MEDWIN PUBLISHERS
Committed to Create Value for Researchers

International Journal of Forensic Sciences
ISSN: 2573-1734

Detecting Deception Using MCI in Twitter

Morgan CA*, Barthen J and Mol SA

Department of National Security, Professor & Forensic Psychiatrist, Henry C Lee College of Criminal Justice and Forensic Sciences, University of New Haven, USA

***Corresponding author:** Charles A Morgan, Department of National Security, Professor & Forensic Psychiatrist, Henry C Lee College of Criminal Justice and Forensic Sciences, University of New Haven, USA, Tel: (203) 932-1154; Email: CMorgan@newhaven.edu

Mini Review
Volume 9 Issue 1
Received Date: February 02, 2024
Published Date: February 14, 2024
DOI: 10.23880/ijfsc-16000358

Abstract

The purpose of the study was to test whether Modified Cognitive Interviewing (MCI) is an effective method for detecting deceptive human eyewitness accounts in a computer-mediated computer platform with limited text space (such as in Twitter). 44 college students were the participants in this study, where they either had to perform, or pretend that they had performed a cognitive task. Of the 44 participants, 15 performed the task and reported truthfully about their activities; 16 performed the task and denied having participated in the task; 13 read the instructions about the cognitive task and when interviewed claimed to have actually performed the task. The transcripts of interviews, conducted in Twitter, were rated by individuals trained in cognitive interviewing; forensic speech variables (response length (RL), unique word (UW) count and type-token ratio (TTR) were coded from transcripts. Human rater judgments and computer-based speech analysis performed better than chance; computer based judgments were superior to the human judgments (i.e., 79% vs. 54%, respectively). Speech content variables derived from MCI differed significantly, and in different ways, between the truthful and false claimant participants and also between the truthful and denial type participants. MCI derived statement analysis methods are a scientifically valid method, when used in Twitter, that can be used by professionals tasked with distinguishing between true claims, false claims and denials.

Keywords: Social Media; Eyewitness Memory; Lying; “X”; Statement Analysis

Abbreviations: RL: Response Length UW: Unique Word; TTR: Type Token Ratio; MCI: Modified Cognitive Interviewing.

Introduction

Modified Cognitive Interviewing (MCI) has proven to be an effective method to detect deception [1-4].

However, the increase in interactions via computer-mediated communication (CMC) is resulting in an interest in the dynamics of deception in online environments.

Over the past 10 years, numerous detecting deception studies have consistently demonstrated that an analysis of speech content – not voice stress - can consistently result in detecting deception rates at or above 82% - a rate that is significantly higher than that achieved by polygraph (i.e. 65%) or by chance (i.e. 50%) [1,2,5-10].

These studies are important and demonstrate that speech content analysis – when derived through the use of a specific interviewing technique known as modified cognitive interviewing (MCI)–can be used to detect deception at rates significantly above other current approaches used by professionals.

At present, no studies have examined whether this MCI could effectively be used to discriminate truthful from denials or false claims in persons questioned in the computer-mediated environment of Twitter about the same task. This absence of research data means that our current understanding of how best to detect deceptive communications in social media is limited. Although one might assume that some of the language characteristics associated with deception that have been detected in traditional interviewing environments will also be manifested when people communicate in social media settings, at present there is a lack of research that might inform such an assumption. Thus, understanding whether alterations in speech content vary significantly between truthful claims, false claims or denials in social media would clarify not only whether features of deception generalize across communication environments, but also how well the method can be applied by professionals in the real world.

The present study was designed to examine A) whether MCI, performed in a limited text based environment (Twitter), derived speech content could be used to distinguish truthful accounts from false claims or denials in people who performed a cognitive activity; and B) the efficacy of MCI in Twitter compared to results previously derived from MCI in other formats.

Methods

Subjects: 44 students at the University of New Haven were the participants of this study. Each was given an oral briefing about the project and each provided written informed consent prior to participation in the study.

Design

This study design consisted of two phases: In Phase One participants engaged in, or only read about, a cognitive task; in Phase Two participants were interviewed about their claimed activity by interviewers through Twitter. When interviewing participants, the interviewers used a Modified Cognitive Interview (MCI) [1-4]. The questions were adapted to fit into one Tweet each (140 characters). Transcripts from the Twitter interviews with participants were rated by human raters and also used to generate speech content variables for computer analysis.

Phase One: Task Exposure

Depending on the randomization system, participants engaged in, or only read about the cognitive task: All truthful persons completed the task; deceptive participants assigned to the "denial" group also completed the task. Deceptive participants assigned to the "fabrication" group were only permitted to read the instructions of the task.

The task involved participants in a series of timed trials during which they had to make use of a set of shapes (i.e. using a commercially available game called Tangoes) to construct an image that matched the figure shown to them by the instructor administering the task. This game is mentally challenging and as addressed in previous studies Morgan CA, et al. [4] requires significant mental effort for the participant to complete in an accurate manner. The task was considered "complete" when participants completed the task or when the time expired (whichever came first).

After completing their task, participants assigned to the Truthful condition were told that they would be interviewed through Twitter, about how they had spent their time. They were instructed to respond openly and honestly about the nature of their activities. Conversely, after completing their task, participants assigned to the "Denial" deceptive condition, were told that they could not report on their activities and were instructed to lie when interviewed.

Participants assigned to the "false claim" deceptive condition were given written detailed instructions about the task. Each was given 10 minutes to study the materials. Each was told when given the instructions that they would have to lie and claim that they had actually performed the task when interviewed.

Phase 2: The Modified Cognitive Interview

Interviews conducted in the social media platform Twitter. Every participant was assigned a pre-registered Twitter account, in which they had to answer the MCI questions. Each answer could only consist of one Tweet, this means that they only had 140 characters to formulate there answer. The interview was conducted with multiple participants simultaneously, ranging from groups of four to eight participants interviews at the same time.

Transcripts of the Twitter interviews were created and edited* so that only the parts that provided actual memory remained. The transcripts were used to calculate the key Speech Content variables that would be used in the actuarial database: Response Length [i.e. total words uttered by participant], Unique Word Count [total number of unique words spoken] and the Type-Token Ratio (TTR) [i.e. the ratio of Unique Words/Response Length.] Finally, and in order to assess whether the best discrimination would occur when using the interview as a whole, we created these three variables from the entire speech/statement set acquire during the cognitive interview.

*This was done by deleting, filler words (such as; just, like and only) and repeating of the questions.

Seven human raters, mostly law enforcement professionals, trained in MCI, were the raters in this study.

Each independently reviewed the typed interview transcripts of the MCI through an online survey. After reading a transcript, each rendered a judgment about a participants' status (Truthful/Deceptive). If a rater judged a participant to be deceptive, their judgment was coded as a "1"; if genuine, their judgment was coded as a "0". Individual Cross-tab analyses were performed using the variables Genuine Status (i.e. the true assignment of the participant) and each individual's judgment Scores.

Forensic Statement Analysis

In order to assess whether Response Length, Unique Word Count, TTR for the task differed between truthful and deceptive participants we performed General Linear Model Univariate Analyses of Variance using Group (Truthful, False Claim, Denial) as the independent variable and Speech Content (i.e. TTR, Response Length and Unique Word count from each prompt of the MCI) as the dependent variables.

Tukey post hoc tests were used to evaluate how speech content variables differed amongst the three groups (Truthful, False Claim, Denial).

Results

Speech Content Variables

Univariate Analyses of Variance for the cognitive task indicated the presence of significant differences between the three groups of participants on the speech content variables. With respect to the variable TTR, a significant difference was only noted for: Prompt One ($F(2,41) = 4.13$; $p=0.023$).

Significant differences between the groups of participants were also noted for the speech content variable RL in response to Prompt One ($F(2,41) = 2.36$; $p= 0.107$), Prompt Two ($F(2,41) = 3.83$; $p=0.030$); Prompt Four ($F(2,41) = 2.64$; $p=0.084$), and the prompt total ($F(2,41) = 3.06$; $p=0.058$).

Finally, significant differences between the Groups of Participants were noted for the speech content variable UW for: Prompt Two ($F(2,41) = 3.55$; $p=0.038$); Prompt Three ($F(2,41) = 2.36$; $p=0.107$); Prompt Four ($F(2,41) = 3.28$; $p=0.048$), and the prompt total ($F(2,41) = 4.66$; $p=0.015$).

These differences were due to the fact that, unlike truthful participants, false claims and denial participants failed to expand on their accounts when exposed to the mnemonic prompts.

Raters

The human raters performed slightly better than chance at distinguishing true from false accounts. Cross tab analyses indicated that the accuracy of the raters was 54% (range 35% to 66%).

Discussion

When exposed to the MCI in Twitter, genuine, false claimants and denial participants behaved in significantly different ways. Although the groups were indistinguishable from each other in the initial detailed prompt, they differed significantly when exposed to the second phase of the MCI (i.e., the mnemonic prompts). Truthful participants exhibited greater Unique Word counts and Response Lengths. The denial and false claim participants were unresponsive to the memory prompts and offered few additional details about their experience.

Within the context, it is reasonable to hypothesize that this reduced responsiveness to the mnemonic prompts could be due to the increased cognitive load associated with lying [11] and a desire to "tell their story" and stick to it so as to be believed, just as in Morgan CA, et al. [4].

Rater judgments were only slightly better than chance (54%). it is possible that raters were unable to assess the transcripts due to the short response length caused by the limited response possibility in Twitter. This finding is lower to that noted for raters in our previous study Morgan CA, et al. [4], where human raters performed significantly better (65%). Further studies may evaluate whether the rater accuracy can be increased. It bears emphasizing however, that as in many studies, the classification accuracies based on the speech content analysis by the computer were higher than those demonstrated by the human raters (78.5% vs. 54%, respectively) [12,13].

As in previous studies, the application of the MCI method was effective at discriminating between genuine and false eyewitness accounts. As noted above, however, this field of inquiry is new and future research is needed in order to clarify the extent to which the current method and accuracies may apply to different types of experiences reported in social media. Although the current finding that MCI was effective for detecting deception in social media suggests this approach to detecting deception may have wider applicability in computer-mediated communications, it is important to acknowledge that social media communication platforms vary in the degree to which people text or video their observations. Therefore, a conservative interpretation of the present findings would be that they should be understood within the context of platforms in which people are sending text based communications.

References

1. Morgan CA, Colwell K, Hazlett GA (2011) Efficacy of Forensic Statement Analysis in Distinguishing Truthful from Deceptive Eyewitness Accounts of Highly Stressful Events. Journal of Forensic Sciences 56(5): 1227-1234.

2. Morgan CA, Rabinowitz YG, Hilts D, Weller CE, Coric V (2013) Efficacy of modified cognitive interviewing compared to human judgments in detecting deception related to bio threat activities. Journal of Strategic Security 6(3): 100-119.

3. Morgan CA, Rabinowitz YG, Leidy R, Coric V (2014) Efficacy of combining interview techniques in detecting deception related to bio-threat issues. Behavioral Sciences & the Law 32(3): 269-285.

4. Morgan CA, Rabinowitz YG, Palin B, Kennedy K (2015) Who Should You Trust? Discriminating Genuine from Deceptive Eyewitness Accounts. The Open Criminology Journal 8: 49-59.

5. Vrij A (2008) Detecting lies and deceit: Pitfalls (Chapter 5). Wiley, Chinchester, UK, pp: 373-388.

6. Colwell K, Hiscock CK, Memon A (2002) Interview techniques and the assessment of statement credibility. Applied Cognitive Psychology 16(3): 287-300.

7. Colwell K, Hiscock-Anisman C, Memon A, Rachel A, Colwell L (2007) Vividness and spontaneity of statement detail characteristics as predictors of witness credibility. American Journal of Forensic Psychology 25(1): 5-30.

8. Morgan CA, Steffian G, Colwell K, Coric V, Hazlett G (2007) Efficacy of forensic statement analysis in distinguishing true from false eyewitness accounts. Journal of Intelligence Community Research and Development (JICRD) pp: 1-35.

9. Morgan CA, Mishara A, Christian J, Hazlett G (2008) Detecting deception through automated analysis of translated speech: Credibility assessments of Arabic-speaking interviewees. Journal of Intelligence Community Research and Development pp: 1-22.

10. Morgan CA, Rabinowitz R, Christian J, Hazlett G (2009) Detecting deception in Vietnamese: efficacy of forensic statement analysis when interviewing via an interpreter. Journal of Intelligence Community Research and Development pp: 1-38.

11. Vrij A, Fisher R, Mann S, Leal S (2009) Increasing cognitive load in interviews to detect deceit. In: Williamson T, Milne B, et al. (Eds.), International developments in investigative interviewing, Willan Publishing, Canada, pp: 176-189.

12. Morgan CA, Hazlett G, Doran A, Garrett S, Hoyt G, et al. (2004) Accuracy of eyewitness memory for persons encountered during exposure to highly intense stress. International Journal of Law and Psychiatry 27(3): 265-279.

13. Morgan CA, Bhatt S (2012) A critical assessment guide to the topic of detecting deception: identifying methods that work and those that don't. Journal of Intelligence Community Research and Development pp: 1-76.