Journal of Strategic Security

Volume 6
Number 3 *Volume 6, No. 3: Fall 2013*

Article 9

Efficacy of Modified Cognitive Interviewing, Compared to Human Judgments in Detecting Deception Related to Bio-threat Activities

Charles A. Morgan Dr.
M-3 Forensic Consulting, camorgan3rd@gmail.com

Yaron G. Rabinowitz Dr.
Center for Research and Development, rubes0509@gmail.com

Deborah Hilts
Center for Research and Development, hilts@centerrd.com

Craig E. Weller
Center for Research and Development, crg.weller@gmail.com

Vladimir Coric Dr.
Center for Research and Development, coric@centerrd.com

pp. 100-119

Recommended Citation
Morgan, Charles A. Dr.; Rabinowitz, Yaron G. Dr.; Hilts, Deborah; Weller, Craig E.; and Coric, Vladimir Dr.. "Efficacy of Modified Cognitive Interviewing, Compared to Human Judgments in Detecting Deception Related to Bio-threat Activities." *Journal of Strategic Security* 6, no. 3 (2013) : 100-119.
DOI: http://dx.doi.org/10.5038/1944-0472.6.3.9
Available at: https://digitalcommons.usf.edu/jss/vol6/iss3/9

Efficacy of Modified Cognitive Interviewing, Compared to Human Judgments in Detecting Deception Related to Bio-threat Activities

Abstract

National security professionals have few scientifically valid methods for detecting deception in people who deny being involved in illicit activities relevant to national security. Numerous detecting deception studies have demonstrated that the Modified Cognitive Interviewing (MCI) method is one such method - yielding detecting deception rates (i.e. 80-85%) that are significantly above those achieved by chance (i.e. 50%) or by human judgments (i.e. 54-56%). To date, however, no MCI studies have involved dilemmas of ethological interest to national security professionals. This project begins to address this gap in the scientific literature. In it, we compared the efficacy of MCI to that of human judgments for detecting deception in scientists with expertise in biological materials. Sixty-four scientists were recruited for study; 12 met with a “terrorist” and were paid to make biological materials for illicit purposes. All 64 scientists were interviewed by investigators with law enforcement experience about the bio-threat issue. MCI elicited speech content differences in deceptive, compared to truthful scientists. This resulted in a classification accuracy of 84.4%; Accuracies for Human Judgments (interviewers/raters) were 54% and 46%, respectively. MCI required little time and its efficacy suggests it is reasonable to recommend its use to national security experts.

Introduction

Government officials in many countries are tasked with national security issues and must often rely on subjective judgments in order to determine whether a person being questioned is being truthful or deceptive about what they know or what they have done. However confident professionals may feel about their judgments, current scientific evidence demonstrates that the level of accuracy of human judgments about lying are only at, or slightly above, levels one might achieve by chance.[1] This relative inability to detect deception is true for a range of professional groups: police officers, judges, psychiatrists, university students, and agents from government law enforcement agencies.[2] Thus, it seems reasonable to explore whether alternate approaches might result in professionals being able to make more accurate assessments.

The majority of scientific studies on detecting deception (to include studies on devices such as the polygraph) are based on a model in which there is a presumption that liars are *more afraid* than truth tellers and as a result, will show signs of *increased physiological activity* (i.e., increased signs of autonomic arousal). In spite of this widespread – and popular – assumption, meta-analyses of studies based on this model[3] provide evidence that the *increased arousal* hypothesis yields detecting deception rates that are only modestly above rates expected by chance (i.e. 52-62%).[4]

By contrast, data acquired over the past decade from studies based on the *cognitive load* model (i.e., a model in which the telling of a lie is posited to require *more* mental work than the act of telling the truth) show that it is reliably more useful and may yield detecting deception rates far above those of chance or traditional approaches (i.e., 82-92%).[5] These data suggest that

[1] C. F. Bond Jr, K. N. Kahler, and L. M. Paolicelli, "The Miscommunication of Deception: An Adaptive Perspective," *Journal of Experimental Social Psychology* 21 (1985): 331-45; G. Hazlett and C. A. Morgan III, "Efficacy of Two Deception Detection Strategies When Assessing Individuals within Cross-Cultural Circumstances: Scientific Technical Report," (2011); C. A. Morgan, III, K. Colwell, and G. Hazlett, "Efficacy of Forensic Statement Analysis in Distinguishing Truthful from Deceptive Eyewitness Accounts of Highly Stressful Events," *Journal of Forensic Science* 56:5 (2011): 1227-34; A. Vrij and L. Akehurst, "Verbal Communication and Credibility: Statement Validity Assessment," in *Psychology and Law: Truthfulness, Accuracy and Credibility*, ed. A. Memon, A. Vrij, and R. Bull (Maidenhead, Great Britain: McGraw-Hill, 1998), 3-31.

[2] B. M. DePaulo and R.L. Pfeifer, "On-the-Job Experience and Skill at Detecting Deception," *Journal of Applied Social Psychology* 16:3 (1986): 249-67. P. Ekman and M. O'Sullivan, "Who Can Catch a Liar?," *American Psychologist* 46:9 (1991): 913-20; Vrij and Akehurst, "Verbal Communication and Credibility: Statement Validity Assessment," 3-31.

[3] Studies in which researchers have used the polygraph, the PCASS, lasar dopler technology, voice stress analysis devices. All of these methods are based on the idea that the act of lying will produce alterations in the sympathetic nervous system that can be detected through changes in heart rate, blood pressure, respiration rate, skin conductance, or voice quality.

[4] MI-5 Center for the Protection of National Infrastructure (CPNI), "Government Report: Detecting Deception: Guidance on Tools and Techniques," (May 2009). B. M. DePaulo et al., "Cues to Deception," *Psychological bulletin* 129:1 (2003): 74-118. J. D. Harnsberger et al., "Stress and Deception in Speech: Evaluating Layered Voice Analysis," *Journal of Forensic Science* 54:3 (2009): 642-50. Hazlett and Morgan, "Efficacy of Two Deception Detection Strategies When Assessing Individuals within Cross-Cultural Circumstances: Scientific Technical Report." Harry Hollien et al., "Evaluation of the Nitv Cvsa," *Journal of Forensic Sciences* 53:1 (2008): 183-93.

[5] K. Colwell et al., "Vividness and Spontaneity of Statement Detail Charateristics as Predictors of Witness Credibility," *American Journal of Forensic Psychology* 25:1 (2007): 5-30; C. A. Morgan, III et al., "Efficacy of Verbal and Global Judgment Cues in the Detection of Deception in Moroccans Interviewed Via an Interpreter," *Journal of Intelligence Community Research and Development* (2008a); C. A. Morgan, III et al., "Detecting

credibility assessment methods based on the *cognitive load* model of deception may offer a way to improve current credibility assessment procedures.

Over the past decade, numerous detecting deception studies have confirmed that Modified Cognitive Interviewing (MCI)[6], is: a) effective for detecting lies of both omission and fabrication; b) valid when used in cross-cultural settings[7]; and c) consistently associated with classification accuracies[8] at or above a level of 80 percent – a rate that is significantly above those achieved by professional judgments, polygraphy, or non-verbal behavior analysis.[9]

The design of the present study was developed in response to meetings with the grant sponsor designed to include a scenario that was ethologically valid for the sponsor. As a result, we created a scenario that would let us assess how well the MCI would work when assessing a group of interest to the government (i.e., scientists) an issue of interest to the government (i.e.,

---

Deception through Automated Analysis of Translated Speech: Credibility Assessments of Arabic-Speaking Interviewees," ibid.(2008b); C. A. Morgan, III et al., "Detecting Deception in Vietnamese: Efficacy of Forensic Statement Analysis When Interviewing Via an Interpreter," ibid.(2009b); Morgan, Colwell, and Hazlett, "Efficacy of Forensic Statement Analysis in Distinguishing Truthful from Deceptive Eyewitness Accounts of Highly Stressful Events," 1227-34; A. Vrij et al., "Detecting Deception by Manipulating Cognitive Load," *Trends in Cognitive Sciences* 10:4 (2006); A. Vrij et al., "Increasing Cognitive Load to Facilitate Lie Detection: The Benefit of Recalling an Event in Reverse Order," *Law and Human Behavior* 32:3 (2008).

[6] Although both traditional and modified versions include a detailed recall prompt to initiate the interview, Traditional and formal cognitive interviewing involves both a larger number of questions and prompts compared to the modified version as well as a more rigid adherence to sequence and phrasing of the prompts. For example, whereas traditional versions of the method include up to 10 memory prompts (e.g. for visual, auditory, olfactory, tactile sensory, gustatory, thoughts, emotions, alternative perspective, temporal reversal of account and a second full recall of story), Modified Cognitive Interviewing uses 4 prompts (e.g. visual, auditory, personal feelings, temporal reversal). As a result MCI is shorter and places less of a demand on the interviewee to keep repeating their story as many times. In addition, in MCI the phrasing of the prompts is less formal. For example, instead of the more formal prompt "I'd like to you think about your story and starting at the beginning tell me everything you saw during this time," the visual prompt in the MCI may consist of "What did that look like?" or "So if I had been with you during this time, what would I have seen?" or "To help me get a picture in my mind of this event, tell me what I would have seen" or "If this had been on the news, what would the TV camera show us?"

[7] K. Colwell, C. K. Hiscock, and A. Memon, "Interviewing Techniques and the Assessment of Statement Credibility," *Applied Cognitive Psychology* 16:3 (2002): 287-300; Colwell et al., "Vividness and Spontaneity of Statement Detail Charateristics as Predictors of Witness Credibility," 5-30; A. Memon et al., "Distinguishing Truthful from Invented Accounts Using Reality Monitoring Criteria," *Legal and Criminological Psychology* 15, no. 2 (2010): 177-94; C. A. Morgan, III et al., "Detecting Deception in Arabic: Efficacy of Forced-Choice Testing Dilemmas in Morrocans," *Journal of Intelligence Community Research and Development*, August (2007); C. A. Morgan, III et al., "Efficacy of Verbal and Global Judgment Cues in the Detection of Deception in Moroccans Interviewed Via an Interpreter," ibid.(2008a); C. A. Morgan, III et al., "Detecting Deception through Automated Analysis of Translated Speech: Credibility Assessments of Arabic-Speaking Interviewees," ibid.(2008b); C. A. Morgan, III and G. Hazlett, "Efficacy of Forced Choice Testing in Detecting Deception in Russian," ibid.January(2009a); C. A. Morgan, III et al., "Detecting Deception in Vietnamese: Efficacy of Forensic Statement Analysis When Interviewing Via an Interpreter," ibid. (2009b); C. A. Morgan, III et al., "Efficacy of Automated Forced Choice Testing Dilemmas in Vietnamese," ibid.June(2010); Morgan, Colwell, and Hazlett, "Efficacy of Forensic Statement Analysis in Distinguishing Truthful from Deceptive Eyewitness Accounts of Highly Stressful Events," 1227-34; Vrij et al., "Increasing Cognitive Load to Facilitate Lie Detection: The Benefit of Recalling an Event in Reverse Order," 253-65.

[8] Classification accuracy is determined by calculating the percentage of liars correctly identified as liars and adding this to the percentage of truthful persons correctly identified as truthful and dividing this sum by 2.

[9] DePaulo et al., "Cues to Deception," 74-118; Morgan, Colwell, and Hazlett, "Efficacy of Forensic Statement Analysis in Distinguishing Truthful from Deceptive Eyewitness Accounts of Highly Stressful Events," 1227-34.

https://digitalcommons.usf.edu/jss/vol6/iss3/9
DOI: http://dx.doi.org/10.5038/1944-0472.6.3.9

bioterrorism) and circumstances of interest to the government (i.e., a base rate of lying that was less than the tradition base rate of 50% that is used in most studies).

Our decision to recruit from a population of well-educated individuals knowledgeable about biological materials was based on the following reasoning: 1) We reasoned that this educated population was more representative of individuals targeted for recruitment by parties interested in bioterrorism activities than are college students who may lack seasoned experience in biological experiments; 2) We reasoned that people with expert knowledge about and familiarity with biological materials might, if lying about their experience, experience less cognitive load and, as a result, make detecting deception more difficult. Such information would inform us about whether and to what degree MCI-based detection deception interviews would be practical if used in real world situations involving populations of highly educated subject matter experts. In this study we also employed methods designed to improve our understanding as to how well the MCI method would work if used in real world settings and contexts. Specifically, we:

1) *Increased the complexity of behaviors about which participants lie*. Whereas the participants in most deception studies are asked to lie about an action of low complexity (e.g., having removed the answer key for an exam paper from a professor's office), participants in this study would participate in, and subsequently lie about, a complex set of actions (i.e., being recruited by a 'terrorist' buyer; surreptitiously growing a specific biological culture over the course of one week; preparing the finished culture for delivery; and delivering the package to said buyer). These elements served two important goals: a) the set of activities rendered the paradigm closer to real world dilemmas; and b) the complexity of the actions increased the potential for the memory about these activities to be reasonably rich.

2) *Lowered the base rates of deception from the traditional rate of 50 percent.* In nearly all studies of deception, the base rates of lying are set at 50 percent. Although this feature of scientific design offers investigators a reasonable opportunity to determine whether bio-behavioral 'signals' differ between liars and truth-tellers, such a base rate does not reflect most situations of interest in the real world where professionals are trying to detect a small group of liars in a larger group of truthful persons or a small group of truthful persons in a larger group of liars. In this study we wanted to assess how well validated methods of detecting deception would perform when used in conditions where the base rate of lying was 18 percent.

Hypothesis

We hypothesized that deceptive participants would speak less and use fewer unique words than would truthful participants when interviewed about their activities.

Methodology

Subject Recruitment

The study was reviewed and approved by the New England Independent Review Board (NEIRB). Participants were recruited via advertisements in a local newspaper. All participants had a master's degree or doctoral degree and met the requirement for 'hands-on' experience

working with biological materials in a laboratory environment. All participants were healthy and free of psychiatric or medical illness. All participants provided written, informed consent.

Table 1: Subject demographics.

	N	%		N	%
Gender			Ethnicity		
Male	31	52	Caucasian	49	77
Female	33	48	Black or African American	0	0
Age			Hispanic or Latino/Latina	3	4
Median age: 33.95			Asian or Asian American	10	16
18 to 24	5	8	Am Indian, Pac Islander, Alaska Native	1	1
25 to 34	38	59	Other	1	1
35 to 44	11	17	Income Level		
45 to 54	8	13	0 to $19,999	2	3
55 to 64	2	3	$20,000 to $39,999	11	17
Education Level			$40,000 to $59,999	19	30
Bachelor's Degree	16	25	$60,000 to $79,999	13	20
Master's Degree	34	53	$80,000 to $99,999	11	17
Ph.D.	14	22	$100,000 to $119,999	6	9
			$120,000 to $139,999	1	2
			$140,000 to $159,999	1	2

Subject Randomization

All participants in the study knew that this study was designed to assess lying and truth-telling. So as not to alter their behavior with respect to our primary dependent measures, participants were not made aware of our hypotheses about alterations in speech content data. Participants were assigned to truthful or deceptive groups using a pseudo-randomization method in order to ensure a low base rate of deception (i.e. 18%).

Instructions for Participants Assigned to the Truthful Group

Individuals assigned to the truthful group were instructed to visit a local coffee shop[10] for 30 minutes. All were told that while in the coffee shop they were to pay attention to what occurred around them during their time in the coffee shop and that they would later be quizzed about their memory about their time in the coffee shop. Participants were free to buy coffee, read the newspaper, and/or speak to other customers if they wished. Participants were told they would receive a minimum of $50 dollars for their participation in the study.

One week later participants returned to the research laboratory and were then told:

> "You are going to be interviewed by a member of our research team who is a law enforcement officer. Please be completely truthful when you are interviewed by the law enforcement officer; remember that this person does not know whether we are paying

[10] Blue State Coffee, Wall Street, New Haven, Connecticut.

https://digitalcommons.usf.edu/jss/vol6/iss3/9
DOI: http://dx.doi.org/10.5038/1944-0472.6.3.9

> you to lie or tell the truth about your time in the coffee shop; If they decide that you are lying you will lose the extra money (i.e. $200) that you might have gained in this study so it is important that you appear as honest and sincere as possible when you are interviewed about your time at the coffee shop."[11]

Instructions for Participants Assigned to the Deceptive Group

Prior to visiting the coffee shop, deceptive participants also told the information noted above. However, they were also were shown a picture of the 'terrorist' (i.e. a member of the research team who posed as a criminal offering money for the production of biological materials). They were also given his telephone number and dialed it to speak with him. He then gave them directions to a local coffee shop. After ending the telephone call, participants went to the coffee shop and met with him for 30 minutes. During their meeting, participants received physical materials (i.e., flour, a small quantity of yeast starter hidden in a pill bottle, and several bowls and containers in which to store the culture) and instructions for creating and growing a biological culture (i.e., a yeast culture). The 'terrorist' reviewed instructions with the participant on how to ensure proper growth of the yeast, on the volume required, and on methods for packaging the culture prior to shipping. He then scheduled a follow-up meeting with participants in order to pay them for having prepared the biological culture.

One week later, and after their second meeting with the 'terrorist', deceptive participants returned to the laboratory for their interview. Each was reminded by the research coordinator that when they were interviewed by the law enforcement investigator, they would have to deny having any knowledge about the 'terrorist', his identity, his activities, or having any involvement in growing the yeast culture. The research coordinator reminded deceptive participants that they were allowed to be truthful about having visiting the coffee shop. Each was reminded to appear 'as honest and sincere as possible' when with the law enforcement investigator so as not to raise suspicion on the part of the investigator. Each was reminded that they risked losing the bonus research money (i.e. $750) if the investigator correctly detected that they were lying.[12]

Interview Procedure

The law-enforcement professionals who conducted the modified cognitive interviews in this study were selected because each had more than 15-20 years of real world experience in law enforcement investigative settings. None were novices with respect to interviewing or to the task of assessing for deception. Each interviewer was blind to the status of participants and to the base rate of lying in the study. The MCI interview was conducted using a semi structured format and was *identical* for all participants. While interviewed, all participants sat in a chair facing the interviewer. Participants were told by the interviewer:

> "Well, as I said, my name is _____ and I am part of a team that is investigating some suspicious activity that has been occurring around the Blue State Coffee shop. We've gathered evidence to support our investigation and now we are speaking with people who we believe have information to help us. Our investigation has

[11] As in previous studies, all participants were paid in full for their participation in the study.
[12] All participants were paid in full for their participation in the study.

> shown that people, like you, who have knowledge of making biological materials, have been meeting with another person to discuss these matters. Our investigation has also shown that this "other person" is paying people to actually make the biological material. As you can see this is a very serious matter and so I must tell you it is important that everything you say during our interview is the absolute truth. I'd now like to ask you some questions about your week. Within the last week, did you visit the Blue State Coffee shop?"

Once the person affirmed they had visited the coffee shop, the interviewer initiated the MCI. As shown in table 2, the version of the MCI used for this investigation consisted of five traditional mnemonic prompts and a sixth, experimental prompt. After completing the MCI, participants were escorted from the interview room and debriefed about the study.

Table 2: MCI prompts.

#	Name	Description
1	Full Recall	Tell me everything you remember from your time in the coffee shop. Be as detailed as you can be; Begin with when you entered the shop and end with the time-point you exited the shop. Don't leave anything out even if you think is it trivial or unimportant.
2	Visual	If I had been with you in the coffee shop what would I have seen from the time you entered the shop until the time you left?
3	Auditory	If I had been with you in the coffee shop what would I have heard from the time you entered the shop until the time you left?
4	Emotional	What was the experience in the coffee shop like for you?
5	Temporal	Please start with the last thing you remember and tell me, in reverse order, everything you remember. Like we were running a movie backwards. Start with the last thing that happened (i.e. you leaving the shop) and finish with the first thing you remember doing (i.e. coming into the shop).
6	Mistakes	Do you think that you left anything out or made any mistakes in what you have told me of your memories from the coffee shop?

Data Analysis

Interviewer Judgments

At the conclusion of each interview, the interviewer made a judgment about the status (Truthful/Deceptive) of the person they interviewed. If the interviewer thought the person was deceptive, the judgment was coded as a '1'; if the interviewee was thought to be truthful, the judgment was coded as a '0'. These values were entered into a statistical spreadsheet[13].

Cross-tab analyses were performed using the variables Genuine Status (i.e. the true assignment of the participant) and Interviewer Judgment. Based on these it was possible to calculate the values for True Positive (i.e., deceptive persons correctly judged to be deceptive), False Positive (i.e., truthful persons erroneously judged to be deceptive), True Negative (i.e., truthful persons correctly judged to be truthful) and False Negative judgments (i.e., deceptive persons erroneously judged to be truthful).

[13] Spss Ver. 19, IBM, Armonk, NY.

https://digitalcommons.usf.edu/jss/vol6/iss3/9
DOI: http://dx.doi.org/10.5038/1944-0472.6.3.9

Following this, we calculated variables of Sensitivity, Specificity, Positive Predictive Value, Negative Predictive Value, Positive Likelihood Ratio (LR+), Negative Likelihood Ratio (LR-), and Classification Accuracy.[14]

Professional Rater Judgments

The five raters in this project had more than 15-20 years of real world experience in law enforcement investigations or real world experience in DoD or Intelligence settings that involved credibility assessment type work. None was a novice to the subject matter of detecting deception. Each of the five raters independently viewed each of the 64 videos and made judgments about participants' status (Truthful/Deceptive). If the rater judged an interviewee to be deceptive, the judgment was coded as a '1'; if an interviewee was judged to be truthful, the judgment was coded as a '0'. These values were entered into a statistical spreadsheet[15] and a 'summary judgment' score for each interview was created. The values of such scores ranged from 0-1. All interviews with a summary score less than 0.5 were coded as 'truthful'; all with a summary score greater than 0.5 were coded 'deceptive'; all that received a score of 0.5 were coded as 'Tied/Undecided'. Final Rater Judgment Scores for Truthful interviews were coded as '-1', those that were tied as a '0', and those that were deemed deceptive as '+1'. These values were entered into a statistical spreadsheet.

Cross-tab analyses were performed using the variables Genuine Status (i.e. the true assignment of the participant) and the final summary Rater Judgment Scores. Based on the cross-tabulation results, values for the predictive nature of the raters' judgments were calculated using the same methods as used for the interviewer's judgments.[16]

MCI Speech Content Data

Using transcripts of the MCI, we calculated variables of Response Length (RL)[the total number of words uttered in response to MCI prompts], Unique Word (UW) count [the total number of unique words uttered in response to MCI prompts] and the type-token ratio (TTR) [the ratio of UW to RL]. General Linear Model Multivariate Analyses of Variance[17] were performed using Status (Truthful, Deceptive) as the independent variable and the Speech Content variables (i.e. TTR, RL and UW from each of the five prompts of the MCI) as the dependent variables. ROC Curves and Graphs were generated to evaluate which variables were best at discriminating between the two groups of participants. In order to calculate the classification accuracy of the models, stepwise and forward binary logistic regressions were performed using the most useful speech content variables. Following this, we calculated the True Positive, True Negative, False Positive, and False Negative values from the regression. Following this, we calculated Sensitivity, Specificity, PPV, NPV, LR+, LR-, and Classification accuracy.

---

[14] *Sensitivity* = (TP/(TP+FN);*Specificity* = (TN/(TN+FP); Positive Predictive Value = (TP/(FP+TP); Negative Predictive Value = (TN/(FN+TN); LR+ = (*Sensitivity/1-Specificity*)
LR- = (*1-Specificity/Sensitivity*); Classification Accuracy = (%TP + %TN)/2

[15] SPSS.

[16] D. G. Altman and J.M. Bland, "Statistics Notes: Diagnostic Tests 1: Sensitivity and Specificity," *BMJ* 308, no. 6943 (1994a): 1552; D. G. Altman and J. M. Bland, "Statistics Notes: Diagnostic Tests 2: Predictive Values," ibid.309, no. 6947 (1994b): 168-69.

[17] SPSS.

Results

Interviewer Judgments

At the conclusion of their interviews, each interviewer rendered a judgments as to the veracity of the participant. As noted in table 3, the interviewers identified a total of 47 people as truthful and 17 as deceptive. However, of the 47 people they identified as truthful, 8 were actually deceptive individuals. Similarly, of the 17 people the interviewers identified as deceptive, 13 were misidentified as deceptive and were, in fact, truthful individuals. Thus, interviewers correctly classified 39 of 52 truthful people (i.e. 75%) and 4 of 12 deceptive people (33%). Thus, the classification accuracy was (75+33)/2 or **54 percent.**

Table 3: Cross-tab calculations using interview judgments and true status of participants.

	True Status		
Interviewer Judgment	**Truthful**	**Deceptive**	**Total**
Truthful	39	8	47
Deceptive	13	4	17
Total	52	12	64

The *Sensitivity* of the professional judgments (i.e., TP/(TP+FN) = 4/12) was 33%; The *Specificity* of the professional judgments (i.e., TN/(TN+FP) = 39/52), was 75%; The Positive Predictive Value for professional judgments (i.e., TP/(FP+TP) = 4/17), was 24%; The Negative Predictive Value for professional judgments (i.e., TN/(FN+TN) = 39/47), was 83%; The LR+ for professional judgments (*Sensitivity/1-Specificity*) was .33/.25, or 1.32; The LR- of the professional judgments (*1-Specificity/Sensitivity*) was .25/.33, or 0.75.

Although interview judgments were not highly accurate with respect to participants' Status (truthful/deceptive), a significant relationship was observed between ethnicity of participants and the judgments of the interviewers. Participants who were not Caucasian were significantly more likely to be judged as 'liars' by the interviewers than were participants who were Caucasian (Pearson Chi-Square = 8.27; $p<0.04$).

Professional Rater Judgments

Cross-tab calculations using rater judgments and genuine status of the participants are displayed in table 4. Due to a technical malfunction, the video component (but not the audio component) of two participants randomized to the truthful condition were compromised. Therefore, unlike the interviewers who rated all 64 participants, the raters were only able to view 62 videos of participants (50 truthful; 12 deceptive). The raters identified 33 people as truthful and 29 as deceptive. However, of the 33 people they identified as truthful, seven were actually deceptive individuals. Of the 29 people the raters judged to be deceptive, only five were genuinely deceptive individuals. Thus, raters correctly classified 26 of the 52 truthful people whose videos they viewed (i.e. 50%) and five of 12 deceptive people (42%). The raters misidentified 58% of the deceptive people and 46% of the truthful people. The classification accuracy for raters was: (50+42)/2 or **46 percent.** The *Sensitivity* of the professional judgments (TP/(TP+FN) = 5/12 = 42%; the *Specificity* of the professional judgments (TN/(TN+FP) = 26/50, or 52%; the Positive Predictive Value for professional judgments (TP/(FP+TP) = 5/29, or 17%; the Negative

https://digitalcommons.usf.edu/jss/vol6/iss3/9
DOI: http://dx.doi.org/10.5038/1944-0472.6.3.9

Predictive Value for professional judgments (TN/(FN+TN) = 26/33, or 79%; LR+ for professional judgments (*Sensitivity/1-Specificity*) = .42/.52, or 0.81; LR- of the professional judgments (*1-Specificity/Sensitivity*) = .52/.42, or 1.2. While these indicate that the *sensitivity* of rater judgments were superior to those of the interviewers, the *specificity* of rater judgments was lower.

Table 4: Cross-tab calculations using rater judgments and true status of participants.

	True Status		
Rater Judgment	**Truthful**	**Deceptive**	**Total**
Truthful	26	7	33
Deceptive	24	5	29
Total	50	12	62

Analysis of MCI Elicited Speech Content

Task: Memory of Events While Person Was in the Blue State Coffee Shop

The primary Speech Content variables [Type-Token Ratio (TTR), Response Length (RL), and Unique Words (UW)] were calculated from transcripts of the audio-recordings of all 64 participants. In order to detect whether these variables differed between deceptive and truthful participants, we executed General Linear Model Multivariate Analyses of variance using TTR, RL, and UW from each prompt of the MCI and for the full story as the dependent variables and Status (Deceptive/Truthful) as the fixed factor. The model was significant for both the intercept ($F [1,18] = 848$; $p<0.000$] and for Status ($F [1,18] = 2.7$; $p<0.004$). Tests of between subjects effects for Status with respect to the separate components of the MCI (i.e. the individual prompts) and for the full story (i.e. the prompts combined together) indicated the following: MCI Prompt 1 (RL: $F [1,62] = 8.7$; $p<0.004$; UW: $F [1,62] = 11.3$; $p<0.001$; TTR: $F [1,62] = 14.6$; $p<0.000$); MCI Prompt 3 (RL: $F [1,62] = 5.8$; $p<0.02$; UW: $F [1,62] = 7.4$; $p<0.008$; TTR: $F [1,62] = 11.6$; $p<0.001$); MCI Prompt 5 (RL: $F [1,62] = 8.7$; $p<0.005$; UW: $F [1,62] = 12.8$; $p<0.001$; TTR: $F [1,62] = 2.9$; $p<0.09$); MCI Full Story (RL: $F [1,62] = 7.7$; $p<0.007$; UW: $F [1,62] = 8.7$; $p<0.004$; TTR: $F [1,62] = 3.7$; $p<0.06$).

Table 5 also gives the specific values for TTR, RL, and UW for each of the prompts as well as for the full interview (i.e. all the responses combined from the MCI) for both deceptive and truthful participants.

Table 5: Speech variables for truthful and deceptive participants during MCI.

	Truthful	Deceptive	Signif.
Prompt 1-Full Recall			
TTR	.53 (SD=0.11)	0.67 (SD=.08)	p<0.000
RL	235.8 (SD=175)	84.5 (SD=50)	p<0.004
UW	108.9 (SD=55)	53.3 (SD=29)	p<0.001
Prompt 2-Visual			
TTR	0.61 (SD=.16)	0.67 (SD=.15)	p<0.2
RL	148.9 (SD=134)	93.3 (SD=77)	p<0.2
UW	73.0 (SD=43)	53.8 (SD=33)	p<0.2
Prompt 3-Auditory			
TTR	.64 (SD=.12)	0.77 (SD=.16)	p<0.001
RL	113.3 (SD=82)	54.2 (SD=38)	p<0.02
UW	64.5 (SD=33)	36.9 (SD=22)	p<0.008
Prompt 4-Emotional			
TTR	0.72 (SD=.14)	0.78 (SD=.10)	p<0.2
RL	49.5 (SD=36)	40.8 (SD=19)	p<0.4
UW	31.8 (SD=18)	30.5 (SD=11)	p<0.8
Prompt 5-Temporal			
TTR	0.51 (SD=.90)	0.56 (SD=.13)	p<0.09
RL	182.3 (SD=93)	98.5 (SD=66)	p<0.005
UW	86.0 (SD=33)	49.7 (SD=24)	p<0.001
Full Story*			
TTR	0.36 (SD=.11)	.42 (SD=.08)	p<0.06
RL	729.7 (SD=405)	371.3 (SD=59)	p<0.004
UW	364.3 (SD=145)	224.1 (SD=93)	p<0.002
*Includes data from all prompts			

Binary Logistic Regression analyses indicated the classification accuracy achieved using RL and UW from the first MCI prompt was 84.4% (96% for truthful; 33% for deceptive). Separate Binary Logistic regression analysis using the RL and UW for the Full Story (FS) [i.e. all the data from all five prompts] resulted in a classification accuracy of 84.4 percent (98% for truthful; 25% for deceptive participants). The Binary Regression model correctly classified three of the 12 liars, classified 51 of 52 truth-tellers and misclassified one truthful person as a liar. Thus, the True Positive value was three and the False Negative value was nine. The True Negative value was 51 and the False Negative value was one.

Thus the *Sensitivity* (i.e. TP/(TP+FN) is 3/(3+9), or **27 percent.** The Specificity of the test (TN/(FP+TN) was 51/(51+1) or **98 percent.** The low Sensitivity indicates that the test did not capture many of the deceptive persons; the high Specificity, however, suggests that the speech content may work very well as a 'negative screening tool' – meaning that if the test classifies a person as 'truthful/innocent', the person is very likely to be innocent.

In this study, we were more interested in finding which members of the group were lying. Thus, PPV was calculated to determine whether a person identified by the regression as a liar was truly

https://digitalcommons.usf.edu/jss/vol6/iss3/9
DOI: http://dx.doi.org/10.5038/1944-0472.6.3.9

a liar.[18] The PPV [TP/(TP+FP)] = 3/(3+1), or **75 percent**. This high value indicates that the majority of people identified by the test as 'Liars' were, in fact, liars and that the false positive risk (i.e. calling a truthful person a liar) was small. The Negative Predictive Value (NPV) of a test is determined by dividing the number of 'true negative calls' [i.e. that the person is innocent] by the total number of 'negative calls' one makes using the test, or TN/(TN+FN). The NPV was 51/(51+9), or **88 percent**. This high value indicates that the majority of 'negative tests' are valid and that there were very few false negatives (i.e. few liars considered to be truth-tellers) by this test.[19] However, given that the prevalence of lying was low, it was more meaningful to calculate *likelihood ratios* with respect to determining status because -unlike PPV and NPV – likelihood ratios are not dependent on the 'prevalence of lying' within the group. These are presented below following the Receiver Operating Characteristics (ROC) calculations.

Most detecting deception studies report on the differences between the means of deceptive and truthful *groups.* However, knowing the means and group differences are not often useful to real world investigators who must make judgments about the status of specific individuals. ROC calculations offer a way for professionals to examine data in a way that permits one to make judgments at the individual as opposed to the group level. To illustrate this within the context of the present study, we performed ROC analyses using all MCI variables that differed significantly between groups (see Table 6) and used as the targeted status being 'Truthful.' Table 6 shows the Area under the Curve, and Asymptotic Significance for each of the variables of interest.

Table 6: Area under the curve associated with speech content variables generated in MCI.

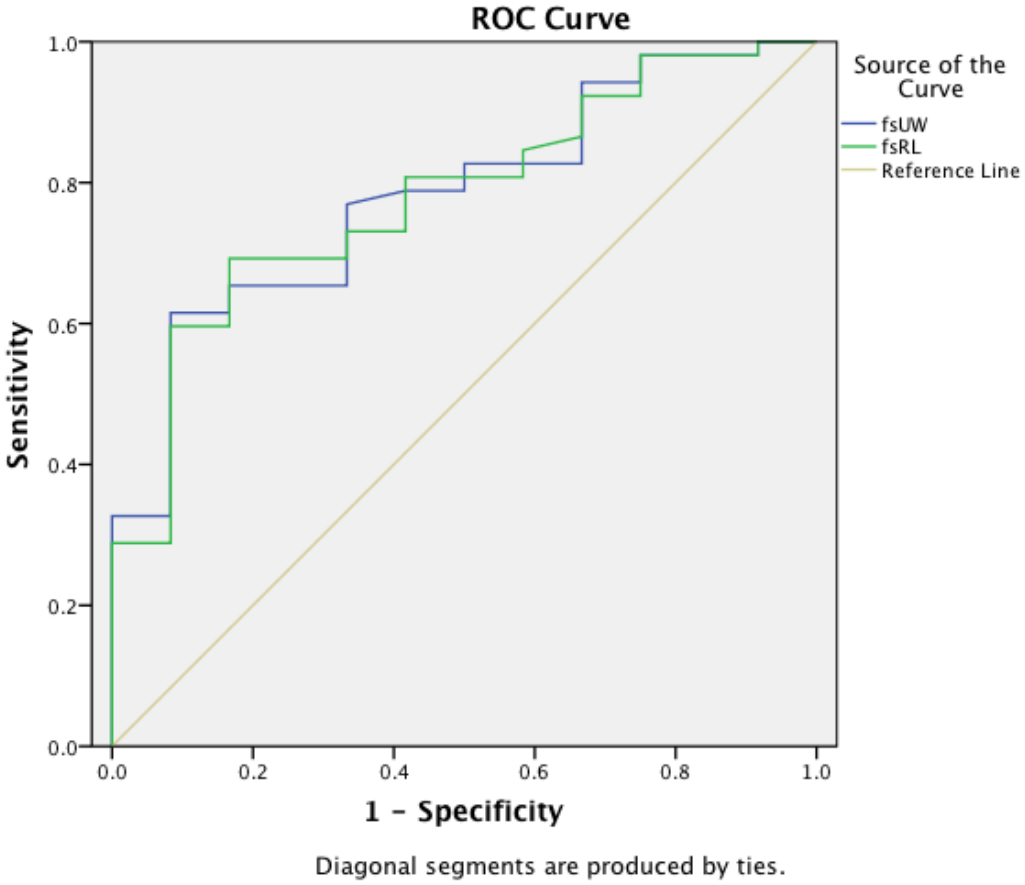
	Area	Asymptotic Sig.
Prompt 1-Full Recall		
TTR	0.18	p<0.000
RL	0.83	p<0.000
UW	0.84	p<0.000
Prompt 3-Auditory		
TTR	0.26	p<0.01
RL	0.73	p<0.02
UW	0.73	p<0.02
Prompt 5-Temporal		
RL	0.79	p<0.002
UW	0.82	p<0.001
Full Story*		
RL	0.78	p<0.003
UW	0.78	p<0.003
*Includes data from all prompts		

[18] This is akin to using a medical test, finding a positive result, and determining how likely it is that the test has truly identified someone with the disease of interest.

[19] NOTE: It is useful to remember that the PPV is not "intrinsic" to the test – PPV depends on the prevalence of the condition one is trying to detect. PPV is directly proportional to the prevalence of liars in the group. If our group of 64 scientists had included a greater proportion of deceptive/liars, then the PPV would have been HIGHER, and the NPV, LOWER.

To demonstrate how ROC analysis can give one the ability to make judgments about the Status (Deceptive/Truthful) of any single individual in this project, we calculated the ROC graph and table (see Figure 1 and Tables 7 & 8) for the speech data from the MCI 'full story.'

Figure 1: ROC graph depicting MCI Full Story Variables RL and UW.



As noted in Figure 1, both MCI full story variables the area under the curve is significant and serves to indicate that each variable can be useful in making predictions about status. As shown in tables 7 and 8 below, if *Response Length* is used as the primary variable by which one is to make a decision, the probability of being wrong (i.e. *1-Specificity*) in judging a person as who has spoken 884 or more words as 'innocent/truthful', is approximately less than 0.1%; Similarly, the probabilities of being wrong if the person has spoken 458 words or 531 words are approximately 17% and 8%, respectively. If, instead of *Response Length* one used *Unique Word count* as the primary variable, then the probabilities of being wrong in judging as 'innocent/truthful' a person who has spoken 284, 300, or 423 unique words would be 17%, 8% and 0.1%, respectively.

https://digitalcommons.usf.edu/jss/vol6/iss3/9
DOI: http://dx.doi.org/10.5038/1944-0472.6.3.9

Table 7: Full MCI Response Length ROC columns and reference data for Sensitivity and I-Specificity

Table Result Variables(s)	Positive if Greater than or Equal to	Sensitivity	I-Specificity
FS* RL			
	107.00	1.00	1.00
	117.50	1.00	0.92
	142.00	0.98	0.92
	163.50	0.98	0.83
	197.50	0.98	0.75
	228.00	0.96	0.75
	252.00	0.94	0.75
	274.00	0.92	0.75
	277.50	0.92	0.67
	281.50	0.90	0.67
	293.00	0.88	0.67
	313.50	0.86	0.67
	325.00	0.84	0.58
	343.50	0.82	0.58
	362.00	0.81	0.58
	365.00	0.81	0.50
	373.50	0.81	0.42
	382.00	0.79	0.42
	388.00	0.77	0.42
	392.50	0.75	0.42
	394.50	0.73	0.42
	397.50	0.73	0.33
	403.00	0.71	0.33
	425.00	0.69	0.33
	449.50	0.69	0.25
	458.00	0.69	0.17
	465.00	0.67	0.17
	479.00	0.65	0.17
	504.50	0.64	0.17
	523.50	0.62	0.17
	526.50	0.60	0.17
	531.00	0.60	0.08
	552.50	0.58	0.08
	577.00	0.56	0.08
	600.00	0.54	0.08
	619.00	0.52	0.08
	654.50	0.50	0.08
	694.50	0.48	0.08
	706.50	0.46	0.08

	726.00	0.44	0.08
	748.00	0.42	0.08
	757.00	0.40	0.08
	764.00	0.39	0.08
	772.00	0.37	0.08
	786.50	0.35	0.08
	824.00	0.33	0.08
	856.00	0.31	0.08
	866.00	0.29	0.08
	884.50	0.29	0.00
	928.00	0.27	0.00
	1000.50	0.25	0.00
	1079.00	0.23	0.00
	1157.50	0.21	0.00
	1214.50	0.19	0.00
	1231.10	0.17	0.00
	1242.10	0.15	0.00
	1251.00	0.14	0.00
	1273.00	0.12	0.00
	1332.50	0.10	0.00
	1373.50	0.08	0.00
	1448.00	0.06	0.00
	1532.00	0.04	0.00
	1658.50	0.02	0.00
	1774.00	0.00	0.00

*Includes data from all prompts

https://digitalcommons.usf.edu/jss/vol6/iss3/9
DOI: http://dx.doi.org/10.5038/1944-0472.6.3.9

Table 8: Full MCI Unique Word count ROC columns and reference data for Sensitivity and I-Specificity.

Table Result Variables(s)	Positive if Greater than or Equal to	Sensitivity	I-Specificity
FS* UW			
	87.00	1.00	1.00
	91.50	1.00	0.92
	100.50	0.98	0.92
	112.50	0.98	0.83
	140.50	0.98	0.75
	163.50	0.96	0.75
	171.00	0.94	0.75
	180.50	0.94	0.67
	184.50	0.92	0.67
	189.50	0.90	0.67
	194.50	0.89	0.67
	195.50	0.87	0.67
	198.00	0.85	0.67
	211.00	0.83	0.67
	227.00	0.83	0.58
	232.50	0.83	0.50
	234.00	0.81	0.50
	236.00	0.79	0.50
	239.50	0.79	0.42
	245.00	0.77	0.33
	249.00	0.75	0.33
	253.00	0.73	0.33
	258.00	0.71	0.33
	260.50	0.69	0.33
	261.50	0.67	0.33
	263.50	0.65	0.33
	273.50	0.65	0.25
	284.00	0.65	0.17
	291.00	0.64	0.17
	297.50	0.62	0.17
	300.00	0.62	0.08
	302.00	0.60	0.08
	303.50	0.58	0.08
	307.50	0.56	0.08
	321.50	0.54	0.08
	340.50	0.52	0.08
	357.50	0.50	0.08
	366.50	0.48	0.08
	370.00	0.46	0.08

	375.50	0.44	0.08
	380.00	0.42	0.08
	393.00	0.40	0.08
	404.50	0.39	0.08
	407.00	0.37	0.08
	409.50	0.35	0.08
	415.00	0.33	0.08
	422.50	0.33	0.00
	433.00	0.31	0.00
	449.50	0.29	0.00
	463.50	0.27	0.00
	481.00	0.25	0.00
	501.50	0.23	0.00
	518.50	0.21	0.00
	529.00	0.19	0.00
	538.50	0.17	0.00
	548.00	0.15	0.00
	553.50	0.14	0.00
	559.50	0.12	0.00
	566.50	0.10	0.00
	571.50	0.08	0.00
	578.50	0.06	0.00
	624.50	0.04	0.00
	673.00	0.02	0.00
	683.00	0.00	0.00

*Includes data from all prompts

Using the ROC tables we calculated positive likelihood ratios (LR+) [i.e., *Sensitivity/1-Specificity*].[20] An *LR+* greater than 1 indicates that the test result is associated with the 'condition' one is trying to detect; a value of less than one is associated with the absence of the 'condition' one is trying to detect. *LR+* values that lie close to '1' have little practical import in that the 'post-test probability' is little different from the 'pre-test probability'. When *LR+* values are greater than or equal to 5 (or the *LR- less than 0.2)* they can serve well as a screening tool.

In this study, the ROC tables were calculated to detect 'truthful' individuals. Calculations using an RL value of 362 words, would result in an LR+ of .81/.58, or 1.39. This value is very close to 1 and is unlikely to be useful for 'screening' the group.[21] However, calculations using a value of 600 words result in an LR+ of .54/.08, or 6.7. This value would be very useful for 'screening' this group of scientists and letting people go who are likely to be innocent. Using the ROC table (see above) an LR+ of 5 or greater can be achieved by using a 'cutoff' RL value of 531 words. If one were to "screen" or "triage" the group of scientists in the present study and "released"

[20] This is the equivalent of calculating the probability that a person who tests 'positive' truly has the condition one is trying to detect, divided by the probability of a person who tests 'positive' who does not have the condition one is trying to detect.

[21] Harrell F, Califf R, Pryor D, Lee K, Rosati R (1982). "Evaluating the Yield of Medical Tests". *JAMA* **247** (18): 2543–2546.

https://digitalcommons.usf.edu/jss/vol6/iss3/9
DOI: http://dx.doi.org/10.5038/1944-0472.6.3.9

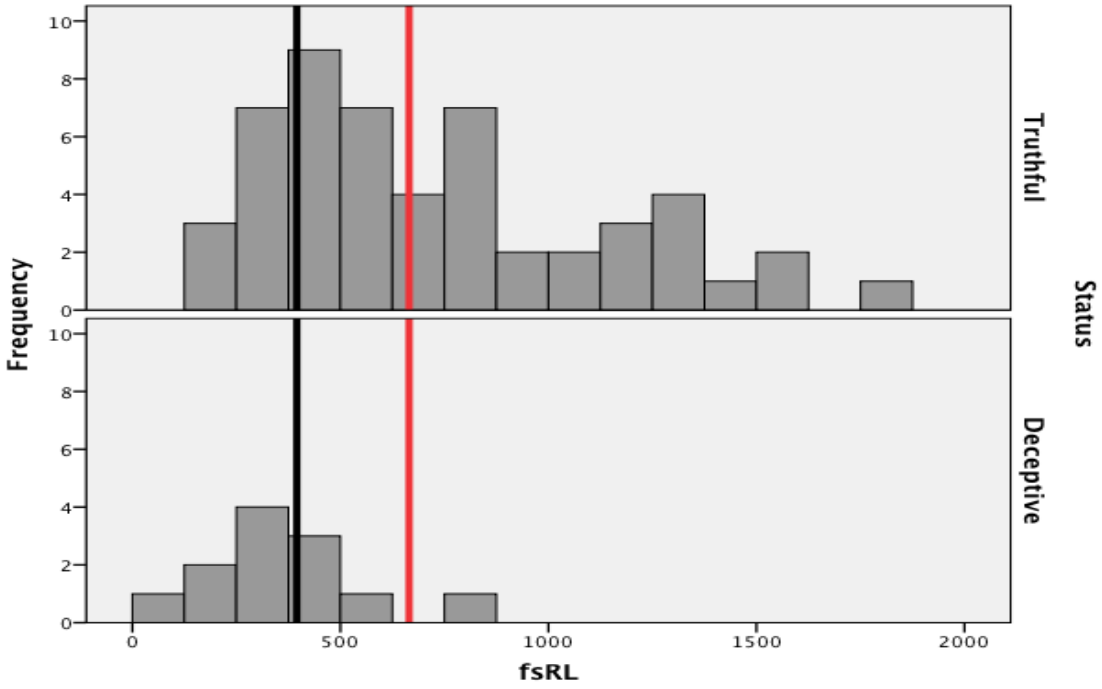
individual scientists based on whether they had spoken 531 or more words in their interview (i.e. a full MCI RL value of 531), one would effectively reduce the number of scientists from 64 to 31. This smaller group of 31 scientists would contain all but one of the 12 deceptive scientists. Having a smaller number people to question further reduces the workload on the police doing an investigation, while offering a greater chance to detect scientists who were knowledgeable about the bio-threat.

Figure 2 illustrates this point in a different way by using histograms. The red line indicates the mean[22] RL for the entire group of 64 scientists. If investigators used the mean RL for the full MCI (i.e. the red line) as the 'cut point', and released everyone who spoke more than the group mean, they would be left with a smaller group of people that included 11 of the 12 (i.e. 92%) of the deceptive participants. If, instead, investigators wanted to detain fewer people and used, in stead of the group mean, a "cut off" point corresponding to the bottom 3rd of the entire group of 64 scientists (i.e. the black line), this would result in them detaining an even smaller group of scientists (i.e. 17) - but one that included seven of the 12 liars (i.e. 58%).

It is important to emphasize that the decision about which "cut points" ought be used by investigators is completely arbitrary. The ROC data simply provide information about the probability that one may err when judging a scientist's status. Unlike statistical analyses in which one reports whether observed "differences" between groups are likely to have occurred by chance – as reflected in p values that are less than or equal to 0.05 – ROC data permit a professional to choose any cut point on a ROC table. The investigator can then see, for any value selected, the likelihood that he or she is wrong in their decision to release a scientist. As noted in the tables, "cut-points" associated with 1-specificities less than 50% offer an advantage to human judgments.

[22] Average.

Figure 2: FS RL in truthful and deceptive participants



Discussion and Conclusion

Our overall aim in this project was to assess the efficacy of MCI when applied to a group of interest (i.e. scientists) with expertise in an issue of interest (i.e. production of biological materials) under conditions of interest (i.e. low base rate deception). Based on the present findings, we conclude that MCI is effective in that judgments based on MCI data were superior to judgments made by Interviewers and Professional Raters.[23] The MCI retained practical efficacy when applied to lower base rate conditions of deception and has the potential to enhance credibility assessments performed by national security professionals.

The modest detecting deception levels for human judgments noted in this study are not an artifact of 'inexperienced personnel.' To the contrary, the interviewers and raters who participated in this study were selected because each had years of real world experience in law enforcement investigations settings, DoD, or Intelligence settings that involved credibility assessment type work; none were novices. The pooled scores from these professionals are within the range previously noted in the scientific literature for laymen and professionals alike when trying to 'detect deception' and suggest that the present findings are not anomalous.[24] This said, we were surprised that our interviewers who performed the MCIs did not do better at detecting deception. This may be due to the fact that although they executed the method accurately for the study, the interviewers were not highly experienced with the MCI and resorted, when forming their judgments about a person, to heuristics from their professional line of work. As such we

[23] All raters had over 15 to 20 years of experience in credibility assessments related to law enforcement or national security.
[24] DePaulo et al., "Cues to Deception," 74-118; Hazlett and Morgan III, "Efficacy of Two Deception Detection Strategies When Assessing Individuals within Cross-Cultural Circumstances: Scientific Technical Report."

https://digitalcommons.usf.edu/jss/vol6/iss3/9
DOI: http://dx.doi.org/10.5038/1944-0472.6.3.9

believe these results realistically reflect the limited capabilities of even the most experienced professionals in detecting deception under more complex circumstances and the potential offered by methods such as MCI to said professional groups.

In the real world, investigators must often decide how to efficiently and effectively triage a large group of people down to a smaller, more manageable group of individuals who they would like to interview further. When triaging a group, they hope that the smaller group will contain people who are genuinely knowledgeable about the issue they are investigating. The current data illustrate how MCI elicited speech data might be used in real situations. In this study, the MCI did not require much time to execute (i.e. it took, on average, four to eight minutes per person) and the information obtained in the MCI could be used to triage the group rather effectively. Hypothetically, if interviewers in this study had used the MCI in order to calculate how much each scientist had talked about their time in the coffee shop (i.e. the Response Length, or RL) and then decided to 'detain' all scientists who had talked less than the average scientist for the group as a whole (i.e. used a RL *below* the mean of the entire group of 64 scientists,) *all but one* deceptive scientist would have been kept in the smaller, 'detained group'. This illustrates how a low tech method like MCI might significantly assist investigators by helping them work efficiently and reduce their workload while increasing the likelihood of finding the deceptive people.

As in previous MCI deception studies, the liars in this study exhibited reduced verbal productivity.[25] The present findings illustrate that well educated laboratory scientists were not immune to this MCI 'deception induced' effect. It is reasonable, therefore, to speculate that the MCI will prove useful when used in real world settings involving highly educated individuals and when used under conditions of low base rate deception. It bears noting that a number of truthful participants spoke very little during the MCI. Some truthful participants were very nervous and this may have caused them to speak less during the MCI – and look like the group of liars. In our view, to optimize the effectiveness of MCI techniques, it is important to minimize the level of anxiety experienced by an interviewee so that a sense of anxiety or nervousness does not inhibit the degree to which they may speak in response to the mnemonic prompts of the MCI. If a person is at ease, there is a greater likelihood that deception related increases in cognitive load are responsible for the reduction in overall effectiveness of the memory prompts.

Not all of the MCI prompts elicited responses that differed between liars and truth-tellers. In this study the response to the visual and the emotional prompts did not differ between liars and truth-tellers. However responses to the auditory and temporal (reverse order) prompts did differ significantly between the groups. The reasons for this are not clear. It is possible that this group of intelligent laboratory personnel may have found it easier to report visual content – as opposed to auditory content – since reporting visual observations may be a more familiar task to these scientists and, as a result, created less cognitive load. Alternatively, it is possible that because deceptive persons, while meeting with the 'terrorist' in the coffee shop, were less attentive to the

[25] Vrij et al., "Increasing Cognitive Load to Facilitate Lie Detection: The Benefit of Recalling an Event in Reverse Order," 253-65; Colwell, Hiscock, and Memon, "Interviewing Techniques and the Assessment of Statement Credibility," 287-300; Colwell et al., "Vividness and Spontaneity of Statement Detail Charateristics as Predictors of Witness Credibility," 5-30; Morgan, Colwell, and Hazlett, "Efficacy of Forensic Statement Analysis in Distinguishing Truthful from Deceptive Eyewitness Accounts of Highly Stressful Events."

auditory stimuli and, as a result, had less memory information to access when presented the MCI auditory prompt. Although we do not yet understand the varied effectiveness of MCI prompts this phenomenon has been observed in previous investigations.[26] Thus, it seems practical and useful to continue using an MCI format that includes a variety of mnemonic prompts as this maximizes the opportunity for discerning between truthful and deceptive responses. Future studies designed to assess why and how individuals differ in their responses to the prompts might elucidate this issue.

Acknowledgements

This project was funded by a grant from the U.S. Federal Bureau of Investigations (FBI) High-Value Interrogation Group (HIG) research division. We would like to thank Jim Kline, Wes Clark and Susan Hill for their assistance on this project.

[26] Hazlett and Morgan III, "Efficacy of Two Deception Detection Strategies When Assessing Individuals within Cross-Cultural Circumstances: Scientific Technical Report; Morgan, Colwell, and Hazlett, "Efficacy of Forensic Statement Analysis in Distinguishing Truthful from Deceptive Eyewitness Accounts of Highly Stressful Events," 1227-34.

https://digitalcommons.usf.edu/jss/vol6/iss3/9
DOI: http://dx.doi.org/10.5038/1944-0472.6.3.9