Detection.....	3
Interview Techniques.....	5
Phase 1: Method.....	13
Phase 1: Results and Discussion.....	22
Participant Perceptions of Interview Context.....	22
Interviewer and Screening System Accuracy.....	23
Phase 2: Method.....	28
Phase2: Results and Discussion.....	30
Observer Video Ratings.....	30
PBCAT.....	31
Predicting Malintent across the Interview Context.....	34
General Discussion.....	36
Malintent Detection vs. Deception Detection.....	36
Effect of Interview Method.....	37
The Screening Process.....	38

Limitations and Future Research.....	39
Policy Implications.....	40
Conclusion.....	41
References.....	43
Appendices.....	59
Curriculum Vita.....	64

List of Tables

	Page
Table 1: Study Design and Conditions.....	47
Table 2: Purpose, Tone, and Posture of Each Interview Type.....	48
Table 3: Means and Standard Deviations of Malintent and Truthful Participant Ratings of Anxiety, Guilt, and Preparedness by Interview Method.....	49
Table 4: Percentage of Correct Classification Accuracy Throughout the Screening Process.....	50
Table 5: Percentage of Correct Classification by Interview Condition for Interviews, Observers, and the PBCAT.....	51
Table 6: Signal Detection Estimates Means and Standard Deviations by Interview Condition.....	52
Table 7: Means and Standard Deviations for PBCAT Cues by Interview Condition.....	53
Table 8: Hierarchical Logistic Regression Using Screening Decisions Points to Predict Malintent Classification Accuracy.....	54
Table 9: Hierarchical Logistic Regression Using Screening Decisions to Predict Malintent Classification Accuracy by Interview Method.....	55

List of Figures

	Page
Figure 1: Mean state STAI scores pre and post screening by interview condition.....	56
Figure 2: Malintent decision points throughout the screening process.....	57
Figure 3: PBCAT mean score decision accuracy rates by malintent status and interview condition	58

Examination of Investigative Interviewing Techniques for Use in Secondary Screening for Malintent

On December 25, 2009 Umar Farouk Abdulmutallab attempted to detonate an explosive device that was sewn into his underwear while on board a Detroit-bound flight (Nossiter, 2010). Although his attempt failed, the incident reminded US citizens of our nation's vulnerability to terrorist attacks. With the continued existence of threats to national security since the September 11th attacks, research initiatives have focused on developing screening procedures, methods, and technologies that might aid in the prevention of future attacks. For example, travelers may be asked to go through full-body scanners to ensure no contraband is hidden on their person. With the development of these screening procedures and technologies, corresponding questioning and interviewing methods need to be designed and tested in order to increase the effectiveness of the screening process.

A distinction can be made between two types of screening: primary screening and secondary screening. *Primary screening* generally involves a brief, initial interaction with the broad purpose of early recognition of suspicious individuals. Screeners could be looking for contraband, like weapons or illegal substances, or looking for suspicious behavior to identify potential human threats. For example, when re-entering the country travelers must go through a brief security screening with a Customs Agent during which they may be asked a series of questions about where they traveled and the purpose for travel. Using the above definition, these types of interactions would be considered primary screening. If a person's behavior or property is deemed suspicious during primary screening (or through random screening selections), the person may be sent for additional evaluation, which is referred to as secondary screening. *Secondary screening* may consist of a more extensive physical search of the person or their property and may include a more in-depth interview. In the Customs inspection example, a traveler may be sent to secondary screening if the primary screening agent deems the traveler's responses suspicious. During the secondary screening interview, the agent must decide whether or not

the traveler poses a security threat and whether the situation requires further action. In sum, while primary screening may establish a person's behavior or property as suspicious, secondary screening is vital in making determinations of whether or not further detention is warranted.

Regardless of whether screening is in the primary or secondary phase, the type of intelligence being sought may vary. The majority of research and technology development to date has focused on screening for contraband, suspicious behavior, and detecting past behavior. A perhaps more difficult task, and arguably more central to national security, is screening for *planned future events*. It has recently been noted that such research on true and false intent has been missing in the field of legal psychology (Granhag, 2010; Martin, Martin, & Coskren, 2007). Granhag (2010) has defined intent as the "mental state preceding a corresponding action." Given the significance of situations in which assessments of the veracity of one's intentions is warranted, it is surprising that research in this area is lacking. Therefore, researchers have called for the development of empirically based methods to detect future true and malicious intent (Granhag, 2010; Martin et al., 2007).

Theory of Malintent. As a first step in the development of empirically based methods for screening of intent, Martin, Martin, and Coskren (2007) developed the theory of malintent. The theory defines malintent as the "intention or desire to cause harm" and is thought of in the specific context of intending harm to the country's citizens or its infrastructure (Martin et al., 2007, p. 1). The theory draws from deception detection research to predict the behavioral and psychophysiological cues that are manifestations of the underlying emotional, behavioral, and cognitive processes a person with malintent may experience. The theory of malintent hypothesizes that identifying an individual with malintent becomes more likely as the magnitude and consequences of the event increase and as the distance to execution of the event decreases. It is predicted that deceptive individuals and those with malintent may experience similar emotional, behavioral, and cognitive processes and therefore display similar cues. While deception detection research focuses on the physiological and behavioral cues a person

exhibits while lying about a past event, malintent detection focuses on the cues a person exhibits regarding a planned future act. The current project seeks to develop and test methods for secondary screening for malicious intent by drawing on previous research on deception detection, intent detection, and investigative interviewing.

Deception Detection

As previously stated, research on future intent and cues to malicious intent are lacking (Granhag, 2010; Martin et al., 2007). However, looking to decades of research on deception detection has provided insight into the cues that may be indicative of false intent. Unfortunately, research has shown that both lay people and law enforcement professionals usually perform at near chance levels in deception detection tasks (Bond & DePaulo, 2006; Vrij, 2008). For law enforcement professionals, this poor performance could be due to a guilt bias and erroneous beliefs about which deception cues are actually indicative of deception (Kassin, Meissner, & Norwick, 2005; Meissner & Kassin, 2002). Nevertheless, improving individuals' ability to detect deception has proved to be a difficult undertaking.

Recent meta-analyses have examined the verbal (i.e., number of details, statement length), nonverbal (i.e., head nods, gestures), and paraverbal cues (i.e., pitch, vocal tension) to deception in order to determine those with the most empirical support. The DePaulo et al. (2003) meta-analysis examined 1338 estimates of 158 deception cues. Deception was significantly associated with several paraverbal and nonverbal cues including increases in pitch and vocal tension and decreases in the frequency of illustrators. Several verbal cues showed significant relationships with deception including decreases in verbal immediacy (clear and direct speech), the number of details revealed, logical structure of the account, overall talking time, spontaneous corrections, and an admitted lack of memory. Sporer and Schwandt (2006; 2007) have also conducted two meta-analyses on the nonverbal and paraverbal cues to deception. In their first meta-analysis of paraverbal cues, results from 41 studies were combined and analyzed. Results indicated that liars exhibited higher vocal pitch, longer response

latencies, more speech errors, and shorter messages (Sporer & Schwandt, 2006). In their second meta-analysis, results from 54 studies were combined and analyzed in order to examine moderating factors of nonverbal cues of deceit (Sporer & Schwandt, 2007). Results indicated that three behavioral cues were predictive of deception: decreased nodding, decreased leg and foot movements, and decreased hand movements. When participants were deceptive about facts versus their feelings, there was a decrease in eye contact, head nodding, and illustrators as well as an increase in head movements (Sporer & Schwandt, 2007). Taken together, these three meta-analyses reveal some consistency in the literature for cues such as pitch, length of account, and illustrators; however, there is also great variability in the types of cues displayed, and the effect sizes for all of the findings described above are in the small to medium range.

In addition to the variability in findings and small effect sizes, other criticisms of the deception detection literature exist. One of the most common criticisms of the research is that most of the studies to date have been conducted in laboratory settings using college students as participants and therefore may not be generalizable to real world scenarios (DePaulo et al., 2003; Hazlett, 2006). Also, participants are usually instructed to lie or tell the truth and may not be highly motivated to be perceived as truthful (DePaulo et al., 2003; Hazlett, 2006). Finally, an additional criticism of previous research is the lack of cross-cultural investigations of deception detection which would be particularly useful in intelligence gathering and screening contexts (Hazlett, 2006). In an attempt to further the literature and avoid some potential limitations, the current project will recruit from the general population to achieve a cross-cultural sample, provide participants the option of engaging in a malicious task (to ensure motivation and intent), and use a highly realistic security screening paradigm (for improved generalizability).

The preceding review of the deception detection literature shows that prior research has focused almost exclusively on cues to deception elicited when lying about a past event. As previously noted, the theory of malintent predicts that both deceptive individuals and those with malintent may

experience similar emotional, behavioral, and cognitive processes (Martin et al., 2007). Therefore, it is predicted that individuals with malintent may exhibit a similar pattern of physiological and behavioral cues when discussing future intent as deceptive individuals lying about past behavior.

Recently researchers have proposed new interviewing methods that may work to exploit the differences between liars and truth tellers by enhancing deception cues, rather than relying upon the passive observation of cues elicited through traditional methods (Vrij, Fisher, Mann, & Leal, 2010). Training interviewers to actively elicit cues using these strategic interviewing approaches may therefore improve the ability to detect deception (Vrij et al., 2010). The current study sought to further investigate this area of research by modifying these methods to fit the security screening context and examine the utility of these methods for malintent detection. Before discussing the methodology, I first review the relevant research on investigative interviewing methods currently used by law enforcement for confession elicitation, information gathering, and deception detection. I then examine some of the information gathering cognitive-based approaches in order to evaluate which methods may be applicable to screening for future intent.

Interview Techniques

There are two contrasting interview techniques often discussed in the literature: traditional (anxiety-based) approaches and information gathering (cognitive-based) approaches.

Traditional (anxiety-based) approaches. Traditional interviewing approaches involve confronting the individual with an accusation in an attempt to increase feelings of anxiety and pressure in the guilty suspect but not the innocent (Vrij, Mann, & Fisher, 2006). For example, the interviewer may say, "I know you committed the theft. Your reactions make me think you're trying to hide something." The underlying theory holds that liars will be more nervous and tense than truth tellers during the interview and this anxiety will manifest itself through verbal and nonverbal behavioral indicators of

deception (Vrij et al., 2010). This is the type of approach that law enforcement officers use on a regular basis.

There are several popular manuals and texts on interrogation including Walters' (1996) *Kinesic Interviewing*, Royal and Schutt's (1976) approach, and Reid and colleagues' (e.g., Inbau et al., 2004) technique (see Neuman & Salinas-Serrano, 2006). These techniques are generally based on the idea that deceit on the part of the suspect can be discovered through psychological manipulation. The hypothesis is that increasing levels of anxiety will force a person to confess in order to reduce anxiety, and that even before the confession a person may engage in physical activities to reduce this anxiety (Gudjonsson, 2003; Neuman & Salinas-Serrano, 2006). As a result of this increasing anxiety, suspects are believed to engage in various verbal and nonverbal behaviors that serve as either signals of discomfort or are the result of a suspect's attempts at reducing the anxiety – and these verbal and nonverbal behaviors therefore serve as potential cues to the veracity of the suspect's statements.

In order to increase the anxiety level of the suspect, the interviewer may engage in several behaviors such as asking guilt presumptive questions and shutting down denials of guilt. The interviewer may also use maximization techniques, which are confrontational in nature and use intimidation to persuade the suspect to confess (Kassin & McNall, 1991). For example, the interviewer may present false evidence of guilt, exaggerate the seriousness of the offense, or maintain a firm stance on the guilt of the suspect. The interviewer may also use minimization techniques, which are “soft sell” tactics that attempt to mislead the suspect into a false sense of security (Kassin & McNall, 1991). For example, the interviewer may justify the situation, sympathize with the suspect to gain trust, or reduce the perceived consequences of confessing to the offense.

An example of these types of techniques is the Behavior Analysis Interview (BAI), which is the first step of the Reid Technique. This preinterrogation interview is conducted to identify any deception on the part of the suspect through his or her verbal and nonverbal behavior. However, research

suggests that law enforcement officers, as well as the average person, perform only slightly better than chance when attempting to detect truth and deception (Bond & DePaulo, 2006; Vrij, 2008). Results from one study investigating the effectiveness of the BAI indicated that it actually produced a pattern in which false presumptions of guilt could be placed on innocent suspects, which is opposite of the intended purpose of the BAI (Vrij et al., 2006). Despite the research indicating people are poor lie detectors, the Reid Technique encourages investigators to interrogate suspects only after they have been deemed deceptive during the BAI (Neuman & Salinas-Serrano, 2006). The literature suggests that relying on the verbal and nonverbal behaviors of suspects during these types of interviews leads to inaccurate judgments of deception and these judgments can negatively impact decisions to detain suspects.

Vrij, Mann, and Fisher (2006) compared the experiences of truthful and deceptive participants who underwent an anxiety-based interview or an information gathering interview. Results indicated that participants in the anxiety-based accusatory condition felt more uncomfortable and exhibited fewer nonverbal cues to deception than those in the information gathering interview (Vrij et al., 2006). In another study comparing interview styles, the researchers found that accusatory interviews led to shorter denials and elicited fewer verbal cues to deception than information gathering interviews (Vrij, Mann, Kristen, & Fisher, 2007). Additionally, when the videotaped interviews were subsequently shown to police officers, those viewing anxiety-based interviews made more false accusations and were more confident in their incorrect veracity judgments (Vrij et al., 2007). Taken together, the results of these studies indicate that, despite their popularity in the law enforcement community, accusatorial, anxiety-based interviews may not be the most effective approach for eliciting information and detecting deception. In fact, it appears that these types of interviews can lead to increased confidence in incorrect decisions.

Information gathering (cognitive-based) approaches. While traditional anxiety-based interviewing approaches are based on the hypothesis that liars will exhibit more cues with increased fear, pressure, and anxiety, cognitive-based information gathering approaches are built on the idea that lying requires more cognitive resources than truth telling. Therefore, if the interview places additional cognitive demand on the liar, cues to deception are more likely to be elicited (Vrij, Fisher, Mann, & Leal, 2006). According to this theory, lying is more cognitively demanding than truth telling for several reasons. First, generating a lie while at the same time suppressing the truth is an intentional and deliberate act that requires more cognitive resources than telling the truth (Vrij et al., 2010). Furthermore, in addition to generating the lie, liars must monitor and manage their own behavior to avoid detection and observe the responses of the interviewer to determine if the interviewer believes the story (Vrij et al., 2010). Psychophysiological research has confirmed predictions that lying increases cognitive load. Using fMRI, researchers have shown that areas of the brain associated with higher order and executive functions are activated when someone is lying (Spence et al., 2004) and that this increased cognitive demand reduces tonic (basal or resting) arousal in liars in the form of decreased skin conductance and blink rate (Leal, Vrij, Fisher, & van Hooff, 2008).

Interviewers can exploit the fact that liars have fewer cognitive resources available by asking interviewees to complete additional cognitively demanding tasks during the interview. This is thought to result in increased displays of cognitive load in liars including a slower rate of speech with more pauses and hesitations, less overall movement, and fewer central details (Vrij et al., 2010). In short, it is hypothesized that these interventions will be more challenging for liars compared to truth tellers (Vrij et al., 2010). Strategic techniques suggested for increasing cognitive load include asking interviewees to tell their story in reverse order, instructing interviewees to maintain eye contact with the interviewer, and having interviewees complete additional (unexpected) tasks like drawing during the interview (Vrij et al., 2010).

Vrij and colleagues have begun to evaluate the utility of these cognitive load inducing interventions systematically (Vrij et al., 2008). In order to examine the effects of reverse order recall, participants were randomly assigned to either a truthful or deceptive condition. Those in the truthful condition were instructed to play a game with a confederate. During this time, another individual entered the room, wiped information off a blackboard, and left. Shortly after that interruption, another individual entered the room looking for a missing wallet. Liars did not participate in this staged event but were given a detailed synopsis. Participants were then asked to recall as much of the event as possible and half of the liars and truth tellers were also asked to recall their account in reverse order. Results indicated that those participants in the reverse order condition displayed more deceptive cues and signs of cognitive load than those in the control condition including more speech hesitations and errors as well as more leg and foot movements (Vrij et al., 2008). The researchers then showed the videotapes of the interviews to police officers and asked them to make veracity judgments. None of the police officers had received any training in deception detection. Results indicated that deception detection rates significantly improved and were above chance levels in the reverse order recall condition compared to the control condition (Vrij et al., 2008).

A similar procedure to the one described above was employed to test the effect of another technique for increasing cognitive load, namely maintaining eye contact with the interviewer. In this study, half of the participants were instructed to maintain eye contact with the interviewer when providing their accounts and half were given no instructions (Vrij et al., 2010). Compared to truth tellers, liars in the eye contact condition provided fewer temporal and spatial details, which are signs of both cognitive load and deception, and gave their accounts in sequential order. As in the previous study, independent raters viewed the videos of the interviews and provided determinations of statement veracity. Although there was no difference in discrimination accuracy between the control and eye contact conditions, raters were able to discriminate lies from truths when asked on a Likert-type scale

about truthfulness (Vrij et al., 2010). Another suggested manner of increasing the cognitive demand placed on an individual during an interview is to instruct participants to engage in a secondary, distracter task while being interviewed; however, to my knowledge, results related to this intervention have not yet been published (Vrij et al., 2010).

In addition to the cognitive load inducing interview techniques described above, other strategic interviewing methods have been proposed that would also mentally challenge liars compared to truth tellers. Asking unanticipated questions is one such method that may induce cognitive load and shows promise in discriminating liars from truth tellers (Vrij et al., 2009). To test this strategic technique, truthful participants went to lunch in pairs while deceptive participants were told to pretend they had done the same activity. Participants were then interviewed about their experience. They were asked spatial and temporal questions and were asked to draw the restaurant's layout, which was an unexpected task. Based on the responses to just these questions, participants were accurately classified by condition 80% of the time (Vrij et al., 2009). These results suggest that asking unexpected questions about topics central to the target event to unsuspecting liars can elicit cues to deceit and aid in deception detection (Vrij et al., 2009). Taken together, the results of these studies indicate that increasing cognitive load can actively elicit cues to deception, which may be more effective than the observation of cues previously used in conjunction with traditional anxiety-based methods (Evans, Meissner, & Brandon, 2010; Vrij et al., 2010).

The current research project compared the use of anxiety-based and cognitive-based interview approaches in secondary screening for intent in a highly realistic scenario. As reviewed above, previous research demonstrates that increasing the cognitive load of deceptive interviewees can elicit cues to deception. The current research modified the cognitive-based and anxiety-based techniques to screen for intent in an attempt to determine malintent status from the cues elicited from participants being interviewed regarding their future intent.

The Current Study

The current project was completed as part of the Future Attribute Screening Technologies (FAST) program funded by the Department of Homeland Security (DHS). Using the theory of malintent (Martin et al., 2007) as its foundation, the FAST program's objective is to increase the accuracy and validity of the identification of persons planning destructive acts through the development and testing of innovative technologies combined with theory-based, applied social and behavioral research. The FAST program is in its fourth year of applied research. The program recently shifted focus from screening for malintent at points of entry and large events to screening of employees for malintent at stadiums and arenas.

Purpose. The purpose of the current study was to develop and test operationally relevant interview methods that will maximize the amount of accurate information elicited and improve malintent discrimination within a limited period of time. The interview conditions examined were based on the anxiety-based and cognitive-based interviewing techniques previously described, as well as a direct questioning approach that served as a baseline comparison condition. The project was conducted as part of a larger screening study using a novel paradigm developed specifically for the FAST program. The current study received Institutional Review Board approval from the New England Institutional Review Board (NEIRB), which is an independent, accredited IRB located in the Boston area.

Interview method utility was evaluated using several dependent measures. First, the interviewers made determinations of malintent status in order to investigate accuracy as a function of interview method. These determinations of malintent were compared to the real time algorithm analysis of the primary screening system (which relies upon remote, noncontact sensors of physiological responses including pupil diameter, gross body movement, thermal signature, respiration, and cardiac activity) in order to determine if the secondary interview adds any additional diagnostic information to the entire screening process. Participants were asked to provide information about the nature of the

interview and their experiences during the interview including how much pressure, stress, and anxiety they felt during the screening process and interview in order to gain insight into the underlying emotional processes occurring during the interview. In addition, both trained and naïve independent raters reviewed the video recordings of the interviews to examine the verbal, nonverbal, and paraverbal indicators of malintent and make determinations of malintent status.

Hypotheses for the current study are as follows:

H1: Participants with malintent will feel more pressure, stress, and anxiety based on self report measures of pressure, stress, and anxiety, and scores on the State Trait Anxiety Inventory (STAI: Spielberger, Gorush, Luschene, Vagg, & Jacobs, 1983) during the screening process and interview, especially in the anxiety-based interview condition.

H2: The secondary screening determinations of malintent status given by the interviewers will provide diagnostic information to complement the primary screening system by increasing malintent detection accuracy, especially in the cognitive-based interview condition.

H3: The cognitive-based interview approach will increase malintent discrimination accuracy of the interviewers and independent raters by eliciting greater verbal, behavioral, and paraverbal cues of malintent as coded via the Psychologically Based Credibility Assessment Tool (PBCAT: Evans, Meissner, & Brandon, 2010) compared to either the direct questioning or anxiety-based interview approach

Phase 1

Method

Participants. Participants ($N = 144$) with experience in the food service industry were recruited through advertisements in Boston area media outlets. Based on previous studies (Russano, Meissner, Narchet, & Kassin, 2005; Lassiter, Greers, Handley, Weiland, & Munhall, 2002), a cell size of 30 was estimated to achieve an effect size of $d = .50$ with power = .80 at alpha = .05 level. Due to resource limitations and data collection issues, a cell size of 24 was achieved. Participants for this experiment were males (55.6%) and females (43.4%) with a mean age of 38.9 ($SD = 11.42$) and a high school education plus some additional education (years of education $M = 13.71$, $SD = 2.0$). At the time of recruitment, participants were told that they would be working as catering staff for a special event inside a large technology facility and would be required to go through security screening prior to entering the building. Participants were told they would be paid \$15 per hour of work. The experiment took place over two days at two locations and lasted approximately 3 hours, although participants were led to believe they would be working an additional half day at the event.

Design. Participants were randomly assigned to one of six experimental conditions based on a 2 (malintent vs. truthful) x 3 (interview method: direct vs. cognitive-based vs. anxiety-based) between-subjects factorial design (see Table 1). Participants in the malintent condition were randomly assigned to either the malintent without device condition or the malintent smuggle condition, which are described in more detail below. Although two types of malintent were examined as part of the larger experiment, these were collapsed for the purposes of this project.

Procedure. Individuals who responded to advertisements completed a brief telephone screening for initial eligibility with a researcher. This assessment identified potential participants as ineligible due to mental or physical conditions that would negatively impact participation or due to the

use of any medications or drugs that would negatively impact sensor data collection. Once deemed eligible, participants scheduled an appointment for the first research session.

Day One Briefing. During the first research session, participants met with a researcher to review the consent form and purpose of the study. Participants were told that they were being hired to work as catering staff for a special event with several VIPs in attendance. In order to enter the facility they would have to go through a security screening module that was being tested. In actuality, the study examined the utility of sensor technology in malintent detection in primary and secondary screening. Although participants were not told about the malintent condition during the informed consent interview, they were informed that some people may find portions of the study to be stressful. After signing the consent form, participants completed several assessments including a demographic information questionnaire and the State Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983) in order to obtain a baseline measure of anxiety. After completing these scales, participants were reminded of the next scheduled appointment which was on the day of the fictitious VIP event.

Day Two Briefing. At the beginning of the second research session, the researcher reviewed the consent form with the participant to ensure that the person continued to consent to participate. Participants were informed that they would be working as catering staff at a special event in a nearby technology building. In order to enter the facility, they would need to go through a new security screening system because of the VIPs in attendance at the event. Participants were told to expect two officers. The first would ask for personal identification and an official work schedule which was provided by the researcher. Inside the screening module, a second officer would ask 12 yes or no questions while multiple noncontact sensors measured their physiology. A third officer acted as the secondary screening interviewer. Participants were informed that they would only undergo secondary screening if they were flagged during primary screening; in actuality, all participants were sent to secondary screening.

Three officers from a private security agency filled the role of all three security officers. The officers were all male with a mean age of 25.67 ($SD = 2.52$) and at least a high school education (years of education $M = 13.0$, $SD = 1.0$). All three served in the military ($M = 7.0$ years, $SD = 1.73$) and had a mean of 4.33 years of security experience ($SD = 2.52$). All three officers reported they were fairly confident in their ability to detect malintent ($M = 2.67$, $SD = .58$ on a 5-point Likert scale) and were very motivated to detect malintent ($M = 4.0$, $SD = 1.73$ on a 5-point Likert scale).

Once participants agreed to work for the day and go through the experimental screening module, they were randomly assigned to either the truthful or malintent conditions. Participants assigned to the truthful condition were told to pass through the screening module, answer all the officer's questions truthfully, and continue to the employee entrance for work. Those participants randomly assigned to one of the malintent conditions were told that they had been selected to complete an additional task and could earn a \$100 bonus, but the task may be considered stressful or cause some additional anxiety. Participants then decided whether or not to complete the additional task. Those participants ($n = 12$ out of 144) that decided not to complete the additional task were debriefed, paid the full \$90, and escorted out of the facility.

After being assigned to a condition, participants were given an envelope and left alone in a testing room to review its contents. Participants received a CD to listen to the instructions as well as a written copy to review. The envelope also contained a copy of their work schedule and a red index card. The index card served as a way for participants to signal their desire to withdraw from the study at any time. After passing through screening, participants were told to enter the facility and receive further instructions about their job from the catering manager.

Malintent participants were given one of two sets of instructions, either for the malintent without device condition or the malintent smuggle condition. Both conditions involve working with a partner to complete a malicious task. Malintent without device participants were told to go through the

screening module, enter the facility, and meet their partner. Their partner would have an explosive device that would cause a disturbance at the event and the participant's job was to detonate the device. Other participants randomly assigned to the smuggling malintent condition were asked to smuggle a camera into the event, meet their partner inside, and take pictures of the VIPs illegally. Malintent participants were told that if they were caught by the officers, both they and their partner would lose any money they had already earned and would not be allowed to work at the event.

Participants in all conditions were told that if the officers found their identification or answers to any of the questions suspicious, they would be flagged for secondary screening. Secondary screening was described as an in depth interview that could last from 1 to 3 hours. The threat of further screening and detention, and the bonus for malintent participants, were intended to increase motivation to proceed through screening undetected. In reality, all participants went to secondary screening and it is this portion of the experiment that is the focal point of this project.

After reviewing the instruction packets, all participants were given about 10 min to plan their activities for the remainder of the day. In order to facilitate the planning phase, participants were given a fictitious map of the facility that included the location of the screening modules, employee entrance, kitchen, and event room. Participants were also given the daily menu, list of tasks to be completed, and the seating arrangement for the event room. Truthful participants were instructed to review directions to the kitchen where they were to report to the catering manager and to familiarize themselves with the menu and tasks. Malintent participants were told to plan out the execution of their primary task and also create a cover story in case they are stopped by a security officer.

Day Two Screening. After participants finished reviewing the instructions and planning their tasks, the researcher entered the testing room and introduced participants to their partner who was actually a confederate. The use of the confederate was intended to increase ecological validity, increase motivation to complete the task, and control flow of the experiment. Participants then walked to the

screening module with the confederate, engaging in a scripted interaction according to condition, and entered the line for the screening. The confederate attempted to have each participant enter the queue first. In some sessions, the participant wanted the confederate to go through the screening first. In these situations, the confederate was held in screening for five min in order to maintain realism and was then escorted out of screening surreptitiously.

The first officer asked for identification and the work schedule and escorted participants into the screening module. Once inside the screening module, the second officer asked 12 questions requiring a yes or no response based on the Utah Relevant Relevant Paradigm (RRP) developed by Kircher and colleagues at the University of Utah (J. Kircher, personal communication, July, 2008). This paradigm includes three question types: two types are relevant questions likely to trigger a behavioral or psychophysiological response, and the other type is neutral questions that serve as controls. One set of the relevant questions (R1) pertain to smuggling a recording device. For example, participants were asked, "Do you intend to use an illegal recording device inside the facility?" The second set of relevant questions (R2) relate to causing a disturbance at the event. For example, participants were asked, "Do you intend to cause a disturbance inside the facility?" Neutral questions inquire about the day of the week or current month. In comparing the two sets of relevant questions, participants with malintent should respond more strongly to the questions pertaining to the specific type of malintent they have (smuggling in a camera or detonating an explosive device). Each participant was asked 12 questions (3 R1 questions, 3 R2 questions, and 6 neutral questions). After each question there was a 17 second pause to allow the participants' physiology to return to baseline levels. At the end of this series of questions the officer informed participants that they would need to answer a few additional questions in an adjoining room.

In order to investigate the utility of the described interviewing approaches in malintent detection, all participants were escorted into the secondary screening room. Sending all participants for

further screening also allowed the researchers to examine participants' physiological reactions while being informed of the need for additional questioning. The secondary screening room was located inside the mobile screening module, adjoining the primary screening room, and consisted of a small room outfitted with audio and video surveillance equipment and two chairs. Once the participant was seated, a third officer entered the secondary screening room, introduced himself, and sat directly across from the participant. At this point, the officer began interviewing the participant based on one of three randomly assigned scripts: direct questioning, anxiety-based, or cognitive-based approach (see Appendix A).

Each interview began with a brief period of rapport building by manipulating similarity, liking, and empathy of the interviewer. This was accomplished through the interviewer smiling while introducing himself, shaking hands with the participant, and nodding early in the interaction. Rapport building has been shown to have a positive effect on interviews such that the interviewee provides more correct information without an observable increase in inaccurate information (Collins, Lynch, & Frank, 2002). This rapport building period was followed by several interview prompts or questions which varied depending on condition. Each interview lasted about 5 min and all interviewers were extensively trained in each interview approach (see Table 2).

The direct questioning interview served as a baseline comparison condition in which participants were asked basic questions about their intent after leaving the screening. For example, the interviewer asked questions like, "How did you find out about this position?" and "Do you have anything you'd like to tell me about your intentions today?" The interviewers' tone was neutral and posture was more formal.

The anxiety-based interview was modeled after techniques currently used in law enforcement interrogations and consisted of questions intended to increase the anxiety and pressure placed on participants. For instance, the interviewer said, "I don't believe that you are here to work today. Tell me

why you're really here." The interviewers' tone was confrontational and more aggressive than the direct questioning and cognitive load approaches.

The cognitive-based approach incorporated several of the interventions developed by Vrij and colleagues (Vrij et al., 2010) previously described, including reverse order recall and asking unexpected questions such as "How do you get to the kitchen from the screening module?" and "Ok, now can you tell me those directions in reverse order from the kitchen to the screening module?" The interviewer approached the interview in a more information gathering manner using a conversational tone, friendlier demeanor, and more relaxed posture.

Several steps measures were put in place in order to ensure that the officers were delivering the interview scripts consistently and accurately. Interviewers performed only one script each day and were informed the previous day which script would be performed the following day. This was intended to prevent script contamination and interference and allowed the interviewers to rehearse the day's script upon arrival to the mobile screening lab. The researcher observed a majority of secondary interviews and provided immediate feedback to the interviewer about tone and script adherence. Additionally, the interviewers completed a self assessment of script adherence and reported any inconsistencies and anomalies after each participant completed screening. Based on these measures, interviews that varied significantly from the scripts were removed from analyses.

Once the secondary interview was completed, participants were escorted out of the screening module and instructed to continue into the facility through the employee entrance. A researcher was waiting near the screening exit to meet the participant, escort them back to the research office, and debrief participants about the true nature of the study.

Day Two Debriefing. During debriefing, participants completed a series of dependent measures questionnaires. In addition to gathering information about the participants' impressions of primary screening, participants were asked about their impressions of the secondary screening interview.

Participants were asked about the pressure they felt during the interview to reveal information, the amount of stress they felt during the interview, and how anxious they were during the interview. These questions serve as dependent measures and also provide insight into how much participants “bought into” the scenario. Participants were also probed regarding any suspicions they had about the true purpose of the experiment as an additional measure of participant “buy in.” Once all questionnaires were completed, participants were paid, thanked for their time, and escorted out of the facility.

Several questions on the dependent measures questionnaire focused on anxiety and nervousness throughout the screening process including how anxious the participant felt walking to the screening, entering the screening, speaking with officer 1, officer 2, and officer 3. Other variables included how stressed, worried, nervous, and anxious participants felt during the interview. A factor analysis revealed that the factor loadings for these variables ranged from .622 to .878 ($\alpha = .913$), yielding a single measure of anxiety. In addition to the anxiety questions, several questions focused on the guilt participants experienced during screening. These included how guilty participants felt about completing their instructions, lying to the various officers, and lying during the interview. Also included in this factor was how immoral participants believed it was to complete their instructions and lie to the officers. Given the range of factor loadings for these variables (.749 to .910, $\alpha = .869$), these variables were combined into a single measure of guilt. Finally, two questions measured how prepared participants felt and if they believed they had enough time to prepare for their tasks. The factor loading for these measures was .830, again yielding a single measure of preparedness.

Because of the realism of the study, several safety precautions were followed to avoid adverse events. All research assistants were trained and prepared to deal with any unexpected adverse reactions to the study including security breaches and highly emotional participant responses. Additionally, before data collection began, there was a safety and security briefing for all research assistants, security officers, and system operators involved in the experiment.

In addition to the participant questionnaires, other dependent measures were collected. After each participant completed the entire screening process, the officers recorded whether or not they believed the participant had malintent and which specific cues they used to make their determination. The primary goal of the FAST screening technology is to make determinations of malintent based on features from noncontact sensors during primary screening. Participants' predicted malintent status based on sensor information from primary screening was compared with the secondary interviewers' determination of malintent status. This analysis determines whether or not a participant sent to secondary because of suspicion of malintent in primary screening would have subsequently been detected during the secondary interview. The results of this analysis serve as a measure of the diagnosticity and the operational utility of the interview techniques.

Results and Discussion

Participant Perceptions of Interview Context. To assess anxiety before and after screening, participants completed the STAI trait form during the day one briefing and the state form before and after screening on day two. An analysis of variance (ANOVA) was run to assess the effects of malintent status and interview condition on trait anxiety levels. As would be expected based on random assignment, results indicated no significant difference in trait anxiety levels across malintent conditions, $F(1, 138) = .97, ns., \eta_p^2 = .01$, and interview conditions, $F(2, 138) = .48, ns., \eta_p^2 < .01$, and no interaction effects, $F(2, 138) = .48, ns., \eta_p^2 < .01$.

To assess state anxiety before and after screening as a function of malintent status (truthful vs. malintent) and interview condition (direct vs. cognitive vs. anxiety), a $2 \times 2 \times 3$ repeated measures ANOVA was conducted. There was a main effect of time, $F(1, 138) = 56.88, p < .001, \eta_p^2 = .29$, 95% CI [.17, .40] such that state anxiety was higher after screening ($M = 32.34, SD = 10.58$) than before screening ($M = 26.67, SD = 6.68$). There was also a significant time by interview condition interaction, $F(2, 138) = 3.13, p < .05, \eta_p^2 = .04$, 95% CI [.00, .12]. As predicted, participants interviewed using the anxiety-based interview reported higher levels of anxiety post screening than participants in either of the other interview conditions (see Figure 1). While this difference was not found to be significant in post hoc analyses, a review of the effect sizes reveals that effect size was nearly double for the anxiety-based interview ($d = .81$) compared to the cognitive-based ($d = .46$) and direct ($d = .55$) interviews. Interestingly, there was no effect of malintent status on state anxiety, $F(1, 138) = 1.89, ns., \eta_p^2 = .01$.

In addition to the STAI, participants also completed self report measures of anxiety, guilt, and level of preparedness for the day's tasks experienced during the screening process. As previously discussed, related variables were combined into three latent variables: anxiety, guilt, and preparedness (see Table 3 for means and standard deviations of all three factors). A 2×3 ANOVA was conducted to investigate the effect of malintent status and interview condition for each variable.

For participants' self-reported anxiety, results indicated that, as predicted, malintent participants ($M = .18, SD = 1.03$) felt more anxiety throughout the screening process than did truthful participants ($M = -.18, SD = .94$), $F(1, 138) = 4.81, p < .05, \eta_p^2 = .03, 95\% CI [.00, .11]$. There was no effect of interview condition on anxiety, $F(2, 138) = .31, ns., \eta_p^2 = .01$, and no interaction effect, $F(2, 138) = .80, ns., \eta_p^2 = .01$.

For participants' feelings of guilt, a main effect of malintent status was observed, $F(1, 137) = 4.81, p < .05, \eta_p^2 = .03, 95\% CI [.00, .11]$. Malintent participants ($M = .68, SD = .94$) felt more guilt than truthful participants ($M = -.69, SD = .42$). There was no effect of interview condition on feelings of guilt, $F(2, 137) = .46, ns., \eta_p^2 = .01$, and no interaction effect, $F(2, 137) = .21, ns., \eta_p^2 < .01$.

Finally, analysis of participants' feelings of preparedness indicated that malintent participants ($M = -.29, SD = .94$) felt less prepared than truthful participants ($M = .29, SD = .97$), $F(1, 138) = 13.81, p < .001, \eta_p^2 = .09, 95\% CI [.02, .19]$. Additionally, a main effect of interview condition was also observed, $F(2, 138) = 4.33, p < .05, \eta_p^2 = .06, 95\% CI [.00, .14]$. Participants interviewed using the cognitive-based script ($M = -.26, SD = 1.11$) felt less prepared than those interviewed using the direct script ($M = .30, SD = 1.00$), $t(94) = 2.60, p < .05, d = -.53, 95\% CI [-.85, -.25]$. This finding confirms the manipulation that questions in the cognitive-based interview were unanticipated and participants felt they were unable to entirely prepare or rehearse their story for the interview. Feelings of preparedness were also lower for participants in the cognitive-based interview condition compared to the anxiety-based condition ($M = -.04, SD = .80$); however, this difference was not significant, $t(94) = 1.12, ns$. The interaction effect was not significant, $F(2, 138) = .50, ns., \eta_p^2 = .01$.

Interviewer and Screening System Accuracy. The screening process involved several steps with participants interacting with several security officers and the screening system (see Figure 2). Officer 1 greeted participants outside the screening module and escorted them inside where Officer 2 operated the screening system and asked the screening questions. The interviewer then met participants inside

the secondary screening room and further questioned participants based on one of the experimental interview scripts. In order to examine the accuracy of each step of the screening process, malintent decisions were collected from security officers after each participant completed screening. In addition, the screening system provided a classification for each participant based on the physiological sensors (see Table 4).

Officer 1. Officer 1 was the first person participants encountered when entering the screening system. The security officer would greet the participant, ask for their identification and work schedule, and then escort the participant into the primary screening system (see Figure 2). After this brief interaction, the officers were asked to make determinations of the participant's malintent status. Overall, Officer 1 accuracy was near chance (54.9%, $p > .05$) and there was evidence of a truth bias as officers were significantly more accurate in classifying truthful participants (84.5%) when compared with malintent participants (25.4%), $\chi^2 (1) = 50.18, p < .001$ (see Table 4).

Physiological measures. Once questioning in primary screening was complete, the screening system's real time decision algorithm fused data from the physiological sensors in order to make a determination of whether or not a participant needed further screening (see Figure 2). There was no significant difference in classification accuracies between truthful participants and malintent participants, $\chi^2 (1) = 1.20, ns.$ (see Table 4), and overall physiological sensor accuracy was 58.9%, which was greater than chance ($p < .05$). It should be noted that improvements to the real time decision algorithm based on previous data collection and development efforts are ongoing. For this study, upgrades were not implemented due to resource restrictions, but these upgrades will be implemented as the program moves towards a more operational context.

Officer 2. Officer 2 was standing behind a podium inside the mobile screening system. The security officer greeted participants, gave instructions for screening, calibrated the physiological sensors, asked the RRP questions, and escorted participants into secondary screening (see Figure 2).

After this phase of screening, which lasted about 7 min, Officer 2 made a determination of malintent status. Overall, Officer 2 accuracy was again no different than chance (47.2%, $p > .05$) and there was no difference between malintent and truthful classification, $\chi^2 (1) = 1.78$, *ns.* (see Table 4).

Interviewer. After participants had completed the primary screening, they were escorted into the secondary screening room. The interviewer entered the secondary screening room, introduced himself, and then began interviewing the participant based on one of the three interview methods evaluated. Each interview took less than 10 min. After each interview was complete, the interviewer determined whether or not the participant had malintent. A 3 x 2 hierarchical loglinear analysis was performed in order to assess whether interviewer accuracy varied as a function of participant condition and interview condition. Overall classification accuracy was at chance levels (48.3%, $p > .05$). Contrary to my hypothesis, neither the malintent status x interview method interaction, $\chi^2 (2) = 3.50$, *ns.*, nor the main effect of interview method, $\chi^2 (1) = 1.36$, *ns.*, were significant (see Table 5). There was a significant effect of malintent status, $\chi^2 (1) = 51.36$, $p < .001$, such that interviewer accuracy was higher for truthful participants (77.5%) compared to malintent participants (19.4%) (see Table 4). This suggests that interviews displayed a truth bias when making malintent determinations.

Malintent classification. In addition to the accuracy rates of each step in the screening process, we were interested in examining the accuracy rates of the interviewers for malintent participants that the screening system algorithm classified as having malintent. This analysis provides practical information about classification performance and diagnosticity in an operational setting. The question is: if a person has malintent and the screening system recommends further screening, would the interviewer detain the individual? Thirty-eight participants with malintent were correctly classified by the screening system algorithm based on physiological measures. After secondary screening, only seven (18.4%) of these participants were correctly identified by the interviewers as having malintent. Four out of twelve participants were correctly identified using the direct script, two out of fourteen using the

cognitive-based method, and one out of twelve using the anxiety-based method. After primary screening, the screening system correctly identified and recommended 54.3% of malintent participants for secondary screening. Of those sent to secondary screening, 18.4% were correctly identified by the interviewers as having malintent leading to an overall malintent detection rate of 10% with both the screening system and the interviewer correctly identifying seven participants out of 70 participants with malintent. The percentage of truthful participants falsely identified as having malintent was also examined. The primary screening system falsely identified 26 (36.6%) truthful participants as having malintent. In secondary screening, four of those participants were classified by the interviewer as having malintent leading to an overall false alarm rate of 6%.

Summary and Conclusions. Participants believing they were going to work as temporary catering staff went through a screening process in order to enter a mock event. Some of these participants were planning on disrupting the event (malintent) while others were just going to work for the day (truthful). Overall, the screening process made people more anxious, but malintent participants and those participants interviewed using the anxiety-based method felt even more anxious than other participants. Additionally, malintent participants and those interviewed using the cognitive-based script felt less prepared for their assigned tasks, suggesting that participants felt like the cognitive-based questions were unexpected and malintent participants felt they did not have enough time to rehearse their cover story.

Accuracy of the security officers, regardless of where they were positioned in the screening process, was at chance levels ($ps > .05$), which would be expected based on previous research on deception detection performance (Bond & DePaulo, 2006; Vrij, 2008). However, contrary to previous research demonstrating that law enforcement professionals usually display a guilt bias, the contracted security officers in this study actually displayed a truth bias. The security professionals in this study had little previous interviewing and deception detection experience, and this may explain the truth bias that

is sometimes seen in research conducted with non-law enforcement professionals as participants (Meissner & Kassin, 2002; 2004).

Phase 2

Once the screening phase of the experiment was complete, the videos from the secondary screening interviews were processed and two different types of reviews were completed. First, to evaluate independent observers' judgments of the interviews, a subset of the videos ($N = 60$) were reviewed by participants who made veracity determinations for each video. Next, in order to examine whether behavioral cues varied as a function of interview type or malintent status, two trained research assistants reviewed each video and rated them using the Psychologically Based Credibility Assessment Tool (PBCAT).

Method

Participants. Eighty-six participants ($N = 86$) were recruited from the University of Texas at El Paso through the web-based experimental signup software (Sona Systems). Participants were mostly female (62.8%) and Hispanic (72.1%) with a mean age of 21.01 ($SD = 5.20$) years old.

Design. A 2 (malintent vs. truthful) x 3 (cognitive-based vs. anxiety-based vs. direct) mixed factorial design was employed, with participants exposed to both malintent and truthful conditions and interview condition manipulated between subjects.

Procedure. Participants viewed a series of five malintent and five truthful videos from one of the interview conditions. Videos were randomly selected from the larger pool of videos collected in Phase 1 of the study, and presentation order of the videos was randomized across participants. Participants were told that the videos were of people being interviewed about their intentions for the rest of the day and that some were being honest while others were lying about their intentions. For each video, participants provided a veracity judgment. Additionally, participants rated how much pressure the interviewer placed on the participant, how much stress the participant seemed to be experiencing, and how hard the participant had to think about their responses.

PBCAT assessment of cues to deception. Two research assistants trained in using the PBCAT (Appendix B) rated the secondary screening videos. This tool is derived from empirically-based evidence on the cues that have been shown to be the most predictive of deception, and thus could be used to determine malintent status (Evans, Meissner, & Brandon, 2010). During the initial training, the research assistants thoroughly reviewed definitions and examples of each of the cues as well as how to complete and score the tool. Once the videos had been processed, the researchers reviewed 10 sample videos. After discussion about discrepancies, the researchers each reviewed 5 additional videos and compared results. This process was completed until the researchers were within one point on the scale for each of the cues. The researchers then reviewed 140 videos used in the previous analyses. Four videos were not available for analysis due to technical issues with audio or video quality.

The current version of the PBCAT consists of seven cues to truth and four cues to deception that have been empirically validated in previous research (see DePaulo et al., 2003). Initial research using the PBCAT reveals that diagnosticity is improved with interviews employing cognitive load inducing techniques (Evans et al., 2010).

Results and Discussion

Observer Video Ratings. Accuracy of participants' judgments were examined via a 2 (malintent vs. truthful) x 3 (direct vs. cognitive vs. anxiety) repeated measures ANOVA. Results revealed a main effect of veracity, $F(1, 83) = 223.77, p < .001, \eta_p^2 = .73, 95\% \text{ CI } [.63, .77]$, indicating that participants were more accurate on truthful trials (77.7%) when compared with malintent trials (19.1%). These findings are nearly identical to the interviewers and suggest that independent observers also display a truth bias when making veracity judgments of intent (see Table 4). Contrary to my prediction, there was no effect of interview condition, $F(2, 83) = 1.97, ns., \eta_p^2 = .05$, and no significant interaction effect, $F(2, 83) = .79, ns., \eta_p^2 = .02$. The observed effect sizes were small and there was insufficient power to detect the effects.

Veracity judgments were separated into the proportion of correctly identified malintent participants (hits) and the proportion of truthful videos incorrectly identified as malintent (false alarms) for each rater. These rates were then used to compute signal detection estimates of discrimination accuracy (A_z) and response bias (c). To examine whether discrimination varied as a function of interview method, a one-way ANOVA was conducted. Results indicated no significant differences in discrimination across interview methods, $F(2, 83) = .68, ns., \eta_p^2 = .02$. Another ANOVA was conducted in order to examine whether participants displayed a response bias as a function of interview method. Again, results indicated no significant differences in response bias across interview conditions, $F(2, 83) = 1.21, ns., \eta_p^2 = .03$ (see Table 6).

A series of 2 (malintent vs. truthful) by 3 (anxiety-based vs. cognitive-based vs. direct) ANOVAs were conducted to examine raters' impressions of pressure, stress, and level of thought. For pressure ratings, results indicated a main effect of interview condition, $F(2, 54) = 18.27, p < .001, \eta_p^2 = .40, 95\% \text{ CI } [.19, .54]$, such that participants believed interviewers placed more pressure on interviewees in the anxiety-based script ($M = 2.51, SD = .37$) than either the cognitive-based script ($M = 1.96, SD = .32$) or

the direct script ($M = 2.11, SD = .16$), $t_s(38) = 5.09$ and $4.53, p_s < .001, d_s = 1.63$ and 1.44 , 95% CIs [1.47, 1.63], [1.28, 1.51], respectively. A similar pattern was observed when participants rated how much stress the interviewee was under, $F(2, 54) = 5.39, p < .01, \eta_p^2 = .17$, 95% CI [.01, .32]. Participants interviewed using the anxiety-based method ($M = 2.56, SD = .55$) appeared to be under significantly more stress than those participants interviewed using either the cognitive-based method ($M = 2.15, SD = .43$) or the direct method ($M = 2.14, SD = .42$), $t_s(38) = 2.64$ and $2.67, p_s < .05, d_s = .85$ and $.88$, 95% CIs [.61, 1.04], [.64, 1.06], respectively. Raters did not observe any significant differences in how hard participants had to think about their responses as a function of malintent status, $F(1, 54) = .54, ns., \eta_p^2 = .01$ or interview method, $F(2, 54) = .75, ns., \eta_p^2 = .03$.

PBCAT. Two trained research assistants reviewed and coded the secondary interview videos using the PBCAT.

Inter-rater reliability. Once all the videos had been rated by both researchers using the PBCAT, inter-rater reliability was evaluated. This was done in two ways. First, correlations and kappas were calculated for continuous and categorical cues, respectively. For continuous cues, correlations ranged from .73 to .79 with the exception of one cue – namely, how hard the participant had to think ($r = .46$). For categorical cues, kappas ranged from .65 to .77. As an additional measure of reliability, rater agreement was defined as an individual cue score being within 1 point on the scale. For example, if one rater scored a cue as a 0.5 and the other rater scored the same cue as a 1.0, they were considered to be in agreement. However, if one rater scored a cue as a -0.5 and the other rater scored this same cue as a 1.0; this was not considered an agreement. When analyzed in this manner, percentage of agreement between the two raters across cues ranged from 95 – 100%. Based on these two measures of reliability, the scores from one rater were chosen randomly to be analyzed.

PBCAT cues. In order to assess the effects of participant condition and interview type on the 11 PBCAT cues, a 2 (malintent vs. truthful) x 3 (anxiety-based vs. cognitive-based vs. direct interview)

multivariate analysis of variance (MANOVA) was performed. In partial support of predictions, results indicated a significant effect of interview method, $F(22, 250) = 4.83, p < .001, \eta_p^2 = .30, 95\% \text{ CI } [.15, .33]$, but no effect of malintent condition, $F(11, 124) = .75, ns., \eta_p^2 = .06$, and no interaction, $F(22, 250) = 1.25, ns., \eta_p^2 = .10$. Sufficient power was observed to detect the interaction effect had it been present (observed power = .88). The Bonferroni method for controlling for type I error rate was used to evaluate each of the 11 cues using a revised p value of .005. Four cues were found to significantly differentiate between interview conditions. Participants in the cognitive-based interview provided significantly more spatial details than those in either the anxiety or direct interviews, $F(2, 135) = 34.12, p < .001, \eta_p^2 = .34, 95\% \text{ CI } [.21, .44]$ (see Table 7 for means and standard deviations). The cognitive-based interview yielded more details than either of the other two interview methods, $F(2, 135) = 8.19, p < .001, \eta_p^2 = .11, 95\% \text{ CI } [.02, .20]$. Participants in the cognitive condition were judged to have thought harder than those in the anxiety and direct conditions, $F(2, 135) = 9.07, p < .001, \eta_p^2 = .12, 95\% \text{ CI } [.03, .22]$. Finally, those in the cognitive condition also admitted not remembering more often than those in either the anxiety or direct interviews, $F(2, 135) = 13.61, p < .001, \eta_p^2 = .17, 95\% \text{ CI } [.06, .27]$.

In order to examine whether PBCAT cues differentiated malintent from truthful participants within each interview condition, MANOVAs were completed for each interview method. There were no significant differences between malintent and truthful for any of the interview types, (Anxiety $F(10, 36) = 1.12, ns., \eta_p^2 = .24$; Direct, $F(10, 37) = .91, ns., \eta_p^2 = .20$; Cognitive $F(11, 33) = 1.24, ns., \eta_p^2 = .29$).

PBCAT Accuracy. To further investigate the utility of the PBCAT in detecting malintent, a mean score of the 11 cues was calculated. Data was entered such that higher ratings on each individual cue were indicative of truth. A mean score was computed by summing the scores for all of the cues and dividing by the total number of cues (11). Since zero is the median point of each of the cues with negative ratings indicating deception and positive ratings indicating truth, a decision rule was created such that mean scores less than zero would be considered malintent and mean scores greater than zero

would be truthful. The default decision was truthful in cases where the mean was equal to zero in order to replicate real world conditions in which malintent would occur at a low base rate. From this decision, PBCAT accuracy was computed. A 2 (malintent vs. truthful) x 3 (anxiety-based vs. cognitive-based vs. direct) ANOVA was conducted to determine if accuracy varied as a function of either of these variables. Results indicated a significant malintent condition x interview method interaction, $F(2, 135) = 3.89, p < .05, \eta_p^2 = .05, 95\% \text{ CI } [.00, .13]$, such that PBCAT accuracy rates were highest when participants with malintent were interviewed with the cognitive-based script ($M = .71, SD = .46$) compared to any of the other interview methods (anxiety-based $M = .46, SD = .51$; direct $M = .46, SD = .51$). However, for truthful participants, PBCAT accuracy was lowest for those interviewed using the cognitive-based script ($M = .23, SD = .43$) compared to the other interview methods (anxiety-based $M = .48, SD = .51$; direct $M = .46, SD = .51$) (see Figure 3). These findings suggest that the cognitive-based interview facilitated malintent detection, but impeded correct classification of truthful participants.

Summary and Conclusions. After all of the interviews had been conducted, independent observers watched a subset of the videos. As would be expected based on previous literature, observers displayed a truth bias (Meissner & Kassin, 2002; 2004). There was no difference in observer accuracy across interview conditions. In addition to the observers, each video was rated using the PBCAT. More cues to deception were elicited in the cognitive-based script and PBCAT accuracy was highest for malintent participants interviewed using this method. Based on these results, it appears that the cognitive-based interview increases malintent detection when using the PBCAT, but fails to improve interviewer or observer accuracy. The observed effect sizes were small and as a result there was insufficient power to detect significant effects.

Predicting Malintent across the Interview Context

We were also interested in assessing the predictive validity of judgments of veracity made throughout the screening process and interview context. This includes the judgments made by each of the security officers that interacted with participants, the screening system judgments based on psychophysiological and behavioral measures of malintent, the interviewer judgments, and third party observers.

In order to examine the predictive validity of each step of the screening process, a hierarchical logistic regression was performed using the decision from the first officer that checked for identification and work schedule, the primary screening system's algorithm decision, the primary screening officer's decision to send the participant to secondary screening, the interviewers' decision of malintent status, and the computed PBCAT decision previously described entered chronologically (see Figure 2).

Malintent judgments from officer 1 did not significantly contribute to malintent prediction, $\chi^2(1) = 1.62$, *ns.*, and overall classification was 54.8% (see Table 8 for *b*, *Wald*, *ORs*, and classification accuracy for the overall model). The screening system decision was entered into the model next and significantly contributed to the model, $\chi^2(1) = 4.21$, $p < .05$, pseudo $R^2 = .06$, with classification accuracy reaching 59.4%. The officer 2 decision and the interviewer decision were entered separately and neither contributed significantly, $\chi^2(1) = .55$, *ns.*, and $\chi^2(1) = .02$, *ns.*, respectively. Finally, the PBCAT decisions were entered into the model and did not significantly contribute to malintent prediction, $\chi^2(1) = .69$, *ns.* Overall classification accuracy reached 58.0%.

This analysis was further broken down by interview approach to investigate the utility of each interview method. For the cognitive interview method, the interviewer's malintent decision contributed significantly to malintent prediction when entered in the model, $\chi^2(1) = 4.35$, $p < .05$, pseudo $R^2 = .19$ and overall yielded an accuracy rate of 68.1%; however, the interviewer decision was not a significant

predictor of malintent in the overall model (see Table 9). There were no significant predictors for the anxiety-based or direct interview methods (see Table 9).

Summary and Conclusions. The decision points throughout the screening process were examined in order to determine the predictive validity of each step. The system algorithm, which is based on the physiological sensors, was the only significant predictor of malintent in the overall model. Additionally, when the analysis was broken down by interview condition, the interviewer's decision was the only significant predictor in the cognitive-based interview condition. There were no significant predictors in either of the other interview conditions.

General Discussion

Interviewers and independent observers obtained modest malintent classification accuracy rates, regardless of the interview condition. In addition, both of these groups displayed a truth bias when making determination as to malintent status. And, as has been observed in the deception detection literature, malintent detection performance was no better than chance (Bond & DePaulo, 2006; Vrij, 2008). Overall classification accuracies for the individual decision points in the screening process ranged from 46.8% to 58.9% with hit rates for detecting malintent ranging between 19.1% and 54.3%. In both instances, the highest correct classification rates were from the screening system decision algorithm, which is based on noncontact physiological signals. The psychophysiological measures provided the most diagnostic information when attempting to classify malintent.

Malintent Detection vs. Deception Detection

The goal of the screening system evaluated in this study is to detect malintent. Since research on future intent and cues to malintent is lacking (Martin et al., 2007; Granhag, 2010), the theory of malintent draws from deception detection research to provide insight into the behavioral and psychophysiological cues that may be indicative of malintent. The results from the current study, as well as previous studies using the same screening system, demonstrate consistency between psychophysiological indicators of deception and malintent.

One objective of the current study was to investigate verbal and nonverbal malintent cues by examining whether or not cues indicative of deception were also indicative of malintent. The PBCAT is based on empirically validated cues to deception and has been found to be a useful tool in detecting deception (Evans, Meissner, & Brandon, 2010). When raters used the PBCAT to analyze secondary screening videos, certain cues (e.g., spatial details, the overall quantity of details, and admitted lack of memory) were elicited, especially in the cognitive-based script. Additionally, raters believed participants had to think harder about their responses in the cognitive-based interview. These results confirm that

the cognitive-based interview is performing as expected by eliciting more information and placing increased cognitive load on participants. However, the PBCAT cues did not differentiate malintent and truthful participants. While this finding needs to be investigated further and replicated, it may point to differences between verbal and nonverbal cues indicative of malintent and deception.

Previous deception detection research reveals that both professionals and laypeople perform at chance levels in deception detection tasks (Bond & DePaulo, 2006; Vrij, 2008). This same result was observed when both interviewers and independent observers completed malintent detection tasks. Both interviewers and observers obtained low accuracy rates and displayed a truth bias when determining malintent status. Observers may have had difficulty classifying participants due to the context of the interview. Interviewees discussed their intentions and plans for the day which were relatively mundane events that may not engender much suspicion on the part of the observers. Overall, there was little accusation and direct confrontation on the part of the interviewer, especially in the direct and cognitive-based interviews, which may have made this task particularly challenging for observers. Observing and classifying verbal and nonverbal behavior proves to be a difficult task, whether detecting deception or malintent.

Effect of Interview Method

It was predicted that accuracy of interviewers and independent observers would increase for participants interviewed using the cognitive-based method because the increased cognitive load placed on participants with malintent would elicit cues to deception. Results from the PBCAT ratings indicated that several cues to deception were elicited, especially in the cognitive-based interview. In addition, detection of malintent participants using the PBCAT was facilitated when the cognitive-based interview was used. However, truth detection was lowest for the cognitive-based interview condition compared to the anxiety-based and direct interview methods. These results support previous research demonstrating that cognitive load approaches to deception detection are more diagnostic when employed with liars

and demonstrate little gain when used with truth tellers (Evans, Meissner, & Brandon, 2010). Interestingly, when a decision criterion was created using the mean score of the PBCAT, a significant increase in malintent classification accuracy in the cognitive-based interview was observed. In addition, the predictive models of malintent by interview method revealed that the cognitive-based and direct interviewing methods explain more variance in the models and yielded higher accuracy rates than the anxiety-based method.

There was no difference in classification accuracy across any of the interview conditions for the independent observers. While raters using the PBCAT were asked to rate empirically validated cues to deception, independent observers and interviewers were given no instructions about nonverbal or verbal cues to deception or malintent. The difference between observer and interviewer accuracy compared to the PBCAT ratings could be a function of raters and interviewers relying on nondiagnostic cues when making malintent determinations. This finding also supports previous research demonstrating that people display erroneous beliefs about which cues are indicative of deception (Kassin, Meissner, & Norwick, 2005; Meissner & Kassin, 2002). Future work should further investigate the relationship between deception and malintent and work to empirically validate verbal and nonverbal malintent cues.

The Screening Process

The screening system utilizes signals from noncontact physiological sensors in order to make recommendations about secondary screening for malintent. While the security officer's determination factors into the system's algorithm and he or she can ultimately override the recommendation, the psychophysiological measures provided the most diagnostic information when attempting to classify malintent. When the malintent determinations of each step in the screening process were entered into the predictive model, overall accuracy reached 59.4%. This classification rate was identical when examining only primary screening, indicating that in its current state, secondary screening added no

diagnostic information to the screening system. Further, when examining classification accuracy of malintent participants that would have been sent to secondary screening, overall accuracy was only 10%. These accuracy rates are inadequate for screening in general and highlight the need for continued research and development.

It should be noted that this was the first investigation of malintent utilizing only noncontact sensors. Previous research using contact sensors has obtained higher classification results (i.e., classification rates have been as high as 81% accurate). Additionally, the algorithm that was used for this study to fuse the physiological data and maximize malintent detection was created using data from previous investigations and had not been updated for the current investigation. Based on these planned updates to the screening system and algorithm, it is expected that future classification accuracy will improve. In fact, subsequent analyses conducted by the research team on this data using the larger sample of participants and discriminant function analysis have achieved classification accuracy rates of 73%.

Limitations and Future Research

It has been noted that future research should continue investigating the relationship between malintent and deception and specifically work to determine and validate verbal and nonverbal indicators of malintent. Previous research reveals that strategic interview techniques that increase the cognitive load placed on interviewees can actively elicit cues to deception and improve classification accuracy (Vrij et al., 2010). While results from the PBCAT indicate that some cues were elicited in the cognitive-based script that utilized several of these techniques, accuracy rates for interviewers and observers were lower than expected.

The lack of significant findings could be due to several factors. The scenario employed in the current study varied considerably from other studies. Participants were informed that they would need to go through security screening in order to gain access to the facility where they would be working for

the day. In previous research, participants witnessed an event or engaged in another activity and were interviewed about their experience. Some participants told the truth during the interview while others lied. In the current study, participants were interviewed after going through primary screening, but before they completed either a malicious task or went to work for the day. Malintent participants were attempting to conceal their true intent during the interview compared to deceptive participants in previous studies that were attempting to conceal their earlier activities. In addition to differences in the context of the study and experimental scenario, this investigation incorporated a combination of strategic interview techniques compared to previous research that usually employed only one or two techniques. Future research should examine the utility of different combinations of interventions in both false intent and deception detection contexts.

It could be argued that the nonsignificant findings are the result of other factors like poor interview script adherence and weak experimental manipulations. However, interviewers were carefully observed throughout data collection and were given immediate feedback about script variations and anomalies. Results show that malintent participants and those participants interviewed using the anxiety-based interview reported increased anxiety levels indicating that the malintent conditions and anxiety-based interview manipulations were successful in increasing anxiety levels. Further, consistent with predictions and prior research, results from the PBCAT indicated that participants interviewed using the cognitive-based interview were rated as having to think harder about their responses and provided more information in the form of spatial and overall details. These findings provide evidence that the experimental interview conditions were successful.

Policy Implications

Deploying an experimental screening system to an operational setting is the ultimate goal of applied research programs. Before any system can be put into operation, decisions must be made about acceptable classification rates. The acceptable false positive and false negative rates can vary depending

on the operational context of the deployed system. Based on the results of this study, the secondary screening portion of this screening system is not ready for operational use. Future work will need to investigate and refine diagnostic secondary interview methods that can maximize information and intelligence elicitation. This initial attempt demonstrated that cognitive load inducing methods show promise, but other interviewing methods and techniques may further enhance future intent discrimination accuracy. Further, depending on the operational context, different limitations and restrictions may be placed on the interview methods employed. For instance, in some contexts, screeners may be able to interview each individual for an extended period of time. In other operational contexts, getting individuals through screening quickly may be heavily emphasized and interview methods would need to be modified to fit this requirement. The current program continues to investigate the utility of the screening system in various contexts and scenarios in order to increase the accuracy and validity of the identification of persons planning destructive acts. Various secondary screening methods will continue to be investigated as part of this research.

Conclusion

The continued existence of threats to national security requires research focusing on the development screening procedures, methods, and technologies that might aid in the prevention of future attacks. This is not an easy task. Nearly two million passengers travel through airports daily (www.tsa.gov) and close to one million passengers and pedestrians are screened daily at borders and points of entry (www.cbp.gov). These numbers fail to take into account other secure locations and venues, both public and private, where screening for potential security threats is necessary. The majority of people going through screening have no malintent, making this task even more difficult. The ultimate goal is to give screeners a tool to assist in the detection of the malintent. Continued research, creativity, and hard work will ultimately lead to the development of effective screening procedures,

interview methods, and technologies that will increase the accuracy and validity of the identification of persons planning destructive acts.

References

- Bond, C.F. Jr., & DePaulo, B.M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10* (3), 214-234. doi: 10.1007/s10979-006-9016-1
- Collins, R., Lincoln, R., & Frank, M. (2002). The effect of rapport in forensic interviewing. *Psychiatry, Psychology and Law, 9*(1), 69-78. doi: 10.1375/pplt.2002.9.1.69
- DePaulo, B., Lindsay, J., Malone, B., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*(1), 74-118. doi: 10.1037/0033-2909.129.1.74
- Evans, J.R., Meissner, C.A., & Brandon, S. (2010, April). *Development of a Psychologically Based Credibility Assessment Tool (PB-CAT)*. Poster presented at the IC Postdoctoral Research Fellowship Program Colloquium, Tyson's Corner, VA.
- Evans, J.R., Meissner, C.A., & Brandon, S. (2010). *Manipulating cognitive load can enhance the efficacy of deception cues with non-expert observers*. Manuscript in preparation.
- Granhag, P. (2010). On the psycho-legal study of true and false intentions: Dangerous waters and some stepping stones. *The Open Criminology Journal, 3*, 37-43. Retrieved from <http://www.benthamscience.com/open/tocrii/articles/V003/SI0009TOCRII/37TOCRII.pdf>
- Granhag, P., & Knieps, M. (2011). Episodic future thought: Illuminating the trademarks of forming true and false intentions. *Applied Cognitive Psychology, 25* (2), 274-280. doi: 10.1002/acp.1674
- Gudjonsson, G.H. (2003). *The psychology of interrogations and confessions: A handbook*. Chichester, England: John Wiley & Sons, Ltd. doi: 10.1002/acp.1031
- Hazlett, G. (2006). Research on detection of deception: What we know vs. what we think we know. In Intelligence Science Board, *Educating information: Interrogation: Science and art* (pp. 45-61) Washington, DC: National Defense Intelligence College Press.

- Kassin, S.M., & McNall, K. (1991). Police interrogations and confessions: Communicating promises and threats by pragmatic implication. *Law & Human Behavior*, *15*, 233-251.
doi: 10.1007/BF01061711
- Kassin, S.M., Meissner, C.A., & Norwick, R.J. (2005). "I'd know a false confession if I saw one": A comparative study of college students and police investigators. *Law and Human Behavior*, *29*, 211-228. doi: 10.1007/s10979-005-2416-9
- Lassiter, G.D., Greers, A., Handley, I., Weiland, P., & Munhall, P. (2002). Videotaped confessions and interrogations: A simple change in camera perspective alters verdicts in simulated trials. *Journal of Applied Psychology*, *87* (5), 867–874. doi: [10.1037/0021-9010.87.5.867](https://doi.org/10.1037/0021-9010.87.5.867)
- Leal, S., Vrij, A., Fisher, R., & van Hooff, H. (2008). The time of the crime: Cognitively induced tonic arousal suppression when lying in a free recall context. *Acta Psychologica*, *129*, 1-7. doi: 10.1016/j.actpsy.2008.03.015
- Martin, D.J., Martin, J., and Coskren, W.D. (2007). *A Theory of Malintent and Survey of Sensing Abilities for The Future Attribute Screening Technologies (FAST) Program*. (Available from the U.S. Department of Homeland Security, Science and Technology Directorate, Stop 0207, 245 Murray Lane, Washington, DC 20528-0207).
- Meissner, C.A. & Kassin, S.M. (2002). "He's guilty!": Investigator bias in judgments of truth and deception. *Law & Human Behavior*, *26*, 469-480. doi: 10.1023/A:1020278620751
- Meissner, C.A. & Kassin, S.M. (2004). "You're guilty, so just confess!" Cognitive and behavioral confirmation biases in the interrogation room. In D. Lassiter's (Ed.), *Interrogations confessions, and entrapment* (pp. 86-106). New York: Kluwer Academic/Plenum Press.

- Neuman, A. & Salinas-Serrano, D. (2006). Custodial interrogations: What we know, what we do, and what we can learn from law enforcement experiences. In Intelligence Science Board, *Educating information: Interrogation: Science and art* (pp. 141-233) Washington, DC: National Defense Intelligence College Press.
- Nossiter, A. (2010, January 17). The lonely trek to radicalism for terror suspect. *The New York Times*. Retrieved from <http://www.nytimes.com>
- On a typical day in fiscal year 2010. (2011). *U.S. Customs and Border Protection*. Retrieved April 20, 2011 from: http://www.cbp.gov/xp/cgov/about/accomplish/typical_day_fy2010.xml
- Russano, M. B., Meissner, C. M., Narchet, F. M., & Kassin, S. K. (2005). Investigating true and false confessions within a novel experimental paradigm. *Psychological Science, 16*, 481-486. doi: 10.1111/j.0956-7976.2005.01560.x
- Spielberger, C.D., Gorsuch, R.L., Luschene, R.E., Vagg, P.R., & Jacobs, G.A. (1983). *Manual for the State-Trait Anxiety Inventory form Y*. CA: Mind Garden Press.
- Spence, S., Hunter, M., Farrow, T., Green, R., Leung, D., Hughes, C., & Ganesan, V. (2004). A cognitive neurobiological account of deception: Evidence from functional neuroimaging. *Philosophical Transactions of the Royal Society of Biological Sciences, 359 (1451)*, 1755-1762. doi:10.1098/rstb.2004.1555
- Sporer, S., & Schwandt, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology, 28*, 421-446. doi:10.1002/acp.1190
- Sporer, S., & Schwandt, B. (2007). Moderators of nonverbal indicators of deception: A meta-analytic synthesis. *Psychology, Public Policy, and Law, 13(1)*, 1-34. doi:10.1037/1076-8971.13.1.1
- Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*. Chichester, England: John Wiley & Sons, Ltd.

- Vrij, A., Fisher, R., Mann, S., & Leal, S. (2006). Detecting deception by manipulating cognitive load. *TRENDS in Cognitive Sciences, 10*(4), 141-142. doi: 10.1016/j.tics.2006.02.003
- Vrij, A., Fisher, R., Mann, S., & Leal, S. (2010). Lie detection: Pitfalls and opportunities. In G. Lassiter & C. Meissner (Eds.), *Police interrogations and false confessions: Current research, practice, and policy recommendations* (pp. 97-110). Washington, DC: American Psychological Association.
- Vrij, A., Leal, S., Granhag, P.A., Mann, S., Fisher, R. P., Hillman, J., & Sperry, K. (2009). Outsmarting the liars: The benefit of asking unanticipated questions. *Law and Human Behavior, 33*, 159-166. doi: 10.1007/s10979-008-9143-y
- Vrij, A., & Mann, S. (2004). Detecting deception: The benefit of looking at a combination of behavioral, auditory and speech content related cues in a systematic manner. *Group Decision and Negotiation, 13*, 61-79. doi: 10.1023/B:GRUP.0000011946.74290.bc
- Vrij, A., Mann, S., & Fisher, R. (2006). Information gathering vs accusatory interview style: individual differences in respondents' experiences. *Personality and Individual Differences, 41*, 589-599. doi: 10.1016/j.paid.2006.02.014
- Vrij, A., Mann, S., Fisher, R., Leal, S., Milne, R., & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behavior, 32*, 253-265. doi: 10.1007/s10979-007-9103-y
- Vrij, A., Mann, S., Kristen, S., & Fisher, R. (2007). Cues to deception and ability to detect lies as a function of police interview styles. *Law and Human Behavior, 31*, 499-518. doi: 10.1007/s10979-006-9066-4
- What we do. (2011). *Transportation Security Administration*. Retrieved April 20, 2011 from: http://www.tsa.gov/what_we_do/screening/security_checkpoints.shtm

Table 1.

Study Design and Conditions.

Malintent Status	Interview Method		
	Direct	Anxiety-based	Cognitive-based
Malintent	n = 24	n = 24	n = 24
Truthful	n = 24	n = 24	n = 24

Table 2.

Purpose, Tone, and Posture of Each Interview Type.

Interview Type	Purpose	Tone	Posture
Anxiety-based	Increase anxiety level	Confrontational	Rigid
Cognitive-based	Increase cognitive demand	Conversational	Relaxed
Direct	Comparison, baseline	Neutral	Formal

Table 3.

Means and Standard Deviations of Malintent and Truthful Participant Ratings of Anxiety, Guilt, and Preparedness by Interview Method.

Interview	Malintent	Truthful
Direct		
Anxiety	.38 (1.24)	-.21 (1.15)
Guilt	.63 (.98)	-.67 (.53)
Prepared	-.07 (.92)	.67 (.95)
Cognitive		
Anxiety	.19 (.90)	-.23 (.79)
Guilt	.60 (.87)	-.74 (.34)
Prepared	-.44 (1.06)	-.08 (1.16)
Anxiety		
Anxiety	-.02 (.91)	-.11 (.89)
Guilt	.81 (.99)	-.67 (.38)
Prepared	-.36 (.85)	.28 (.62)

Table 4.

Percentage of Correct Classification Accuracy Throughout the Screening Process.

Decision Point	Malintent	Truthful	Overall
Officer 1	25.4%	84.5%	54.9%
System	54.3%	63.4%	58.9%
Officer 2	41.7%	52.8%	47.2%
Interviewer	19.4%	77.5%	48.3%
PBCAT	54.2%	39.1%	46.8%
Observers	19.1%	77.7%	48.4%

Table 5.

Percentage of Correct Classification by Interview Condition for Interviews, Observers, and the PBCAT.

Interview	Interviewers	Observers	PBCAT
Anxiety	50.0%	48.3%	46.8%
Cognitive	41.7%	51.4%	46.8%
Direct	53.6%	45.4%	45.8%

Table 6.

Signal Detection Estimates Means and Standard Deviations by Interview Condition.

Interview	Malintent Accuracy	Truth Accuracy	A_z	C
Anxiety	.21 (.18)	.76 (.24)	.49 (.27)	1.04 (.77)
Cognitive	.19 (.18)	.84 (.17)	.52 (.23)	1.30 (.79)
Direct	.18 (.19)	.73 (.30)	.43 (.22)	1.15 (.99)

Table 7.

Means and Standard Deviations for PBCAT Cues by Interview Condition.

Cues	Interview Method		
	Anxiety	Cognitive	Direct
Auditory Details	.00 (.00)	.04 (.30)	.00 (.00)
Spatial Details	.09 (.28)	1.04 (.77)	.33 (.60)
Temporal Details	.23 (.48)	.29 (.51)	.19 (.49)
Admitted Lack of Memory	.21 (.46)	.82 (.58)	.40 (.64)
Spontaneous Corrections	.15 (.36)	.20 (.46)	.10 (.68)
Overall Quantity of Details	-.06 (.85)	.48 (.75)	-.09 (.77)
Does the Story Make Sense	.09 (.73)	.32 (.74)	.17 (.69)
Thought Hard	.11 (.55)	-.37 (.63)	.10 (.68)
Nervous/Tense	-.18 (.60)	-.18 (.66)	-.33 (.67)
Negativity	-.06 (.54)	.01 (.62)	.15 (.42)
Rate of Speech	-.09 (.48)	-.30 (.53)	-.23 (.48)

Note. Bolded cues were found to significantly differentiate between the cognitive-based interview and the anxiety-based and direct interview methods.

Table 8.

Hierarchical Logistic Regression Using Screening Decisions Points to Predict Malintent Classification Accuracy.

	B	Wald	p	OR	pseudo R ²	95% CI	Accuracy
Officer 1	.60	1.85	.17	1.83	.02	[-.27, 1.48]	54.3
System	.75	4.41	.04	2.11	.06	[.04, 1.45]	59.4
Officer 2	-.24	.46	.50	.79	.06	[-.94, .46]	59.4
Interviewer	.01	.00	.98	1.01	.06	[-.87, .89]	59.4
PBCAT	-.30	.37	.41	.74	.07	[-1.03, .42]	58.0

*p-values bolded are significant at the .05 level

Table 9.

Hierarchical Logistic Regression Using Screening Decisions to Predict Malintent Classification Accuracy by Interview Method.

Interview	B	Wald	p	OR	pseudo R ²	Accuracy
Direct						
Officer 1	.93	1.55	.21	2.53	.06	59.1
System	.77	.66	.25	2.16	.11	65.9
Officer 2	-.78	.68	.25	.46	.15	70.5
Interviewer	.32	.78	.68	1.38	.16	68.2
PBCAT	.120	.68	.86	1.13	.16	68.2
Anxiety						
Officer 1	.88	1.12	.29	2.41	.02	55.3
System	.70	1.10	.29	2.02	.04	53.2
Officer 2	-.03	.00	.96	.97	.04	53.2
Interviewer	.40	.23	.63	1.49	.04	53.2
PBCAT	-.58	.79	.38	.56	.06	55.3
Cognitive						
Officer 1	-.63	.48	.49	.52	.00	51.1
System	1.25	3.53	.06	3.48	.08	61.7
Officer 2	.81	1.22	.27	2.42	.08	61.7
Interviewer	-1.85	3.18	.07	.16	.19	63.8
PBCAT	-.54	.64	.43	.58	.21	68.1

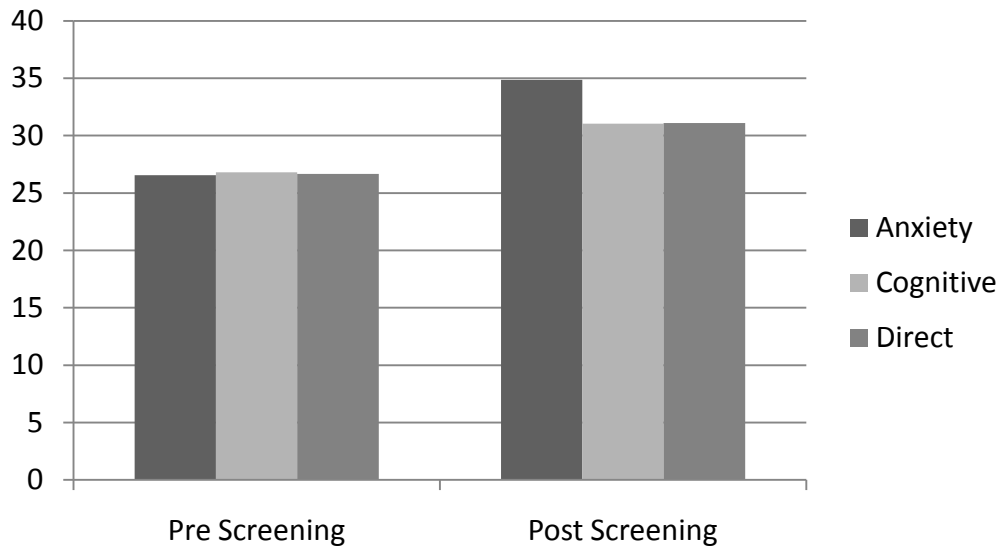


Figure 1. Mean state STAI scores pre and post screening by interview condition.



Figure 2. Malintent decision points throughout the screening process.

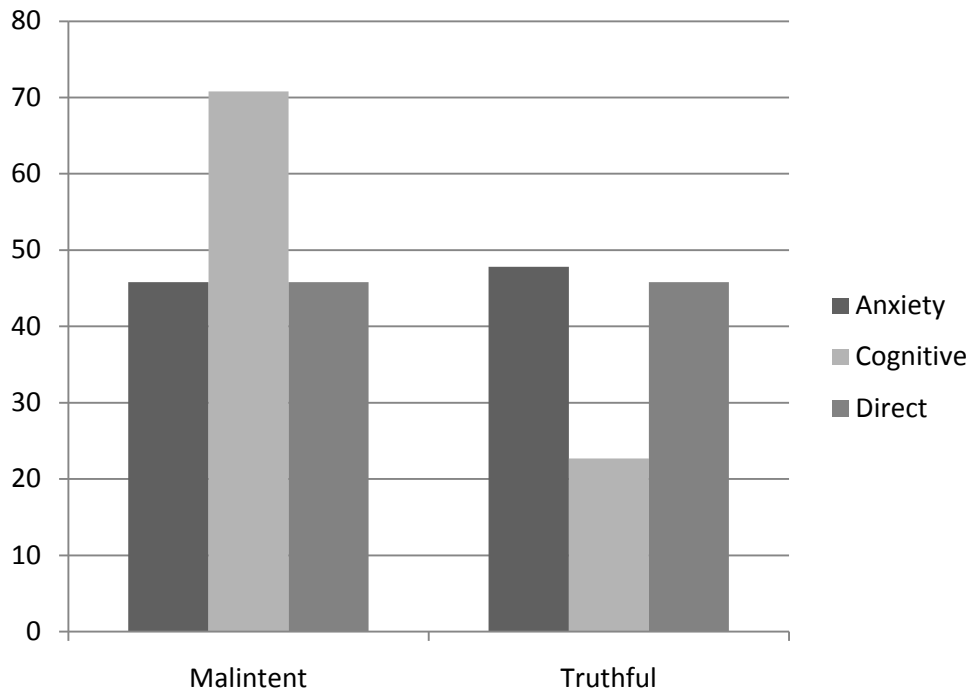


Figure 3. PBCAT mean score decision accuracy rates by malintent status and interview condition.

Appendix A

Secondary Screening Scripts

- The interviewer will wait in the Command and Control module until the participant is finished with Primary Screening.
- The interviewer will then proceed to the Screening module and enter through the back door.
- The interviewer will introduce himself, shake hands with the participant, and smile and nod during the introduction:

Hi, my name is _____, and I am another security officer at this facility.

What's your first name?

How are you doing today?

- The interviewer will then be seated across from the participant in the chair in front of the participant exit door.
- The interviewer will then begin questioning using one of the following scripts:

1. **Direct Questioning Approach (Script D)**

Tone and Posture

- Neutral and formal

Before I let you go into the facility today, I have a few more questions I need to ask you. Please answer all of my questions honestly and be as detailed as possible. This will help me get you into the building to work.

1. *So what are you doing at the facility today?*
 - a. *How did you find out about this position?*
2. *What kind of work do you think you'll specifically be doing today?*
 - a. *Ok, so once you get inside the building, where are you headed?*
3. *Are you planning to engage in any prohibited activities while in the facility today?*
 - a. *Do you understand what activities are prohibited inside the facility?*
 - b. **If No:** *Prohibited activities include carrying a weapon or other contraband, recording any information and smoking inside the facility. Do you plan on engaging in any of these activities?*
4. *Do you have any alternative intentions once you enter the facility you haven't told me about?*
 - a. *Is there anything else you'd like to tell me about your plans for today?*

Ok, thank you for answering all of my questions. You may now enter the facility.

2. Anxiety-based Approach (Script A)

Tone and Posture

- Confrontational and stern

Hi, my name is _____, and I am another security officer at this facility.

What's your first name?

How are you doing today?

Before I let you go into the facility today, I have a few more questions I need to ask you. Please answer all of my questions honestly and be as detailed as possible. This will help me get you into the building to work.

1. *Do you have any idea why we pulled you aside for additional questioning?*
 - a. *This system's never been wrong and it flagged you for additional questioning. Can you think of any reasons why you might have been flagged?*
2. *So, what are you doing at the facility today?*
 - a. *What is your assigned job for today?*
 - b. *I'm not convinced that you're here to work today and you could be in serious trouble if you're not being honest with me. Tell me why you're really here.*
3. *So, you work in the food service industry (insert specific occupation if stated)?*
 - a. *For how long?*
 - b. *Have you ever worked at this facility before?*
 - c. *We don't have many people come and work just for special events. Are you being honest about your plan for today?*
4. *Ok, (Participant's name), you seem like an honest person and I don't want to interfere with your job. If you're up front with me about your intentions at the facility, I can get you out of here quickly and you can get to work. So, where do you plan on going after you get done with screening?*
 - a. *So, if I watched once you left here, I would see you head towards (location participant stated)?*
 - b. *Is there anything else you want to tell me about your intentions here today?*

Ok, thank you for answering all of my questions. You can now go ahead into the facility.

3. Cognitive-based Approach (Script C)

Tone and Posture

- Conversational and relaxed

Hi, my name is _____, and I am another security officer at this facility.

What's your first name?

How are you doing today?

Before I let you go into the facility today, I have a few more questions I need to ask you. Please answer all of my questions honestly and be as detailed as possible. This will help me get you into the building to work.

1. So, why are you here at the facility today?
 - a. Ok, did you do any sort of planning or preparation for your position today?
 - b. As you were doing this planning, did you picture yourself doing anything or picture what the facilities looks like?
 - c. **If yes ask:** Can you please describe what you pictured?
 - i. Is there anything else you can remember about what you pictured?**If no**, continue to the next set of questions.
2. What is being served at the event today?
 - a. Is there anything else you remember from the menu?
3. Now, I'm assuming that since you are working in the building, you know a little about the facility. Can you tell me, in as much detail as possible, how do you get to the kitchen from where we are now?
 - a. Now can you tell me those directions but in reverse order, starting from the kitchen and ending up here?
 - b. Are there any other details you can remember about how to get to the kitchen?
4. Ok, I'd like you please take a minute and draw me a picture of the event room layout. Please be as detailed as possible.
 - a. Is there anything else you'd like to tell me about your plans today?

Ok, thank you for answering all of my questions. You can now go ahead into the facility.

Appendix B

PBCAT

Sensory Details

Not Present Occasional Frequent

Auditory Details

Spatial Details

Temporal Details

Not Present Present Present

(1 or 2 times) (3 or more times)

Admitted Lack of Memory

Spontaneous Corrections

Very Few Details/

Numerous Details/

Very Vague

Unsure

Complete

Overall quantity of Details / Talking time

|-----|-----|-----|-----|-----|-----|-----|-----|

Not at all

Unsure

Completely

Does it Make Sense (contradictions? plausible?)

|-----|-----|-----|-----|-----|-----|-----|-----|

Thought hard Did Not Think Hard Unsure Thought Extremely Hard
|-----|-----|-----|-----|-----|-----|-----|-----|

Extremely Relaxed/ Extremely Tense/
Comfortable Unsure Nervous
Nervousness (tense? fidgeting?)
|-----|-----|-----|-----|-----|-----|-----|-----|

Extremely Positive Neutral/Unsure Extremely Negative
Negativity/Complaints
|-----|-----|-----|-----|-----|-----|-----|-----|

Extremely Fast Unsure Extremely Slow
Rate of Speech
|-----|-----|-----|-----|-----|-----|-----|-----|

Judgment (T/F):

Curriculum Vita

Allyson Judith Horgan was born in Cambridge, Massachusetts and raised in Kingman, Arizona. She began school at Grand Canyon University in Phoenix, Arizona and transferred to Arizona State University's West Campus where she earned her Bachelor of Science degree in Psychology with a minor in Criminal Justice in 2006. While at Arizona State, Allyson worked as a research assistant in the Legal Psychology Research Lab under the direction of Dawn McQuiston, Ph.D. After graduating from Arizona State, Allyson entered the Graduate School at the University of Texas at El Paso to pursue a Ph.D. in Legal Psychology under the direction of Christian A. Meissner, Ph.D.

Allyson has presented her research at several national conferences including the American Psychology-Law Society Conference and the Society for Applied Research in Memory and Cognition Conference.

In the summer of 2009, Allyson completed an internship at Martin Research and Consulting (MRAC) in Cambridge, Massachusetts. She returned to MRAC in January 2010 to complete her dissertation in conjunction with the Future Attribute Screening Technologies Program funded by the Department of Homeland Security. Allyson is currently employed with MRAC as a Research Scientist.

Permanent Address: 2806 Colorado Ave
Kingman, AZ 86401