

2018

# Determining indicators of deception in computer mediated communication using eye tracking

Akmal Mirsadikov  
Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/etd>

## Recommended Citation

Mirsadikov, Akmal, "Determining indicators of deception in computer mediated communication using eye tracking" (2018). *Graduate Theses and Dissertations*. 16641.  
<https://lib.dr.iastate.edu/etd/16641>

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact [digirep@iastate.edu](mailto:digirep@iastate.edu).

**Determining indicators of deception in computer mediated communication using eye tracking**

by

**Akmal Mirsadikov**

A dissertation submitted to the graduate faculty  
in partial fulfillment of the requirements for the degree of

**DOCTOR OF PHILOSOPHY**

Major: Business and Technology (Information Systems)

Program of Study Committee:

Joey George, Major Professor

Kevin Scheibe

Sree Nilakanta

Sekar Raju

Huaiqing Wu

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2018

Copyright © Akmal Mirsadikov, 2018. All rights reserved.

**DEDICATION**

I would like to dedicate this dissertation to my late father, Mirsadikov Mirazam, who has instilled in me love for learning. An avid reader, he wanted people around him to read and enjoy learning and made every effort to educate his children. He would, certainly, have been very proud to see this day. May God bless his soul.

I also dedicate this work to my family: my mom, my wife Gulnora, and my three children Mirmuhammad, Madina, and Mahmud. You all have served as my motivation throughout this journey and continue to be the source of my inspiration.

## TABLE OF CONTENTS

	Page
LIST OF FIGURES .....	vi
LIST OF TABLES .....	vii
ACKNOWLEDGMENTS .....	ix
ABSTRACT .....	x
CHAPTER 1. INTRODUCTION .....	11
CHAPTER 2. DETERMINING INDICATORS OF DECEPTION IN VARYING MEDIA MODES USING EYE TRACKING .....	15
Abstract .....	15
Introduction .....	15
Literature Review .....	20
Deception detection .....	20
Hypotheses development .....	22
Eye tracking .....	34
Research Method .....	38
Pilot study .....	38
Main Study Procedure .....	39
Measures .....	44
Analysis and Results .....	46
Fixation duration on the speaker's eyes .....	48
Fixation frequency on the speaker's eyes .....	49
Fixation duration on the speaker's mouth .....	50
Fixation frequency on the speaker's mouth .....	50
Fixation duration on the speaker's torso .....	50
Fixation frequency on the speaker's torso .....	51
Detection Accuracy .....	51
Analysis of the research model .....	53
Honest messages with sound .....	54
Honest message without sound .....	54
Deceptive message with sound .....	54
Deceptive message without sound .....	55
Discussion .....	55
Implications .....	60
Limitations and future research .....	62
References .....	63
Appendix. IRB Approval .....	69

CHAPTER 3. DETECTING DECEPTION IN TEXT: WHAT YOUR EYES SEE.....	70
Abstract.....	70
Introduction .....	70
Literature review.....	74
Human visual system.....	74
Eye tracking.....	75
Reading.....	76
Eye movements in reading .....	78
Deception.....	82
Text-Based Cues to Deception.....	86
<i>Theme 1: Cognitive load</i> .....	88
<i>Theme 2: Psychological emotions</i> .....	90
<i>Theme 3: Sensory and contextual information</i> .....	93
<i>Theme 4: Memory</i> .....	94
Omitted linguistic cues.....	94
Detection accuracy .....	95
Research Method .....	96
Experimental stimuli .....	96
Pilot study.....	96
Main Study .....	98
Measures.....	100
Analysis and Results.....	101
Theme 1: Cognitive load .....	101
Regression count.....	101
Total time.....	102
Fixations Duration and Count Total .....	102
Causation .....	103
Exclusive.....	104
Motion verbs .....	104
Theme 2: Psychological emotions.....	105
Certainty.....	105
Generalizing terms .....	105
Passive voice verbs .....	105
First-person pronouns .....	106
Third-person pronouns.....	106
Negative emotions .....	106
Negations .....	107
Theme 3: Sensory and contextual information.....	107
Sensory processes .....	107
Temporal.....	107
Spatial .....	108
Theme 4: Memory .....	108
Insight .....	108
Cognitive processes .....	108
Detection accuracy: All cues .....	109
Detection accuracy: By theme .....	111

Discussion.....	115
Implications and future research .....	123
Limitations.....	124
References .....	125
Appendix A. Definitions of Linguistic Cues to Deception (Hauch et al., 2015).....	129
Appendix B. Stimuli Statements.....	131
Statement #1 .....	131
Statement #2 .....	131
Statement #3 .....	131
Statement #4 .....	132
Statement #5 .....	132
Statement #6 .....	133
Statement #7 .....	133
Statement #8 .....	133
Statement #9 .....	134
Statement #10 .....	134
Statement #11 .....	134
Statement #12 .....	135
Appendix C. IRB Approval .....	136
CHAPTER 4. CONCLUSION.....	137

**LIST OF FIGURES**

	Page
Figure 1-1. Research model .....	13
Figure 2-1. Assigned Areas of Interest .....	45
Figure 4-1. Research model .....	137

## LIST OF TABLES

	Page
Table 2-1. Detectability of cues to deception across various media (Lewis, 2009) .....	27
Table 2-2. Distribution of video snippets in two stimulus sets for pilot study .....	38
Table 2-3. Demographic information .....	39
Table 2-4. Stimulus sets selected .....	42
Table 2-5. Arrangement of Treatments per Stimuli Set.....	43
Table 2-6. Distribution of Observations per Treatment for 6 Stimuli Sets.....	44
Table 2-7. Covariance estimates for variable Actor .....	47
Table 2-8. Estimates of means of fixation durations .....	48
Table 2-9. Estimates of mean differences of fixation durations .....	48
Table 2-10. Estimates of means of fixation frequencies.....	49
Table 2-11. Estimates of mean differences of fixation frequencies.....	49
Table 2-12. Parameter Estimates .....	52
Table 3-1. Participant demographic information .....	98
Table 3-2. Pairwise comparisons – mean differences.....	102
Table 3-3. Pairwise comparisons of fixation duration on AOIs - Mean difference.....	103
Table 3-4. Pairwise comparisons of fixation count on AOIs -Mean difference .....	104
Table 3-5. Parameter estimates of logistic regression. Each estimate is analyzed individually.....	110
Table 3-6. Parameter estimates of logistic regression. Each estimate is analyzed individually.....	110
Table 3-7. Parameter estimates of logistic regression for theme 1 variables.....	112
Table 3-8. Parameter estimates of logistic regression for fixation duration of theme 2 variables.....	113



Table 3-9. Parameter estimates of logistic regression for fixation duration of theme 3 variables.....	114
Table 3-10. Parameter estimates of logistic regression for fixation duration of theme 4 variables.....	115
Table 3-11. Results of hypothesis testing. ....	116

## ACKNOWLEDGMENTS

There have been many people who have supported my pursuit of academic career. They have been alongside me for the last five years, guiding me, consoling me, and most of all, believing in me. Without them I would have never been able to accomplish this work. I would like to thank each and every one of them.

I would like to thank my family, especially my wife Gulnora, for her selfless devotion and love, who has carried all the burdens of raising our family on her shoulders and allowed me to focus solely on my education. I thank my mom for her encouragement and prayers. I thank my friend Nuriddin Ikramov for his reviews and edits of my work and guidance in my job pursuit. I thank my children, Mirmuhammad, Madina, and Mahmud for their love and patience and for bearing with me through the completion of my study.

I would like to thank my committee members, Dr. Scheibe, Dr. Nilakanta, Dr. Raju, and Dr. Wu for their guidance and support throughout the course of this research. I would especially like to thank my committee chair, Dr. Joey George, who has put up with my constant interruptions from his work and who has been a true advisor and mentor in almost every matter. Without your guidance and encouragement this journey would have been a dead end. I would also like to thank my late professor, Dr. Brian Mennecke, for his tremendous support and love. I will be forever grateful for your guidance, patience, and support.

In addition, I would also like to thank my friends, colleagues, the department faculty and staff for making my time at Iowa State University a wonderful experience.

## ABSTRACT

This work combines two studies, both identifying indicators of deception through the analysis of the visual attention of a veracity judge. using eye tracking. In the first study, we investigated the effect of the varying media modes on detection accuracy through the analysis of the visual behavior of veracity judges. We employed eye tracking technology to understand where the judges looked at and what impact their visual foci had on their detection performance. We found that the visual foci of the judges varied as a result of the message veracity and media modes. Judges fixated longer and more frequently on the mouth and the torso of the communicators in deceptive messages. In video-only modes, the judges fixated longer on the mouth of the sender. Fixation frequency on the eyes and the mouth of the sender worsened deception detection accuracy. In the second study, we investigated the reading behavior of veracity judges when presented with honest and deceptive statements produced in high-stake, real-life scenarios with potential negative consequences for the individuals who produced those statements. We found that the reading metrics of veracity judges varies across honest and dishonest statements and the linguistic cues that the judges focus on have an effect on deception detection performance.

## CHAPTER 1. INTRODUCTION

Adoption of computer technologies accelerated the shift of our communication into the computer mediated communication (CMC) environment using emails, text messaging, videoconferencing, Voice Over Internet Protocol (VoIP), and other modes. The Internet and especially Web 2.0 altered our means of seeking information and verifying its accuracy. Of course, not all information that we come across or that we are communicated is credible; some is deceptive. In fact, deception is part of our everyday life. Researchers generally categorize lies as outright lies, exaggerations, and subtle lies. Outright lies, also referred to as falsifications or fabrications, are complete lies with no truth in them. A student who states that his paper is ready, while he has not even started writing one, is telling an outright lie. A job applicant who attempts to embellish her résumé by stating that she is competent in Java while her level could at best be categorized as a beginner is exaggerating. Concealing information by omitting details or literal truth designed to mislead the target are examples of subtle lies. In this work, we refer to the person attempting to deceive as a sender or a liar, and the target of deception as a receiver.

Wide-spread adoption of CMC has created new venues for deception. In spite of its prevalence in our daily communication, both private and business, research has shown that we are poor detectors of deception. The relationship between media and deception detection, while emerging, is understudied.

Investigation of deception and its detection has attracted researchers from multiple disciplines. These researchers have attempted to identify reliable indicators of deception with a goal to improve deception detection success rates. Not all deceptive communication necessitates its accurate detection. Some lies are not nefarious in their nature, such as lies

produced to maintain social relationships, and thus do not bear negative consequences. Other lies, however, are more serious in their nature, and identifying those would bring benefits to organizations and to societies. Identifying terrorism plots or serious crime offenders certainly benefits our society. Similarly, hiring viable job candidates, or assigning business contracts to reliable business vendors, is of great importance to organizations.

Traditionally, empirical research in deception detection has relied on trained coders for identifying cues to deception. Such coders watch video footage or listen to audio recordings of liars and truth tellers. They apply certain coding systems to record the frequency and duration of behaviors that the truth tellers and liars display. Next, they compare their results to identify whether a particular nonverbal or verbal behavior was more prevalent in truth tellers or liars. The coders, however, are prone to mistakes: they may misidentify a cue, fail to record one, or make other mistakes, and therefore coding is conducted by multiple coders to minimize the error rate. Contemporary tools and techniques used by deception researchers allow them to overcome the shortcomings of depending on coders. Eye tracking technologies, for example, provide a unique insight into an individual's viewpoint and provide information about what the person looked at, how long they looked, and what pattern her gaze displayed. This technology records all eye movements of the individual and thus provides objective, complete, and accurate data on what the person looked at when assessing the veracity of another person or of a written statement. Access to such data allows researchers to better identify behaviors of senders (in video format) and written cues (in text) associated with accurate detection of deception or truth telling.

Most research on deception detection has investigated deception in face-to-face settings, and thus the relationship between communication media and deception detection is

not well understood. Particularly, the research stream on understanding the relationship between the differences in communication modes and accurate deception detection is scant. The central focus of this work, thus, is to understand the relationship between message veracity, media, and deception detection through the analysis of the point of regard of the veracity judge. Figure 1 shows the guiding research model of this work.

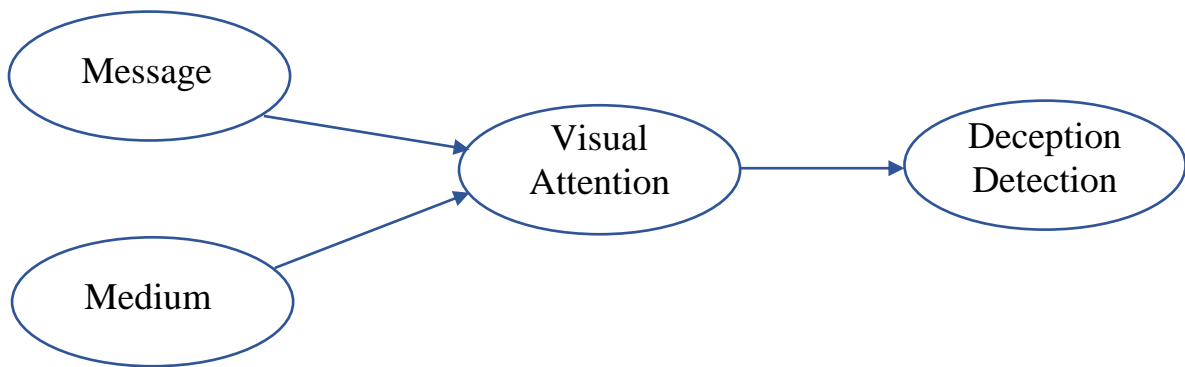


Figure 1-1. Research model

This dissertation is comprised of two studies, each addressing specific research questions pertaining to the research model. In the first study, applying eye tracking technology, we attempt to address the following three research questions:

1. Where do veracity judges look when presented with varying communication modes?
2. Where do veracity judges look when being lied to or told the truth?
3. What is the relationship between what judges focus on and their deception detection success?

In the second study we investigate deception in written text. Specifically, we focus on deception produced in real-life, high-stake settings and attempt to understand what linguistic

cues were noticed by the veracity judges and how those cues affected their deceptive detection performance. Thus, the second study addresses the following two research questions:

1. What do veracity judges look at when assessing the veracity of a text statement written in a high-stakes setting?
2. What is the relationship between what veracity judges choose to focus on and their detection accuracy?

To address these research questions, we conducted experiments by recruiting college students from a large Midwestern University. In the first study, we identified areas that judges focused on when presented with varying media modes and when being lied to and told the truth. We further investigated how what judges looked at affected their detection accuracy. In the second study, we examined linguistic cues that the judges focused on when presented with deceptive and honest written statements and how those linguistic cues affected their detection accuracy. This exploratory study provides a unique approach of studying deception from the perspective of a veracity judge and identifies areas where the judges would focus when communicated through different medium modes and when being lied to.

## CHAPTER 2. DETERMINING INDICATORS OF DECEPTION IN VARYING MEDIA MODES USING EYE TRACKING

Akmal Mirsadikov/Iowa State University

Modified from a manuscript to be submitted to MIS Quarterly

### Abstract

Our daily interactions, both personal and professional, are taking place more and more through mediated modes of communication at this era of interconnectedness. In computer mediated communication (CMC), we seem to continue our habits of face-to-face interactions, including lying. Prior research on deception and its detection has focused on identifying indicators of deception, but deception taking place in CMC has not been researched well, especially in investigating the varying role of media on deception detection performance. In this study we investigated the effect of the varying media modes on detection accuracy through the investigation of the visual behavior of veracity judges. We employed eye tracking technology to understand where the judges looked at and what impact their visual foci had on their detection performance. We found that the visual foci of the judges varied as a result of the message veracity and media mode. Judges fixated longer and more frequently on the mouth and the torso of the communicators in deceptive messages. In video-only modes, the judges fixated longer on the mouth of the sender. Fixation frequency on the eyes and the mouth of the sender worsened deception detection accuracy.

### Introduction

People, as social beings, need to communicate with each other. In doing so, we tend to tell lies, both in the pursuit of self-interest or for altruistic purposes, i.e. for the benefit of others. For example, while a student caught cheating on an exam may lie to avoid being expelled from a school, innocuous statements, such as a nice comment about a friend's



terrible haircut, could be intended to benefit someone other than the deceiver. Research on deception has shown that lying is quite prevalent in daily life. DePaulo and her colleagues suggest that people lie on average one or two times a day (DePaulo et al., 1996). Most lies that people tell are pedestrian in their nature and are about people's preferences, attitudes, and feelings.

People generally prefer to differentiate truth tellers from those attempting to deceive, especially when the stakes of acting on deceptive information are high. Managers making hiring decisions in organizations, for example, must be able to differentiate valid candidates from those who attempt to deceive. Job applicants tend to lie both on their résumés (Guillory and Hancock, 2012) and during job interviews (Fisher, 2014), and the costs of hiring the wrong people could be quite high down the road. Phishing is another form of deception, where perpetrators, through the use of electronic communication channels, attempt to gain certain benefits by persuading the victims to perform certain actions (Abdelhamid et al., 2014). Just in the third quarter of 2017, 296,208 unique phishing reports worldwide were submitted to the Anti-Phishing Working Group (APWG, 2018).

However, previous research has shown that humans are not good at detecting lies. In fact, we can detect lies with about 54% success rate, or slightly better than by chance (Bond and DePaulo, 2006; Miller and Stiff, 1993; Vrij, 2000). This number, however, does not reflect people's real ability to detect deception. The 54% detection rate is based on receivers' ability to correctly determine truth when the message is truthful, and correctly determine deception when the message is deceptive (Bond and DePaulo, 2006). In fact, people are generally good at detecting truth (about 70-80% of the time) but very bad at detecting deception (only 35-40%) (Levine et al., 1999; Sun Park and Levine, 2001).

Research on deception detection is based on the premise that the behaviors of liars, both verbal and non-verbal, are different from those of truth tellers. Researchers have tried to identify those behaviors and other reliable indicators of lying, or cues, that are associated with lying. More than five decades of research on deception detection has not come up with even a single cue that is always associated with lying (akin to Pinocchio's nose), although some verbal and non-verbal cues to deception have been identified to be more reliably associated with deception than other cues on average. A well-cited meta-analysis that investigated many cues that have been associated with lying identified only a small set of cues that are diagnostic of deception (DePaulo et al., 2003).

A large proportion of research on deception has investigated deception in face-to-face (FtF) interactions. Increasingly, however, our daily communication is mediated through various communication technologies (i.e. smart phones, chat rooms, social media boards and messenger apps, FaceTime, emails, and others). People continue to use deception in technology-mediated environments, although the relationship between deception, its detection, and media is not clear (George and Robb, 2008). While the research about detection of deception when communicating face-to-face is well established, the stream of research on detecting deception when a communication medium is involved is relatively less investigated (George et al., 2016; Hancock, 2007). While lying remains a common part of daily interaction in computer-mediated communication or CMC (George and Robb, 2008; McHaney et al., 2017), the detection accuracy in CMC remains equal to that of face-to-face communication (Hancock et al., 2010). This similar detection performance rate raises some interesting questions. How do people assess honest and dishonest information across varying CMC modes? In their discussion of the impact of varying media on the accuracy of deception

detection, Bond and DePaulo (2006) suggest that the video (only) medium generally invites an application of a liar stereotype and results in a poorer detection accuracy rate. In video only conditions, the veracity judges do not have access to the verbal content of the stimuli and depend on visual cues, to which they attend based on their beliefs. In this study, we investigate how the choice of media affects the attention foci of the observers. Thus, the first research question guiding this study attempts to address the effect of media differences on the observer's behavior. Specifically, we attempt to investigate how, if at all, the visual attention of veracity judges varies under varying media.

The research on deception detection proposes multiple explanations for poor detection performance. For example, cognitive biases, including a truth bias, whereby people tend to believe the speaker, or a lie bias, whereby targets decode all incoming messages as deceptive (McCornack and Levine, 1990), have been shown to impede veracity assessment (George et al., 2008). Other factors affecting the accuracy of detection include motivations of both a deceiver and his/her target, their experience in lie detection, message severity, time to rehearse a lie or to prepare a response when asked, and many other reasons. Vrij (2008) suggests that one of the reasons for such poor detection rates is that detectors pay attention to cues that are not reliably associated with deception. Researchers have used different techniques and tools in studying deception and its detection (Bond and DePaulo, 2006). Generally, researchers conduct laboratory-based experiments where participants, mostly students, provide truthful and/or deceptive messages. Their verbal responses and non-verbal behavior, or cues, are coded by trained third party coders, whose objective is to map objective cues to deception (Hartwig and Bond, 2011; Vrij, 2008).

The limitation of this approach is that the coders are prone to mistakes and may not code all verbal and non-verbal cues that veracity judges may observe. To overcome the shortcomings of the human coders and to utilize the progress made in technological innovations to help detect physiological nuances not observable by human detectors, researchers have employed various tools and techniques that assess physiological, vocal, and brain activity measures of deceivers (Granhag et al., 2015). They have embraced the use of polygraphs, voice stress analyzers, thermal imaging cameras, brain scanning technologies such as fMRI and EEG, and eye tracking technology to detect the deceit (Elaad, 2014; Vicianova, 2015). The eye tracking technology, which helps to investigate the gaze behavior of the subjects, enables the researches to tap into objective data not previously available to them. As technology advances, these tools are becoming increasingly affordable, less invasive and easier to use, and promise exciting insights into this topic.

Our study employs eye-tracking technology to objectively determine the cues that veracity judges actually look at when assessing the targets. This approach helps us analyze the eye movement behavior of people who assess the truthfulness of the senders and to compare this behavior to the cues that coders have been trained to code. A closer look at the cues that our participants focus on may reveal new areas not previously mentioned in the deception detection literature, and help us both, to confirm what we already know about cues and potentially to add new ones to the list of detection behavior. Thus, the second research question seeks to investigate how, if at all, the visual attention of veracity judges varies across honest and dishonest messages. As the ultimate purpose of our study is to investigate the relationship between the cues and the detection accuracy, our third research question

seeks to understand, how, if at all, the areas that judges choose to focus on affect their detection accuracy.

Next, we will review the literature on deception detection, media, and eye tracking and discuss theoretical approaches that help us derive our hypotheses. Then, we will present the research method and explanation of the procedures for the pilot and main studies, and present our findings. We conclude the paper with a discussion of results.

## **Literature Review**

### **Deception detection**

People generally want to differentiate truths from lies. We say “generally” because our social interactions would be unbearable if everyone only spoke the truth all the time. A mediocre cook who spent three hours cooking a dinner to please her spouse would not want to hear an honest assessment that any takeout would have been better. There are, however, people whose job it is to accurately distinguish lies from the truth, at least in the professional context. For example, lawyers, prosecutors, law enforcement officials, customs officers, insurance personnel, journalists, judges, and others would want to accurately discriminate between truth and lies to make informed decisions.

What is deception? In this study, we follow Buller and Burgoon (1996), whereby they define deception as “a message knowingly transmitted by a sender to foster a false belief or conclusion by the receiver.” This definition of deception suggests that not all deviations from truth are considered deception. The deceiver must “knowingly” attempt to distort his message or behavior. Thus, the intention of the sender is key in defining the act of deception. Vrij suggests that both successful and failed attempts to deceive are considered deception (Vrij, 2008). Hancock suggests a similar definition for deception that takes place in CMC,

which he refers to as “digital deception” and which incorporates the technologically mediated nature of interaction (Hancock, 2007).

People lie every day about various issues. On average, we lie about one or two times a day, with most lies being about our feelings, preferences, opinions and attitudes (DePaulo et al., 2003). DePaulo and her colleagues (1996) concluded that lying was a commonplace attribute of daily life, wherein people lied in approximately one-third of their daily communication. In a similar study, employing a diary study methodology, Hancock and colleagues found that over a quarter of interactions were deceptive (Hancock et al., 2004). A replication study of Hancock et al. concluded that lying is a common part of everyday discourse, not only in face-to-face interactions, but also increasingly in CMC environments (George and Robb, 2008).

Why do people lie? There are many reasons, and they can be condensed into three broad categories: (1) instrumental - to achieve certain benefit or exercise power, (2) relational - to maintain the desired relationship, and (3) identity based - to avoid embarrassment or to project a desired image (Buller and Burgoon, 1996). In contrast to a lay person’s belief that people lie for self-interested purposes, many lies are intended to benefit others, especially those lies designed to maintain social interactions. Vrij (2008) dubs them “social lies” (p. 20) or “social lubricant,” whereby the sender acts both in self-interest and in the interest of others. Vrij suggests similar motivations for people’s deception: (1) for one’s own benefit or for the benefit of others; (2) to avoid costs or punishment; and (3) for materialistic or psychological reasons (Vrij, 2007).

Given the common use of lies in everyday life and thus our familiarity with lying and our general preference for discerning between lies and truthful messages, why are we so bad

at detecting deception? Deception detection researchers have investigated this and related questions to identify reliable indicators of lying, the reasons for lying, and to find tools and techniques to improve our performance as lie detectors. Of over 150 cues that have been identified and investigated over the years, only a handful of cues to deception have been reliably associated with lying (DePaulo et al., 2003).

People are generally bad at detecting lies because they underestimate their ability to lie and overestimate their ability to detect lies (Elaad, 2003). Vrij (2007) suggests that poor motivation, difficulties associated with lie detection, and common errors in detecting lies lead to many lies being left unnoticed. The deception detection accuracy of professionals, such as police officers and customs officers, is generally not better than that of ordinary people (Bond and DePaulo, 2006), which suggests that lie detection is difficult. People also tend to assess the message veracity based on cues that are not reliably associated with deception, which are also referred to as subjective cues to deception (Strömwall et al., 2004).

### **Hypotheses development**

Research on detecting deception has investigated many theoretical lenses to help identify deception more accurately. One research stream is based on the premise that lying is generally harder than truth telling (Ekman and Friesen, 1969; Zuckerman et al., 1981). The multi-factor theory (Zuckerman et al., 1981) suggests that deception is directly associated with psychological factors such as emotions, cognitive effort, and attempted behavioral control. This approach proposes that lying arouses certain emotions in the sender, such as guilt, fear, and excitement. Next, liars attempt to control their behavior, where they attempt to suppress behaviors that they believe are associated with lying and try to display behaviors that they think are related to being honest. Of course, not many people can skillfully enact the desired behaviors, and sometimes they are not even aware of some of the behaviors they

display when communicating (Vrij, 2008). Lying could be mentally taxing as the sender must cognitively process the content of the deceptive message and must remember the details of his story to make it plausible and coherent to the receiver. At the same time, the sender needs to monitor his own behavior and the reactions of the target for any signals of suspicion (Buller and Burgoon, 1996). This extra cognitive burden resulting from multiple tasks that the sender attempts to handle may result in a sender's inferior performance. The multi-factor theory suggests that psychological factors result in behavioral differences that distinguish truth tellers from those lying, i.e., the more senders experience emotions, cognitive load, and attempt to control their behavior, the more likely they are to display cues to deception. For example, a person feeling guilt may avert his gaze and display more speech hesitation, while fear may cause a higher pitch in a sender's voice or a higher blink rate and result in more verbal mistakes. Similarly, higher cognitive load may lead to a longer latency period, less plausible stories, fewer illustrators, and in fewer head and trunk movements in senders. Senders, who engage in strategic behavioral control, may end up looking longer into the eye of the target and provide stories that sound too polished and rehearsed and may repeat their stories more. Such cues may signal the receiver that the message is not genuine and thus probably deceptive.

Ekman and Friesen's (1969) influential theory of deception, leakage theory, is built around nonverbal behaviors of the sender. Similar to the multi-factor framework's view, the leakage theory suggests that lying produces an emotional response that is manifested behaviorally. In the process of information inhibition or behavior simulation, the senders may "leak" (hence the name of the theory) cues to deception through different parts of the body. The theory proposes that the parts of the body vary in their sending capacity, i.e., visible and



interpretable signals displayed by particular parts of the body. The face, they argue, is the richest part of the human body to transmit discernible and visible information to a receiver and thus is the primary focus of the target. Sending capacity is lower in the hands and lowest in the feet.

Leakage theory also discusses two types of feedback – internal and external. External feedback refers to the behavior of the receiver, which informs the sender what the receiver has perceived and evaluated. The internal feedback, on the other hand, is a conscious awareness of the sender of his own behavior. The sender, in an attempt to deceive the receiver, relies on external feedback and monitors the reactions to the message. Because of the sending capacity of body parts, the authors argue, people focus most on the face when seeking feedback, and less on the other parts of the body. The face is the primary site for the display of affect. Leakage theory suggests that the sender, in order to enact the desired affect to support the message, will most focus on his facial features. Extra effort spent on one body part, namely the sender's face, will inhibit the performance of other body parts (hands and feet), which may “leak” cues to deception that the receiver may notice. Leakage theory thus suggests that receivers who focus more on the hands and feet of the sender are more likely to be more accurate in detecting simulated messages than those receivers who focus on the face of the sender.

Buller and Burgoon (1996) introduced Interpersonal Deception Theory (IDT), which merges deception principles with interpersonal communication principles and offers a process view of deception from an interactive, dynamic perspective. It suggests a continuous interplay of interpretations in a face-to-face communication and subsequent enactments based on such interpretations by the two parties involved in a dyadic exchange. The process

of deception in interpersonal communication takes place in three phases: preinteraction, interaction, and postinteraction. The context and the relationship of parties in the preinteraction phase, moderated by parties' expectations, knowledge, and goals, affect their behaviors and cognitions, which in turn affect the interplay in the interaction phase. The interaction phase refines receiver suspicions, if such exist, and subsequent sender behavior. This iterative process concludes with receivers making an assessment of the sender's veracity and sender's assessment of the receiver's suspicion. As suggested, the sender engages in multiple tasks simultaneously: she attempts to control the information conveyed and the behavior displayed, while monitoring closely the reactions of the receiver and adjusts her behavior accordingly. Such cognitive effort to control multiple tasks may not always be successful and may result in unintended performance. Any display of suspicion from the receiver would affect the sender's subsequent behavior. This dynamic exchange of information and displays from both parties results in either the sender's success or failure.

Researchers also tried to identify the beliefs that ordinary people around the world have about the behaviors of liars (Team, 2006). In this study a group of researchers surveyed participants from 58 countries. The respondents were asked what cues they used to identify a lying person, to which more than a hundred different beliefs were provided. Four of those beliefs were mentioned by more than 25% of the respondents: gaze aversion (64%), liars are nervous (28%), incoherent statements (25%), and body movements (25%). In a similar study, the two most common beliefs were found to be gaze aversion (73%) and body movements (25%) (Mann et al., 2004). These studies suggest that judges in assessing the veracity of the senders will be guided by their beliefs and look for indicators that support their beliefs.

All three theoretical frameworks suggest that behaviors of the truth tellers should be different from those of liars. The research on deception detection has investigated extensively the verbal and non-verbal cues to deception. A meta-analysis by DePaulo and her colleagues (2003) investigated 158 such cues and tried to identify the ones that are reliably associated with lying by analyzing the effect sizes (Cohen's  $d$ ) for those cues. In his work, Vrij (2008) summarized findings from multiple meta-analyses and proposed cues he dubbed "objective."

Our discussion of underlying theoretical factors affecting the behavior of liars suggests that deceivers will be expected to behave differently from truth tellers and the judges may react to these differences by attending to them. Moreover, the beliefs that people hold in assessing the veracity of deceivers may guide their visual attention to "anchor" areas, from which they may shift their gaze when they notice deception leakage. Based on the discussion of psychological factors, the multitasking cognitive load in interpersonal interaction, the sending capacities of the body parts, and the leakage of cues, we propose that senders' lies will be signaled behaviorally and verbally and that receivers will react to them. Our first hypothesis hence suggests:

**Hypothesis 1:** *Visual attention of judges will vary across honest and dishonest messages.*

The majority of the research on deception has focused on deception in face-to-face communication, so the relationship between CMC and deception remains poorly understood (George et al., 2008; George and Robb, 2008). Since lying continues to be a part of daily life and as the variety of CMC technologies have been widely adopted and used for ordinary communication, has the choice of media affected deceptive behavior and its detection? The long list of cues investigated by comprehensive meta-analyses were originally researched in

face-to-face interactions, yet many of those cues may not be available under different media modes. Any communication medium will have a reduced number of cues available for deception detection, compared to synchronous face-to-face communication (Daft and Lengel, 1986; George et al., 2016). Furthermore, media vary in their ability to transmit those cues. For example, any cues that are visually and vocally detectable (i.e., pressed lips, body movement, word repetition, logical structure of the message, and others) are available in full audio-visual media modes and cannot be detected in text only media. Similarly, cues that are detectable in audio modes (high pitched voice, latency period, and others) cannot be detected in written modes. There seems to be a hierarchy of cues available for detection afforded by the medium, where full audio-visual modes are able to transmit the most cues and text only mode can transmit the least, while the audio only mode is placed somewhere in between.

Table 2-1 shows the detectability of cues under different communication modes.

Table 2-1. Detectability of cues to deception across various media (Lewis, 2009)

<b>Behavior</b>	<b>Video</b>	<b>Audio</b>	<b>Written</b>
Less talking time	Detectable	Detectable	
Fewer details	Detectable	Detectable	Detectable
More pressed lips	Detectable		
Less plausibility	Detectable	Detectable	Detectable
Less logical structure	Detectable	Detectable	Detectable
More discrepancies and ambivalence	Detectable	Detectable	Detectable
Less verbal and vocal involvement	Detectable	Detectable	
Fewer illustrators	Detectable	Detectable	Detectable
Less verbal immediacy (all categories)	Detectable	Detectable	Detectable
Less verbal and vocal immediacy (impressions)	Detectable	Detectable	
More verbal and vocal uncertainty (impressions)	Detectable	Detectable	
More chin raises	Detectable		
More word and phrase repetitions	Detectable	Detectable	Detectable
Less cooperative	Detectable	Detectable	
More negative statements and complaints	Detectable	Detectable	Detectable
Less facial pleasantness	Detectable		
More nervous and tense (overall)	Detectable	Detectable	

*Table 2-1 continued*

<b>Behavior</b>	<b>Video</b>	<b>Audio</b>	<b>Written</b>
More vocal tension	Detectable	Detectable	
Higher frequency, pitch	Detectable	Detectable	
More pupil dilation	Detectable		
More fidgeting	Detectable		
Fewer spontaneous corrections	Detectable	Detectable	
Less admitted lack of memory	Detectable	Detectable	Detectable
More related external associations	Detectable	Detectable	Detectable

Following the reasoning from face-to-face communication and the potential variety of cues suggested by the aforementioned theories, one would expect that in CMC, the communication modes that are capable of transmitting the highest number of cues would provide better opportunities for the receiver to detect deception and hence, by similar reasoning, one would expect a sender to avoid such types of media. However, the empirical evidence does not provide support for such a straightforward relationship and suggests mixed findings (Burgoon et al., 2008; Burgoon et al., 2010). While some studies have found support of the direct relationship between media and deception detection (Burgoon et al., 2008; Dunbar et al., 2015), some found no support of such a relationship (Burgoon et al., 2010). Two studies by George and his colleague found an indirect, mediated relationship between medium and detection success. In one study, they found an indirect effect of media on detection success through probing. The use of richer media was associated with more probing, which in turn led to more accurate detection (George et al., 2008). In another study, they found a mediated relationship between media and deception detection through sender credibility. When the media increased perceived sender credibility, the accuracy of distinguishing between lies and deception deteriorated (George et al., 2014 2014). There is a

wide range of explanations, and the available empirical evidence may explain this discrepancy.

The academic information systems domain investigates, among many other topics, how information systems help to improve task performance. Media synchronicity theory (MST) was introduced as a theory of communication performance and extended beyond the question of media choice for certain types of tasks. Pointing out the shortcomings of media richness theory (Dennis and Kinney, 1998), which suggests that people are better off choosing specific media based on the nature of the message they want to transmit (Daft and Lengel, 1986), MST argues that matching media capabilities to communication processes improves communication performance (Dennis et al., 2008). Instead of focusing on the communication task, the authors propose focusing on the underlying processes, or steps, as they put it, and suggest that every communication task is composed of different mixes of conveyance and convergence processes. Conveyance processes include gathering and transmission of new information and processing it within an individual to build a mental model of the situation. Convergence processes of communication refer to the processes aimed at building shared understanding among the communicating individuals.

Defining synchronicity as a shared pattern of coordinated synchronous behavior among individuals as they work together, the authors define media synchronicity as “the extent to which the capabilities of a communication medium enable individuals to achieve synchronicity.” (p. 581). The authors suggest five media capabilities that influence media synchronicity: (1) transmission velocity, or the speed at which a medium can deliver a message among communicators, (2) parallelism, or the number of transmissions that a medium can transmit at the same time, (3) symbol sets, or variety of cues or symbols a

medium can transmit, (4) rehearsability, or the extent to which the communication medium allows the sender to rehearse or edit a message before sending, and (5) reprocessability, or the extent to which the communication medium allows the receiver to reexamine or revisit the message either during the process of communication or later on. MST proposes that for conveyance processes use of media supporting lower media synchronicity should result in better communication performance, whereas for convergence processes, use of media with high media synchronicity should improve communication performance. While MST was introduced as a theory of communication where both information sender and receiver are honest and work toward building shared understanding, the same theory was empirically tested to explain strategic choice of media in deceptive communication (George et al., 2013). The researchers sought to investigate whether those intending to deceive preferred certain types of media, as well as the reasons for preferring a certain medium over another. MST was found to be an accurate predictor of media choice, where transmission velocity and symbol sets were overwhelmingly favored by those seeking convergence, and rehearsability and reprocessability were favored by those seeking conveyance.

MST offers a different approach for the assessment of communication medium by focusing on the individual capabilities of a medium, instead of a holistic view. While a full audio-visual communication mode may offer a larger number of symbol sets and provide a synchronous, interpersonal interaction, it lacks in rehearsability and reprocessability. A veracity judge would not be able to revisit the content of the message or would not have a chance to craft better counter arguments or to probe more researched questions. Video modes may also introduce visual and demeanor biases (Burgoon et al., 2008).

In another study, researchers focused on key media characteristics to explain and predict the relationship between media choice and deception (Carlson et al., 2004). They pointed out six key media features: synchronicity, symbol variety, cue multiplicity, tailorability, reprocessability, and rehearsability. These media characteristics are closely aligned with the media capabilities in MST. Synchronicity of a medium is similar to transmission velocity in MST; symbol variety is similar to symbol sets; cue multiplicity is generally the same as parallelism in MST; tailorability refers to the ability of the medium to allow the author of the message to customize the communication event and to personalize it to the needs of a recipient; reprocessability and rehearsability are the same as the identical name media capabilities in MST. Based on these characteristics, Carlson et al. (2004) suggested that deceivers would deceive best when a medium features “higher levels of symbol variety, tailorability, and rehearsability and lower levels of cue multiplicity and reprocessability” (p.20).

Bond and DePaulo (2006) analyzed the deception research findings in light of the double-standard framework. According to this framework, people in general view lying as a negative quality and truth-telling as a virtue. Thus, when judging people who lie, they perceive them as wrong and unacceptable. Yet when they themselves lie, they take a more practical approach and suggest that their lying is innocuous or even is a sanctioned practice. Similarly, they assume the reasons for lying by others as being “nefarious” (p.216). This implies that people have stereotypes of liars, who are “stricken with shame, wracked by the threat of exposure” and thus “leak signs of their inner torment.” (p. 216). This stereotypical approach to assessing deception can lead to inaccurate decisions. Senders who fit this stereotype of a liar may be interpreted as liars, even though there could be other factors that



cause such behavior. For example, a husband suspected of cheating on his spouse may display anxiety or anger, the manifestation of which may very much resemble guilt of acceptance or an attempt to cover up misconduct. On the other hand, those who do not exhibit the stereotypical behavior are most likely to be believed. Empirical results from prior research have shown that deception detection judges base their assessment of a person's truthfulness on the person's demeanor, referred to as a demeanor bias: people who appear most honest when lying are people who also appear most honest when telling the truth (Bond and DePaulo, 2008; Bond Jr and Atoum, 2000; DePaulo and Rosenthal, 1979; Levine et al., 2011).

Based on this double-standard framework, we suggest that people might look for the stereotypical behavior of the message sender in assessing the veracity of the sender. Speakers who fit these stereotypes are expected to be assessed as liars. As discussed earlier, media may vary in how much they invite application of a stereotype by affording different levels of detectability. Communication media that transmit more symbol variety will allow the receivers to build a more thoughtful assessment of the sender, instead of relying solely on a stereotypical image of a liar. A medium with less symbol variety may leave fewer options for a thoughtful assessment and lead the receivers to rely more on their stereotypes. The same message transmitted through different media with varying media capacities is expected to be interpreted differently. Under different communication contexts the sender and receiver cognition and behavior is expected to vary systematically (Buller and Burgoon, 1996). An audiovisual image of a speaker allows a receiver a wider range of behavioral cues than a video image with no sound, which should in turn affect the visual foci of the judges. Hence, we propose our second hypothesis:

**Hypothesis 2:** *Visual attention of judges exposed to full audiovisual stimuli should vary from visual attention of judges exposed to video stimuli only.*

Our earlier discussion of the verbal and non-verbal cues to deception and their detectability under varying media implies that the detection accuracy of the veracity judges is contingent on, among other factors, the cues that judges attend to. In both full audio-visual and video only stimuli, the judges have all visual cues available and only full audio-visual mode affords vocal and paralinguistic cues, and as predicted by Hypothesis 2, the variance in visual foci should result from differences in media capabilities of the two modes. This expected variance in visual attention of judges should manifest itself in varying eye gaze metrics, including the fixation duration and frequency on certain areas of interest (explained further). In the discussion of biases that receivers may have towards the sender, Buller and Burgoon (1996) suggest that when receivers become participants in the communication, instead of observers, they attend more to facial cues. The authors suggest that facial cues are less informative, whereas vocal cues are more informative of deception. Similarly, the application of the stereotypical double-standard framework (Bond and DePaulo, 2006) by the veracity judges can lead to inaccurate decisions, which should contribute to their detection performance. A video message with access to verbal content should provide more content for analysis and thus should not depend solely on stereotypical assessment. In line with this reasoning, Vrij suggests that instead of focusing on single, independent cues to deception, a focus on a combination or a cluster of cues would lead to better detection (Vrij, 2008). Thus our third hypothesis posits:

**Hypothesis 3:** *Variance in visual attention will contribute to the deception detection accuracy.*

In deception detection studies, as suggested earlier, researchers attempt to compare “objective” deceptive behavior, which is usually provided by trained third party coders, with “subjective” assessments provided by deception judges (Vrij, 2008). The scarcity of objective cues, however, makes the judgment task of coders error prone. People are generally not good at accurately reporting what cues they used in making judgments. Their assessments may not be completely based on self-reported cues but might rather be based on intuitive and implicit cognitive processes of which they may not be consciously aware (Hartwig and Bond, 2011). This lack of accurate reporting could be corroborated with an external measure of their behavior, i.e., through eye tracking (Granhag et al., 2015). Eye tracking of elements that lie-catchers looked at, but failed to report in their decision making, could shed some light onto elements that are in the realms of implicit decision making.

### **Eye tracking**

Eye tracking refers to a technique whereby a viewer’s eye movements are measured and the focus of eye gaze is captured so that the researcher knows where the person is looking at any given time. While the early application of this technique involved very invasive methods, advances in technology have enabled more unobtrusive approaches. Today, most commercially available eye tracking systems use a video-based corneal reflection method to measure point-of-regard of a viewer (Duchowski, 2007). Utilizing the information from both the center of the pupil and corneal reflection, this method allows the researcher to disassociate eye movements from the head movements (Duchowski, 2003; Jacob and Karn, 2003).

Studying eye movement data has given researchers an insight into the viewers’ problem solving, reasoning, mental imagery, and search strategies (Poole and Ball, 2006).

This approach has been widely adopted in psychology, human-computer interaction,

marketing, medicine, and other disciplines, as well as in the commercial sector (Poole and Ball, 2006). The main measurements used in eye tracking research are gaze fixations, which are moments when eyes are relatively stable and associated with taking in information, and saccades, quick eye movements between fixations. Because information is believed to be processed only during fixations, most information produced from eye tracking is about fixations (Djamasbi, 2014). Eye tracking systems also provide other metrics, such as pupil size and blink rates. Eye tracking provides objective information about the viewer's gaze behavior, without having to rely on self-reported information. This is important, because previous research has demonstrated that the actual behavior of participants may differ from self-reporting (Bernard et al., 1984; Eysenbach and Köhler, 2002).

Fixation data can also be analyzed based on a specific targeted area(s) of research interest. This is achieved by identifying specific regions known as areas of interest (AOIs). Different types of fixation data can be used to investigate a user's behavior over an AOI, including fixation duration, fixation frequency, time to first fixation on AOI, percentage of viewers, and other metrics (Djamasbi, 2014; Poole and Ball, 2006). Additionally, AOI analysis helps reduce the size and complexity of eye movement protocols significantly (Salvucci, 1999). Identifying various AOIs allows comparing and contrasting regional data quantitatively and drawing inferences as to what area was more noticeable, or more important, to the viewer in making an assessment (Cyr and Head, 2013; Poole and Ball, 2006). In a similar fashion, viewer behavior under varying media could be contrasted and analyzed. We can quantitatively compare whether viewers' foci vary when viewing a video stimulus with sound vs. without sound. Thus, eye tracking is a superior technology for

getting an objective record of what people are looking at when viewing video, text, or other media, as well as for tapping into veracity judges' viewing behavior.

Because of the benefits of eye tracking technology mentioned above, several studies in deception detection research have used this technique. Pak and Zhou investigated the eye gazing behaviors of deceivers in an online video chatting mock dating experiment (Pak and Zhou, 2013). They used gaze fixation data on AOIs and found that deceivers fixated less on their communication partner, compared to honest participants. They also found that deceivers averted their gaze more during deception than while telling the truth. A different group of researchers sought to use eye tracking technology as a potential sensor within an automated screening paradigm (Derrick et al., 2010 2010). Applying the Guilty Knowledge Test in an eye tracking environment and based on the memory assessments of participants, the study asked some participants to make a mock explosive device and sought to investigate if the gaze behavior of those who built the device would be different from those who did not build it when the device was altered. The results revealed that the participants that built the mock bomb gazed longer at the altered part of the device. In this study, the eye tracking technology helped to correctly classify all those who knew information about the “proper” image of the mock explosive device from those who did not.

One of the topics in deception detection research is whether there are experts in deception detection. While Bond and DePaulo suggested there were no differences between lay people and experts in accurately differentiating lies from truth (Bond and DePaulo, 2008), others found some empirical evidence that experts did better than lay people (O'Sullivan and Ekman, 2004). Bond used eye tracking to find out whether experts did better and what was the nature of their detection behavior (Bond, 2008). For the study, he recruited

both students (lay people) and law enforcement personnel (experts). Those individuals who scored 80% accuracy rate or higher in detecting truth and deception were invited for the second study, and this time their eye movements were recorded and analyzed with eye tracking. Only two people continued to differentiate accurately above 80% in the second study. Interestingly, the gaze fixation data showed that while one expert focused on face areas, the other expert looked more at arm/torso areas when making decisions.

In another study that applied eye tracking technology to detect indicators of deception, researchers investigated senders' lookup patterns and pupil dilation in sender-receiver games that involve truth telling and deception (Wang et al., 2010 2010). They used eye tracking data as a supplement to economic analysis of choices. While senders were not instructed to deceive the receivers, the senders had an incentive to exaggerate the truth. The authors used pupil dilation as an indicator of deception based on the premise that such dilations are associated with stress and cognitive load. In another study, a group of researchers investigated pupil responses and reading behavior of participants when they were lying vs when they were honest (Cook et al., 2012). The deceivers in this study had increased pupil responses and took less time to read deceptive statements than honest statements.

As mentioned in our discussion about the eye tracking methodology, this approach gives an objective insight into what veracity judges look at to determine the sender's credibility. It allows the researcher to designate areas of interest and quantitatively compare fixation durations and frequencies of fixations on such areas and to analyze foci shifts between these areas. Moreover, researches may observe other behavioral cues that judges look at but fail to mention when providing the reasons for their assessment. Eye tracking

methods are starting to be adopted in deception detection research and promise to generate additional insight into this topic.

## **Research Method**

### **Pilot study**

Seven participants (1 female), ranging in age from 19 to 35, participated in the pilot experiment. Participants were recruited through an oral announcement made in an MIS class offered in a business college at a large Midwestern university. During the recruitment, the principal investigator explained the purpose and the general overview of the study. The compensation was \$10 per participant. The participants in the pilot study watched 6 video snippets, half honest and the other half dishonest. Half of the videos had sound, and the other half did not. Two stimulus sets, comprised of six unique snippets, were created (Table 2-2). Participants were randomly assigned to one of two sets. Both stimulus sets included all six snippets; however, snippets in sets varied by whether or not they had sound. As shown in Table 2-2, those snippets that had sound in one set were presented in the other set without sound. The order in which snippets were presented was randomized to minimize order or learning effects. This design allowed the comparison of honest snippets with dishonest, as well as snippets with sound with those that did not have sound.

Table 2-2. Distribution of video snippets in two stimulus sets for pilot study

	<b>Set A</b>	<b>Set B</b>
Snippet 1	Dishonest, No Audio	Dishonest, Audio
Snippet 2	Dishonest, No Audio	Dishonest, Audio
Snippet 3	Dishonest, Audio	Dishonest, No Audio
Snippet 4	Honest, No Audio	Honest, Audio
Snippet 5	Honest, Audio	Honest, No Audio
Snippet 6	Honest, Audio	Honest, No Audio

After a successful completion of the pilot study, the researchers decided to slightly enhance the design of the study (explained below) and proceeded to the main study.

### Main Study Procedure

Participants were recruited through verbal announcements made in multiple classes offered at the college of business. During the announcements, the principal investigator explained the purpose of the study and addressed the questions the audience asked. In addition to providing the instructions on signing up for the experiment, the announcement indicated that the participants would be paid \$10 as compensation. To ease the process, the participants were instructed to sign up for the study through a web link, which indicated all possible time slots and dates the study sessions were offered. Those who signed up were emailed the day before their scheduled appointment and reminded of their appointment. The email included information about the location of the lab and a reminder not to wear mascara or artificial eye lashes. Forty eight university students were recruited in total, however, for various reasons (e.g., too long eye lashes, eye shapes, pupil size, etc.) the eye movements of 6 participants could not be calibrated. These prospective participants were paid \$10 and were let go from the study. Table 2-3 depicts information about the participants who remained in the study.

Table 2-3. Demographic information

		Frequency	Percent
<b>Gender</b>	Male	28	66.7%
	Female	14	33.3%
		<b>42</b>	<b>100.0%</b>
<b>Age</b>	18-24	38	90.5%
	25-34	4	9.5%
		<b>42</b>	<b>100.0%</b>



The research experiment was a 2 x 2 repeated measures, within and between subjects design, varying experimental condition (deception and truth) and medium (audio-visual and video only). Each experimental session was completed in a single event with one participant per session. The experimental sessions were designed to last no more than 1 hour, and most sessions were completed in 30 - 40 minutes.

Participants began by completing an informed consent form. Next, they were instructed about the overall procedure. After a researcher answered any questions participants had, a participant was seated in front of a 22" monitor with an eye-tracking system mounted underneath, providing unobtrusive eye-tracking. The eye-tracking system used in the study was SensoMotoric Instruments (SMI) RED 250 eye tracking system, which records eye data at 250 Hz frequency (i.e., 250 times per second) and allows the capture of participants' eye movements and gaze in remote, contact-free setup, while allowing for free head movement. The system allows the capture of eye movement data even with participants wearing glasses or contact lenses.

Each eye tracking session started with gaze calibration and subsequent validation of calibration. To calibrate gazes properly, this process involved adjustments of participants' chair forward/backwards or upwards/downwards to find the best position for the device to capture the event. In the adjoining room, the principal investigator ran the iViewX software, which runs the calibration, validation, and recording of the experimental session. All validation results of the calibration sessions were in the desired range.

The experimental session was comprised of two parts: a practice task and the main experimental task. The practice task was designed to make sure participants were comfortable with operating the keyboard commands, instructions, and understood the task.

After the researcher confirmed a participant did not have any questions and understood the task fully, the participant proceeded to complete the main task. Each participant was randomly assigned to one of the six stimuli sets described below. Operating the eye tracking software and running all experimental sessions and instructing participants was performed by the principal investigator. All study sessions were run by the same person. For both practice and main tasks, the participants were told that there was no time limit for completing the tasks.

The practice task and the main experimental task involved asking students to watch a number of video snippets. The snippets were recorded interviews of undergraduate students and were originally created for a different study. The interviewees in the videos had willingly enhanced a scholarship application and were then asked to defend the embellished application in an interview. Deception was determined by comparing interviewees' actual résumés with enhanced applications. The interviewees are truthful in two of the four selected videos, and deceptive in the other two.

The resulting set of video stimuli included two females (one black female) and two males (both white). The people in the honest stimuli were the white female and a white male, and the people who delivered the deceptive message were the black female and white male. To eliminate the possibility that the veracity judges might assess the credibility of the people in the video stimuli based on the sender's gender or race, a one-way between subjects analysis of variance was conducted. The analysis compared the effect of "actors" on the veracity assessment in honest and deceptive conditions. There was a significant effect of actor on veracity assessment at the  $p < .05$  level for the two conditions [ $F(1, 164) = 11.448$ ,  $p < .001$ ]. Post-hoc tests indicate no significant difference between honest senders, female and

male, both white (mean difference=0.238, se=0.332, p=1.0), between honest (white female) and deceptive (black female) senders (mean difference=0.571, se=0.332, p=0.523), and between honest (white male) and deceptive (black female) senders (mean difference=0.810, se=0.332, p=0.095). The significant mean differences were between the deceptive (white male) sender and all other senders in video stimuli. He was assessed as the least honest among the four senders. The results of our tests indicate that the veracity judges did not demonstrate any racial or gender bias in their assessment.

We created four additional new video snippets from these videos by removing sound from them. Of these eight video snippets, we created six stimulus sets shown in Table 2-5. The arrangement in stimuli sets is organized in such a way that each study participant should see four videos, half with sound and half without sound, half with truthful and the other half with deceptive statements. Moreover, participants should not be assigned to treatments whereby they see the same interviewee with sound and later without sound, i.e., they should assess each interviewee only once. While 16 variations of four stimulus combinations sets are possible to arrange (Table 2-4), our design necessitated that each set include exactly two honest and exactly two dishonest videos, while exactly two of those had sound and two did not.

Table 2-4. Stimulus sets selected

Clip	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Honest Video1	S	N	S	N	S	N	S	N	S	N	S	N	S	N	S	N
Honest Video2	S	S	N	N	S	S	N	N	S	S	N	N	S	S	N	N
Deception Video1	S	S	S	S	N	N	N	N	S	S	S	S	N	N	N	N
Deception Video2	S	S	S	S	S	S	S	S	N	N	N	N	N	N	N	N

This arrangement resulted in the incomplete block design with randomized order of treatments at each stimuli set.

Table 2-5. Arrangement of Treatments per Stimuli Set

<b>Stimulus set</b>	<b>Sound</b>		<b>No Sound</b>	
1	Video1	Video2	Video3*	Video4*
2	Video1	Video3*	Video2	Video4*
3	Video1	Video4*	Video2	Video3*
4	Video3*	Video4*	Video1	Video2
5	Video2	Video4*	Video1	Video3*
6	Video2	Video3*	Video1	Video4*

\* videos with deception

After each video, participants were asked to judge if the speaker in the video was honest or deceptive. They chose their answer on a 7-point Likert scale, with “1” being "completely dishonest" and “7” being "completely honest." Next, they were asked to explain what they saw in the video that made them conclude the way they indicated a step earlier. They typed their responses into the dialog box. After finishing typing their responses, they clicked on the “submit” button to advance to the next screen. An instruction display appeared before each video that asked them to press the space bar to proceed to the next video. All participants completed these steps for each video snippet they observed. The presentation order of video snippets in the main task was randomized. Video snippets in the practice task were different from those in the main task. Each of our experimental manipulations, honest and deceptive messages, as well as with and without sound, were presented through multiple instances of a stimulus category or stimulus sampling. The need for stimulus sampling exists whenever a dependent variable may be affected as a result of a particular instance of a stimulus (Wells and Windschitl, 1999). After completing both tasks, they were asked to fill

out a demographics questionnaire and were debriefed about the study. Each participant was paid \$10.

Arrangement of treatments in six stimuli sets with one participant per set results in 24 observations, with the distribution shown in Table 2-6. Our dataset represents observations from forty two participants, with 7 participants per stimulus set, and 21 observations per cell.

Table 2-6. Distribution of Observations per Treatment for 6 Stimuli Sets

	<b>Clip</b>	<b>Sound</b>	<b>No Sound</b>
Truthful Message	1	3	3
	2	3	3
Deceptive Message	3	3	3
	4	3	3

## Measures

The participant's assessments of each video snippet on a 7-point Likert scale were collapsed into three categories: scores from 1 to 3 were treated as dishonest, 5 to 7 as honest, and 4 as neutral. Each judge was asked to respond to 4 snippets, resulting in 168 total responses. Of these responses, only twenty two were identified as neutral. These responses were excluded from the analysis, leaving the total response count at 146. The dependent variable, detection accuracy, is a dichotomous variable, wherein participants correctly identify honest messages as honest and deceptive messages as lies. Moreover, using the results from the eye tracking data, we identified three areas of interest (AOI): 1) speaker's eyes, 2) speaker's mouth, and 3) speaker's torso (see Figure 2-1). Gaze behavior in this study was measured in terms of fixation duration and frequencies of fixations on those AOIs. The fixation duration was measured in milliseconds for every time a participant looked within an AOI. In the discussion of the viewing behavior for targeted areas, Djambasi (2014) mentions

fixation duration as an effective method. Fixation frequencies on a specific target area are regularly used as metric for eye behavior and cognitive processing (Cook et al., 2012; Rayner, 1998; Rayner et al., 2012). Frequency of fixation was measured by the number of fixations a participant generated in the area-of-interest. The fixation maps on prespecified areas allowed the comparison and contrasting of regions quantitatively.



Figure 2-1. Assigned Areas of Interest

A one-way between subjects ANOVA was conducted to compare the effect of actor on the fixation duration and frequency on three AOIs (eyes, mouth, and torso). The results of the analyses suggest there was no significant effect of actor on fixation duration on the senders' eyes [ $F(3, 164) = 1.177, p=0.320$ ] and fixation frequency on eyes [ $F(3, 164) =$

1.334,  $p=0.265$ ]. There was a significant effect of actor on fixation duration on the senders' mouth [ $F(3,164) = 5.025, p=0.002$ ] and fixation frequency on the senders' mouth [ $F(3, 164) = 5.222, p=0.002$ ]. Post-hoc tests indicate that the judges looked longer into the mouth of a deceptive sender (white male) (mean difference=5048.71 ms,  $se=1325.04, p=0.001$ ) compared to the mouth of honest sender (white male). The judges also fixated more frequently into the mouth of deceptive senders (white male) (mean difference=5.095,  $se=1.776, p=0.028$ ) and black female (mean difference=6.548,  $se=1.776, p=0.002$ ) compared to the fixation frequency into the mouth of an honest sender (white male). Last, there was a significant effect of actor on fixation duration on the senders' torso [ $F(3,164) = 12.181, p<0.001$ ] and fixation frequency on the senders' torso [ $F(3, 164) = 16.729, p<0.001$ ]. Post-hoc tests indicate that the judges fixated longer on the torso of deceptive senders compared to the torso of honest senders. There was no significant mean difference in fixation duration onto the senders' torso within the honest or deceptive conditions. There was a similar pattern of fixation frequency behavior on the senders' torso: the significant differences were observed between conditions, with no significant difference within conditions. As the results suggest, there was no viable indication that any of the areas of interest invited fixation behavior from the veracity judges due to the factors other than conditions of the experiment.

### **Analysis and Results**

To analyze the relationships between media, message, and attention foci, the data were analyzed using repeated measured mixed model in SPSS 24. After removing the neutral responses, 146 were included in the analysis. The first step was to run the analysis of the random variable "actor" or the person appearing in the video and see if there was a collinearity between the actor and one of our main variables (Deceptive and Audio). This was done to make sure that our manipulations were attributed to the changes in the response

variable and that we did not have any confounding of our manipulations by the over/under presence of a specific actor. Table 2-7 shows the results of the covariance parameters for the random variable actor. No statistical effects of an actor variable were observed, and it was therefore excluded from the analysis reported below.

Table 2-7. Covariance estimates for variable Actor

Name	Estimate	Significance
Eyes	51453438.72	.168
Mouth	55671430.42	.164
Torso	3777849.94	.174

The same analysis indicated that the covariance parameters for four trials were not equal, and thus running Repeated Measures ANOVA was not suitable, as RM ANOVA assumes equal variances and covariances of residuals. Thus running MIXED method assuming heterogeneous variance was a preferred course (assumptions of variances and covariances are more relaxed in the MIXED method).

When we ran the analysis including the actor variable as a random effect, and the two other factors (message and media) as fixed effects, we saw that its covariance estimate was not significant. Running the actor variable alone, we saw no significant effect of this variable, although its significance level significantly increased. This implies that there is some portion of variance due to the actor variable, but it is not statistically significant.

To test the first two hypotheses, we analyzed the effect of the two independent variables on (1) the fixation duration and (2) fixation frequency on three focal areas of a speaker in video snippets: eyes, mouth, and the torso.



### Fixation duration on the speaker's eyes

Tables 2-8 and 2-9 show the estimates and the mean differences of estimates, respectively, of the fixation durations. There was no statistically significant main effect of a message type ( $p=0.151$ ) or media ( $p=0.396$ ), nor was there any significant effect of their interaction ( $p=0.667$ ) on fixation duration on the speaker's eyes.

Table 2-8. Estimates of means of fixation durations

AOI	IV	Condition	Mean	Std. Error	df	95% CI	
						Lower Bound	Upper Bound
Eyes	Message	Deception	7969.838	805.957	133.888	6375.784	9563.892
		Truth	6334.633	797.319	136.08	4757.894	7911.371
	Audio	No Sound	6669.466	812.389	138.013	5063.128	8275.805
		Sound	7635.004	790.764	137.897	6071.413	9198.595
Mouth	Message	Deception	8664.873	678.891	138.691	7322.559	10007.188
		Truth	5905.53	661.814	132.17	4596.412	7214.648
	Audio	No Sound	8472.936	681.604	140.364	7125.399	9820.474
		Sound	6097.467	659.019	137.134	4794.313	7400.621
Torso	Message	Deception	2682.024	274.39	140.013	2139.541	3224.506
		Truth	730.636	258.911	114.205	217.746	1243.526
	Audio	No Sound	1699.249	270.994	133.021	1163.233	2235.265
		Sound	1713.411	262.462	128.287	1194.096	2232.726

Table 2-9. Estimates of mean differences of fixation durations

AOI	IV	Mean Diff.	Std. Error	df	Sig	95% CI	
						Lower Bound	Upper Bound
Eyes	Message	1635.205	1133.704	138.231	0.151	-606.438	3876.848
	Audio	-965.538	1133.704	138.231	0.396	-3207.181	1276.105
Mouth	Message	2759.343	948.099	139.022	0.004	884.786	4633.9
	Audio	2375.469	948.099	139.022	0.013	500.912	4250.026
Torso	Message	1951.388	377.259	131.035	0.000	1205.081	2697.694
	Audio	-14.162	377.259	131.035	0.970	-760.468	732.145

### Fixation frequency on the speaker's eyes

There is a statistically significant main effect of media on the number of fixations on the eyes ( $p=0.023$ ). Judges looked more frequently into the eyes of the speakers when the sound was present (mean diff = 4.454,  $se=1.939$ ) see Tables 2-10 and 2-11. There was no statistically significant main effect of the message ( $p=0.287$ ), nor was there a significant interaction effect ( $p=0.084$ ).

Table 2-10. Estimates of means of fixation frequencies

AOI	IV	Condition	Mean	Std. Error	df	95% CI	
						Lower Bound	Upper Bound
Eyes	Message	Deception	14.739	1.379	136.936	12.013	17.465
		Truth	12.665	1.364	138.218	9.967	15.362
	Audio	No Sound	11.475	1.391	141.164	8.724	14.225
		Sound	15.929	1.351	140.039	13.257	18.6
Mouth	Message	Deception	13.804	0.977	138.658	11.873	15.736
		Truth	9.078	0.941	125.529	7.215	10.942
	Audio	No Sound	11.939	0.974	136.695	10.014	13.865
		Sound	10.943	0.945	134.738	9.075	12.812
Torso	Message	Deception	9.348	0.782	139.43	7.802	10.894
		Truth	2.39	0.751	124.022	0.902	3.877
	Audio	No Sound	5.585	0.777	136.043	4.048	7.122
		Sound	6.152	0.756	135.003	4.657	7.648

Table 2-11. Estimates of mean differences of fixation frequencies

AOI	IV	Mean Diff.	Std. Error	df	Sig	95% CI	
						Lower Bound	Upper Bound
Eyes	Message	2.074	1.939	140.903	0.287	-1.76	5.908
	Audio	-4.454	1.939	140.903	0.023	-8.288	-0.62
Mouth	Message	4.726	1.357	136.046	0.001	2.043	7.409
	Audio	0.996	1.357	136.046	0.464	-1.687	3.679
Torso	Message	6.958	1.084	135.887	0.000	4.813	9.103
	Audio	-0.567	1.084	135.887	0.602	-2.712	1.577

### **Fixation duration on the speaker's mouth**

There is a statistically significant main effect of both independent variables on the gaze fixation on the sender's mouth. There is however no significant interaction effect ( $p=0.80$ ). The results on Tables 8 and 9 suggest that the judges looked longer at the mouth of the speakers when lying (mean=8664.873 ms, se=678.891) than when telling the truth (mean=5905.530 ms, se=661.814), and this difference is statistically significant ( $p=0.004$ ). Similarly, the veracity judges fixated longer on the speaker's mouth when no sound was available (mean=8472.936 ms, se=681.604), and, on average, fixated less (6094.467 ms, se=659.019) when the sound was available. The mean difference (2375.469 ms, se=948.099) is statistically significant ( $p=0.013$ ).

### **Fixation frequency on the speaker's mouth**

The analysis of the effect of the two independent variables on the fixation count suggests no significant main effect of media ( $p=0.464$ ) and interaction ( $p=0.621$ ). There is, however, a significant main effect of media on the frequency of fixations ( $p=0.001$ ): the judges looked more frequently at the mouths of speakers (mean=13.804, se=0.977) when lying than when telling the truth (mean=9.078, se=0.941).

### **Fixation duration on the speaker's torso**

As Tables 2-8 and 2-9 suggest, there is evidence of the significant main effect of the message veracity on the duration the veracity judges fixated on the torso of the senders ( $p<.001$ ). The results of the marginal means suggest that the judges on average fixated longer on the torso (mean=2682.024 ms, se=274.39) when the speaker was lying, and fixated less on average (mean= 730.636 ms, se=258.911), when the speaker was honest. The mean difference (1951.388 ms, se=377.259) is statistically significant ( $p<0.001$ ). There was no

significant main effect of audio on the fixation duration on the speaker's torso ( $p=0.97$ ), and similarly, there was no statistically significant interaction effect ( $p=0.528$ ).

### **Fixation frequency on the speaker's torso**

Similar to the fixation duration results, there is a statistically significant main effect of message on the frequency of fixations for torso ( $p<0.001$ ). The veracity judges looked more frequently at the torso of the speakers when lying (mean=9.348, se=0.782) compared with telling the truth (mean=2.39, se=0.751). There effects of media and the interaction of message and media were not significant ( $p=0.602$  and  $p=0.331$ , respectively).

These findings provide support for Hypotheses 1 and 2. Hypothesis 1 predicted judges would look at different things, depending on whether the message was honest or dishonest. While there were no differences for the eyes, there were differences for the mouth. Judges looked longer at the mouth for liars than truth-tellers, and judges looked at the mouth more often for liars. There was a similar pattern for the torso: Judges looked longer and more often at the torso for liars, compared to truth-tellers. H1 is partially supported. Hypothesis 2 predicted judges would focus their gaze on different things, depending on whether the message had sound or not. Judges looked more frequently at the eyes of senders when there was sound, compared to when the sound was muted. Judges looked longer at the mouth of senders when there was no sound, compared to when it was present. There was no significant difference in the fixation duration and fixation count patterns of judges when looking at the torso of the sender under varying media modes. H2 is partially supported.

### **Detection Accuracy**

Of the 146 responses we collected, 98 were correctly assessed, resulting in 67.1% total accuracy rate. A closer look at the detection rate shows that the judges correctly

identified 38 of 72 lies, or a 52.8% hit rate, and 60 out of 74 truths, or 81.1% true negative rate. Overall, this performance is better than the average detection rates of 54% mentioned earlier in this study. The relationship between the attention foci and detection accuracy was analyzed using repeated measures logistic regression, because the dependent variable, detection accuracy, is a dichotomous variable. On SPSS, the GENLIN command was used, with a binomial distribution and logit as the link function.

With the fixation frequency on three areas of interest (eyes, mouth, and torso) as the independent variable, the results of the logistics regression show that both fixation count on eyes ( $X^2(1, N=146) = 5.837, p=0.016$ ) and the mouth ( $X^2(1, N=146) = 7.162, p=0.007$ ) of the speaker were statistically significant (Table 2-12). The effect of fixation count on torso was not statistically significant ( $p=0.499$ ).

Table 2-12. Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)
			Lower	Upper	Wald Chi-Square	df	Sig.	
(Intercept)	2.021	0.4592	1.121	2.921	19.371	1	0.000	7.547
Fixation Count.eyes	-0.039	0.0163	-0.071	-0.007	5.837	1	0.016	0.961
Fixation Count.mouth	-0.054	0.0203	-0.094	-0.015	7.162	1	0.007	0.947
Fixation Count.torso	-0.015	0.0221	-0.058	0.028	0.457	1	0.499	0.985

The parameter estimates of the logistic regression suggest the log of the odds ratio. Negative numbers decrease the odds of accurate assessment, whereas positive numbers increase the odds. The exponentiated coefficients tell us about the effect on the odds ratio, i.e., the factor by which the odds ratio changes. Numbers less than 1 decrease the odds, numbers greater than 1 increase the odds. Hence the interpretation of an exponentiated

coefficient of 0.961 for fixation count on eyes and the coefficient of 0.910 for fixation count on mouth suggest that we can predict that fixating on those areas of the speaker decreases the odds of successful detection by a factor of 0.961 and 0.947, respectively. Given the variate, the baseline scenario, where fixation count on eyes, mouth, and torso is set to zero will result in 0.88 estimated detection rate (95% CI (0.7541, 0.9488)). Setting the fixation count to zero implies that the judges do not look at those areas at all, which is similar to talking over the phone. Looking at the eyes and the mouth, based on our data, has deteriorated the detection ability of the judges.

With the fixation duration on three areas of interest as the independent variable, the accuracy of classifying correctly truth and lies was not impacted significantly. The main effects of fixation durations on the eyes ( $p=0.677$ ), the mouth ( $p=0.989$ ), and the torso ( $p=0.307$ ) were not statistically significant. Hypothesis 3 predicted judges' veracity assessment would differ, depending on what the judges looked at. While there were no differences in detection success rate based on fixation duration on three AOIs, there were differences for the fixation frequency. Judges who looked more frequently at the eyes and the mouth of the senders performed worse in deception detection than judges who looked less frequently. There was no main effect for looking at the torso of the sender on detection accuracy rate. H3 is partially supported.

### **Analysis of the research model**

Next, we sought to analyze the research model we presented in Chapter 1. In order to analyze the mediating effect of the fixation areas on the detection accuracy, we split the dataset into four scenarios: honest messages with sound, honest messages without sound, deceptive messages with sound, and deceptive messages without sound. This was necessitated by the complex nature of the study design, which included two independent

variables and interaction of those variables, repeated measures within subjects design, and a dichotomous dependent variable.

### **Honest messages with sound**

Of the 150 responses, 39 were in this category. Four of the respondents were exposed to this manipulation twice. Only 3 out of 39 responses were inaccurate, resulting in 92.3% truth accuracy rate. The logistic regression analysis results suggest that fixation duration on the torso of the speaker increases the likelihood of accurate detection by a factor of 1.002 ( $p=0.017$ ). There is no effect on detection accuracy of focusing duration on eyes ( $p=0.15$ ) or on the mouth ( $p=0.518$ ) of the sender. With the fixation frequency on the three AOIs as the IV, there was no significant effect on detection accuracy.

### **Honest message without sound**

Thirty five responses were in this category. Six of the participants were exposed to this manipulation twice. The detection accuracy with no sound was worse than when sound was available; the total accuracy rate dropped from 92.3% when the sound was available to 68.6% with the video only snippets. There was a significant effect for audio,  $F(1, 65.781) = 13.242$ ,  $p=0.001$ . The fixation duration on the three AOIs did not have any significant effect on detection accuracy. Nor did the logistic regression analysis with fixation frequency on the three AOIs have any significant effect on detection accuracy.

### **Deceptive message with sound**

Of the 36 responses in this category, 5 participants were exposed to the manipulation twice. Only 18 out of 36 responses were accurate, resulting in a lie detection success rate of 50%. With the fixation duration on three AOIs as the IV, the logistic regression analysis results suggest no significant effect of fixating on eyes ( $p=0.569$ ), mouth ( $p=0.103$ ), and

torso ( $p=0.208$ ). Similarly, with the fixation frequency on the same three AOIs as the IV, the logistic regression analysis did not indicate any significant effect.

### **Deceptive message without sound**

There were 36 responses in this category and 6 participants were exposed to this manipulation twice. Interestingly, for deceptive messages when the sound is removed, the detection accuracy rate goes up. Twenty out of 36 responses were correctly identified as lies, resulting in 55.6% lie accuracy rate vs 50% when the sound was available, yet there is no significant main effect of audio,  $F(1, 69.984) = 0.222, p=0.939$ . The logistic regression with fixation duration on three AOIs as the IV did not reveal any significant effects of fixation duration on eyes, mouth, and torso on detection accuracy. However, when fixation frequency is on three AOIs is analyzed using the logistic regression, results suggest that fixation frequency on the mouth of the speaker decreases the likelihood of accurate detection by a factor of 0.929 ( $p=0.036$ ). There is no effect on detection accuracy of focusing frequency on eyes ( $p=0.244$ ) or a torso ( $p=0.444$ ) of the sender.

### **Discussion**

This research was guided by three research questions. We sought to understand how, if at all, does media play a role in drawing the visual attention of veracity judges. We also investigated what the judges look at when presented with honest and dishonest messages. Finally, we sought to understand the link between what the judges looked at and how successful they were in correctly classifying the messages. In this study, applying eye tracking technology, and varying the audio availability of the video stimuli, we found that both media and the nature of the message have an effect on where people tend to focus. We also found very slight support for the association between the visual foci and detection success.



The multi-factor, leakage, and interpersonal deception theories propose that the behavior of deceivers should differ from that of the truth tellers. This difference is signaled in behavioral and verbal cues, which could be spotted by the receiver. We investigated the gaze behavior of the judges on three areas, namely, the senders' eyes, mouth, and torso. While the judges fixated longer and more frequently on the senders' eyes, for those who were lying as contrasted with those who were telling the truth, the mean differences in fixation duration and fixation counts were not statistically significant. One explanation for such behavior could result from the beliefs that the judges hold. In a study to investigate the beliefs about deception, a team of scholars from around the world led by Bond conducted a study and interviewed people from multiple countries (Team, 2006). The most popular belief by far around the world, indicated by sixty four percent of respondents, was that gaze aversion indicated dishonesty. The belief in gaze aversion as an indicator of deception may help explain why our participants attended to the senders' eyes during their assessment of interview snippets and why we did not see any significant difference across the message type. If a judge believes that the liar would avert his/her gaze more often, they will naturally tend to focus on the eyes of the speakers and look for the signs of the gaze aversion. Naturally, the duration and the frequency of fixation on the eyes will be longer since that is the primary "base" area. Any nonverbal or verbal behavior displayed by the sender's "non eye" region might trigger the receiver's attention and result in a gaze focus shift. After inspecting the area, the eyes of the receivers "regress" naturally to the base area. When the judges tend to focus on the sender's eyes a lot, naturally, it is hard to see any significant difference in gaze fixation duration and frequency regardless whether the sender was deceptive or honest.

Deceptive messages resulted in longer and more frequent fixations on the sender's mouth. In their meta-analysis of cues to deception, DePaulo and her colleagues found that the liars tend to be less forthcoming, less pleasant, more tense, tell less compelling lies, and include fewer ordinary imperfections and unusual content than truth tellers (DePaulo et al., 2003). Behaviors in liars related to the mouth area such as lips being pressed more frequently and less facial pleasantness were found to be statistically significant. While we did not code the nonverbal behaviors of the senders in stimuli, it is interesting that the mouth area of the liars drew judges' attention. Our findings suggest that visual cues around the mouth are good indicators of deception. Further research might look into more detailed inspection of the behaviors and investigate their relationship to deception.

The judges looked more frequently and longer at the torso of the liars, compared to the torsos of truth-tellers. Although DePaulo and colleagues did not report a link between torso movements and lying in their meta-analysis, prior research has found mixed results. Gross and Levenson found that liars suppress the movement of their trunks when concealing their emotions (Gross and Levenson, 1993). In another study, researchers found deceptive responses marked by decreases in smiling and increases in self-manipulations and postural shifts (McClintock and Hunt, 1975). Vrij reported a few studies where researchers found no difference in trunk movements between liars and truth tellers (Vrij, 2008). As mentioned earlier, the results of the global beliefs about the cues to deception revealed that more than 25% of the respondents associated lying with more frequent body movements (Mann et al., 2004; Team, 2006). Posture shifts and trunk movements seem to attract judges' attention and future research may look into details of the torso movements to investigate whether more nuanced movements are associated with deception. In his discussion on reasons for not

finding consistent and reliable nonverbal cues to deception, Vrij (2008) suggests that it could possibly be due to inadequate scoring systems. Instead of categorizing a cue as a general movement, thus lumping together different types of movement behavior, Vrij recommends to differentiate between specific types of movements. Given the findings for the mouth and torso, and how fixations differ for them between honest and dishonest messages, we found partial support for Hypothesis 1, demonstrating how gaze behavior varied as a result of the message veracity.

As predicted, the veracity judges focused on different areas of the stimuli when the media varied. Specifically, they looked longer at the mouth of the speakers when the sound was absent from the image, and they looked more frequently into the eyes of the speakers when presented with full audio-visual stimuli. Under “normal” media conditions, where sound is available, the judges looked more frequently at the eyes of the senders. As discussed before this might have been triggered by the general beliefs that the veracity judges hold, leading them to focus on the eyes of the senders, as the eyes are believed to be the primary indicators of deception. Yet, under “unusual” media conditions, when sound is not present, the judges seem to have compensated for the lack of audio symbols by fixating their gaze longer on the mouths of the speakers to determine the content. The different levels of symbol set transmission proposed by MST resulted in judges’ foci shift from the primary area (eyes) to the mouth. There was no significant effect of the media factor on the fixation behavior on the senders’ torso. This is interesting as bodily movements are detectable under both full audio-visual and video only modes, and the judges seem to only have reacted when people were lying by fixating longer on their torso area. Overall, we found support for Hypothesis 2 and provided empirical evidence of the media mode effect on the gaze behavior. In a study

investigating the effect of cultural differences on the deception detection performance, George and his colleagues (2018) found that people from different cultures can successfully detect lies when told by people of other cultures. In their experiments, they had judges from the US, India, and Spain, and the judges from all three countries had to assess not only senders from their own culture, but from the two other cultures as well. Thus, an American judge, for example, had to watch, listen, and read statements produced in American English, Indian English and Spanish. How could a US judge assess the veracity of a Spaniard if the American judge did not understand Spanish? The researchers suggested the cues to deception are similar across cultures and universal in their nature, hence providing an equivalent detection rate. When a medium transmits an image and no sound, it is similar, in a way, to the detection process of a person of a different culture. Even when deprived of the paralinguistic cues, the judges seem to compensate by shifting the foci to the mouth of the speaker.

When we used fixation duration on three AOIs as the independent variable to test the link between what judges look at and how that affects their detection accuracy (RQ 3), we found no significant effect. The analysis of frequency of fixations on those AOIs, however, indicated that looking at the eyes and the mouth of the sender significantly diminished the performance of the judges. There was no significant effect of looking at the torso of the receiver on detection accuracy. These findings are in line with the claims of leakage theory, that looking into the face of the sender diminishes detection accuracy (Ekman and Friesen, 1969). We found partial support for Hypothesis 3. Buller and Burgoon (1996) suggest that facial cues are less informative, whereas vocal cues are more informative. Despite the commonly held belief, the research has shown that gaze behavior is not related to deception

(Vrij, 2008), and by focusing on the eyes of the senders, the judges seem to have missed out on other cues.

Finally, the analysis of the research model, introduced in Chapter 1, did not find significant support for the mediating role of the visual attention. In this model, we did not investigate the direct effect of message veracity and media on detection accuracy. Visual attention is a complex topic, affected by both psychological and physiological factors (Duchowski, 2007). Human visual attention is driven by both, the low-level visual features of the scene observed, as well as the higher-level intentional factors, such as cognitive processes navigating the visual focus. Our interpretation of the elements we see, however, does not have to be unanimous. Viewers may attribute completely opposite meanings to the same stimuli they observe. For example, prominence interpretation theory (Fogg, 2003) suggests that when assessing the credibility of a website, users must first see the elements of the website and then interpret that element to come to a conclusion. While gaze aversion could be interpreted as a sign of deception, some judges may interpret it as an attempt to remember the details of the situation being told. Thus, when presented with the task to judge the veracity of the speaker, the judges may pick up visual behaviors, and yet their final assessment will eventually be driven by the interpretations they apply to the elements they have noticed.

### **Implications**

This study contributes both to practice and research. From the research perspective, we have looked deeper into the processes of the veracity assessment by the receiver of the communication. Specifically, we were able to analyze the gaze behavior of the judges and investigate the areas they looked at more often and longer. Our findings provide support for the multi-factor, leakage, and IDT theories. The behavior of liars and truth tellers does seem

to vary, and this variance get noticed by the judges, which is evidenced by the gaze pattern and fixation duration of the judges. To the best of our knowledge, no study to date has looked into the gaze behavior of the veracity judges and studied it under varying media modes and message veracity. Our findings support the postulates of the MST and provide an insight into how variance in symbol set capability of the communication modes effect the behavior of the receiver. In their research of media and cultural differences and deception detection, George and his colleagues (2018) provide a theoretical basis for combining cultural components, media theories, and deception theories into one unified model. Similarly, this research combines media theories with deception theories and provides a framework for testing the effect of media on deception detection.

Increasingly our daily communication, both professional and personal, is taking place through some mediated communication mode. The value afforded by various communication channels continues to result in wide adoption of these technologies. Deception that used to take place in face-to-face communication seems to be migrating into the digital environment. Understanding the effect of the media capabilities on identifying deception is thus becoming increasingly important. Our findings suggest that focusing visual attention on the facial area of the sender diminishes the detection accuracy. While generalizability of our findings is limited, the findings nevertheless inform us of the areas to be cautious of when making veracity judgements. In organizational settings, for example, people in charge of hiring may conduct their interviews through televised modes of communication. These people should be aware of the media capabilities and limitations in transmitting cues.

This study provides a glimpse at what eye tracking technology may provide in the study of deception detection. This study has not looked into the analysis of the written

content provided by the participants. That data will help us discover any discrepancies between the stated observations and more objective eye tracking behavior. Furthermore, a visual analysis of the eye scanpath for participants may reveal a rich insight into what participants tend to focus on first and how their gaze varies as the conversation proceeds. How do the scanpaths of those assessing video only snippets differ from those assessing full audio-visual stimuli? More importantly, we need to better understand the link between the veracity of the message, media and detection accuracy. While we found main effects for varying media and message veracity, our findings of the relationship between the visual foci and deception detection are not consistent. Those are the questions that need to be investigated in future research.

### **Limitations and future research**

This study is not void of limitations. We used student groups from a Midwestern state university as the participants of our study. While students in many experiments are seen as a limitation, we believe that for the purpose of our study, students are not a problem. We were looking at what people look in varying media when they are being told lies or truth, and because students are a part of general population, our findings are not of less importance. Generalizability of our findings, while somewhat limited, is applicable to a broader population. The biggest criticism of deception detection studies comes from the use of sanctioned lies, where recruited participants, usually students, are asked to either lie or tell the truth, and their statements are recorded and studied. In this study we used the stimuli materials from another study. Students had applied for a scholarship and had willingly enhanced their resumes. They were later told that the scholarship announcement was made as a part of a study and were asked to defend their enhanced resumes. Interviews obtained in laboratory studies have certain limitations. In cases when lying is sanctioned and no

consequences are expected, liars are unlikely to display strong emotions of guilt or fear. In real life scenarios, however, they would have had to respond to the questions by the interviewers.

We operationalized the visual attention of the participants by the frequency and duration of gaze fixations devoted to certain areas of interest. Prior research into eye movement behavior has pointed to a link between the foveal vision and attention. Attention, however, is not always associated with overt eye behavior, and study participants may voluntarily disengage their foveal vision from attention. We cannot measure this covert mechanism of attention and instead assume that the participants' foveal focus is aligned with their attention and thus measure only overt attention.

Our research has demonstrated the merits of using eye tracking in studying deception under varying media. While we obtained valuable insight into the processes of visual attention of veracity judges, we believe that future work should look at these processes in an interactive communication where the sender engages in conversation synchronously. Eye tracking technology may help us understand not only what the judges look at, but also at what the senders look at when attempting to deceive, and more importantly, how their gaze foci shifts to the receiver's reactions. Further research may investigate the reaction of senders to the receivers probing and whether receivers' visual attention changes when probing.

### References

- Abdelhamid, N., Ayesh, A., and Thabtah, F. 2014. "Phishing Detection Based Associative Classification Data Mining," *Expert Systems with Applications* (41:13), pp. 5948-5959.
- APWG. 2018. "Phishing Activity Trends Report."
- Bernard, H. R., Killworth, P., Kronenfeld, D., and Sailer, L. 1984. "The Problem of Informant Accuracy: The Validity of Retrospective Data," *Annual review of anthropology*), pp. 495-517.



Bond, C. F., and DePaulo, B. M. 2006. "Accuracy of Deception Judgments," *Personality and social psychology Review* (10:3), pp. 214-234.

Bond, C. F., and DePaulo, B. M. 2008. "Individual Differences in Judging Deception: Accuracy and Bias," *Psychological bulletin* (134:4), p. 477.

Bond, G. D. 2008. "Deception Detection Expertise," *Law and Human Behavior* (32:4), pp. 339-351.

Bond Jr, C. F., and Atoum, A. O. 2000. "International Deception," *Personality and Social Psychology Bulletin* (26:3), pp. 385-395.

Buller, D. B., and Burgoon, J. K. 1996. "Interpersonal Deception Theory," *Communication theory* (6:3), pp. 203-242.

Burgoon, J. K., Blair, J. P., and Strom, R. E. 2008. "Cognitive Biases and Nonverbal Cue Availability in Detecting Deception," *Human Communication Research* (34:4), pp. 572-599.

Burgoon, J. K., Chen, F., and Twitchell, D. P. 2010. "Deception and Its Detection under Synchronous and Asynchronous Computer-Mediated Communication," *Group Decision and Negotiation* (19:4), pp. 345-366.

Carlson, J. R., George, J. F., Burgoon, J. K., Adkins, M., and White, C. H. 2004. "Deception in Computer-Mediated Communication," *Group decision and negotiation* (13:1), pp. 5-28.

Cook, A. E., Hacker, D. J., Webb, A. K., Osher, D., Kristjansson, S. D., Woltz, D. J., and Kircher, J. C. 2012. "Lyin'eyes: Ocular-Motor Measures of Reading Reveal Deception," *Journal of Experimental Psychology: Applied* (18:3), p. 301.

Cyr, D., and Head, M. 2013. "The Impact of Task Framing and Viewing Timing on User Website Perceptions and Viewing Behavior," *International Journal of Human-Computer Studies* (71:12), pp. 1089-1102.

Daft, R. L., and Lengel, R. H. 1986. "Organizational Information Requirements, Media Richness and Structural Design," *Management science* (32:5), pp. 554-571.

Dennis, A. R., Fuller, R. M., and Valacich, J. S. 2008. "Media, Tasks, and Communication Processes: A Theory of Media Synchronicity," *MIS quarterly* (32:3), pp. 575-600.

Dennis, A. R., and Kinney, S. T. 1998. "Testing Media Richness Theory in the New Media: The Effects of Cues, Feedback, and Task Equivocality," *Information systems research* (9:3), pp. 256-274.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., and Epstein, J. A. 1996. "Lying in Everyday Life," *Journal of personality and social psychology* (70:5), p. 979.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., and Cooper, H. 2003. "Cues to Deception," *Psychological bulletin* (129:1), p. 74.

DePaulo, B. M., and Rosenthal, R. 1979. "Telling Lies," *Journal of personality and social psychology* (37:10), p. 1713.

Derrick, D. C., Moffitt, K., and Nunamaker Jr, J. F. 2010. "Eye Gaze Behavior as a Guilty Knowledge Test: Initial Exploration for Use in Automated, Kiosk-Based Screening," *Proceedings of the Hawaii International Conference on System Sciences*, Poipu, HI.

Djamasbi, S. 2014. "Eye Tracking and Web Experience," *AIS Transactions on Human-Computer Interaction* (6:2), pp. 37-54.

Duchowski, A. 2007. *Eye Tracking Methodology: Theory and Practice*. Springer Science & Business Media.

Duchowski, A. T. 2003. "Eye Tracking Methodology: Theory and Practice,").

Dunbar, N. E., Jensen, M. L., Burgoon, J. K., Kelley, K. M., Harrison, K. J., Adame, B. J., and Bernard, D. R. 2015. "Effects of Veracity, Modality, and Sanctioning on Credibility Assessment During Mediated and Unmediated Interviews," *Communication Research* (42:5), pp. 649-674.

Ekman, P., and Friesen, W. V. 1969. "Nonverbal Leakage and Clues to Deception," *Psychiatry* (32:1), pp. 88-106.

Elaad, E. 2003. "Effects of Feedback on the Overestimated Capacity to Detect Lies and the Underestimated Ability to Tell Lies," *Applied Cognitive Psychology* (17:3), pp. 349-363.

Elaad, E. 2014. "Covert Detection of Deception," *Detecting Deception: Current Challenges and Cognitive Approaches*, p. 315.

Eysenbach, G., and Köhler, C. 2002. "How Do Consumers Search for and Appraise Health Information on the World Wide Web? Qualitative Study Using Focus Groups, Usability Tests, and in-Depth Interviews," *Bmj* (324:7337), pp. 573-577.

Fisher, A. 2014. "Resume Lies Are on the Rise," in: *Fortune.com*.

Fogg, B. J. 2003. "Prominence-Interpretation Theory: Explaining How People Assess Credibility Online," *CHI'03 extended abstracts on human factors in computing systems: ACM*, pp. 722-723.

George, J. F., Carlson, J. R., and Valacich, J. S. 2013. "Media Selection as a Strategic Component of Communication," *MIS Quarterly* (37:4), pp. 1233-1251.

George, J. F., Giordano, G., and Tilley, P. A. 2016. "Website Credibility and Deceiver Credibility: Expanding Prominence-Interpretation Theory," *Computers in Human Behavior*

(54), pp. 83-93.

George, J. F., Gupta, M., Giordano, G., Mills, A. M., Tennant, V. M., and Lewis, C. C. 2018. "The Effects of Communication Media and Culture on Deception Detection Accuracy," *MIS Quarterly* (42:2), pp. 551-A559.

George, J. F., Marett, K., and Tilley, P. A. 2008. "The Effects of Warnings, Computer-Based Media, and Probing Activity on Successful Lie Detection," *IEEE Transactions on Professional Communication* (51:1), pp. 1-17.

George, J. F., and Robb, A. 2008. "Deception and Computer-Mediated Communication in Daily Life," *Communication Reports* (21:2), pp. 92-103.

George, J. F., Tilley, P., and Giordano, G. 2014. "Sender Credibility and Deception Detection," *Computers in Human Behavior* (35), pp. 1-11.

Granhag, P. A., Vrij, A., and Verschuere, B. 2015. *Detecting Deception: Current Challenges and Cognitive Approaches*. John Wiley & Sons.

Gross, J. J., and Levenson, R. W. 1993. "Emotional Suppression: Physiology, Self-Report, and Expressive Behavior," *Journal of Personality and social Psychology* (64:6), p. 970.

Guillory, J., and Hancock, J. T. 2012. "The Effect of LinkedIn on Deception in Resumes," *Cyberpsychology, Behavior, and Social Networking* (15:3), pp. 135-140.

Hancock, J. T. 2007. "Digital Deception," *Oxford handbook of internet psychology*, pp. 289-301.

Hancock, J. T., Thom-Santelli, J., and Ritchie, T. 2004. "Deception and Design: The Impact of Communication Technology on Lying Behavior," *Proceedings of the SIGCHI conference on Human factors in computing systems: ACM*, pp. 129-134.

Hancock, J. T., Woodworth, M. T., and Goorha, S. 2010. "See No Evil: The Effect of Communication Medium and Motivation on Deception Detection," *Group Decision and Negotiation* (19:4), pp. 327-343.

Hartwig, M., and Bond, C. F. 2011. "Why Do Lie-Catchers Fail? A Lens Model Meta-Analysis of Human Lie Judgments," *Psychological bulletin* (137:4), p. 643.

Jacob, R., and Karn, K. S. 2003. "Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises," *Mind* (2:3), p. 4.

Levine, T. R., Park, H. S., and McCornack, S. A. 1999. "Accuracy in Detecting Truths and Lies: Documenting the "Veracity Effect", *Communications Monographs* (66:2), pp. 125-144.

Levine, T. R., Serota, K. B., Shulman, H., Clare, D. D., Park, H. S., Shaw, A. S., Shim, J. C., and Lee, J. H. 2011. "Sender Demeanor: Individual Differences in Sender Believability Have a Powerful Impact on Deception Detection Judgments," *Human Communication Research* (37:3), pp. 377-403.

Lewis, C. C. 2009. *To Catch a Liar: A Cross-Cultural Comparison of Computer-Mediated Deceptive Communication*. The Florida State University.

Mann, S., Vrij, A., and Bull, R. 2004. "Detecting True Lies: Police Officers' Ability to Detect Suspects' Lies," *Journal of applied psychology* (89:1), p. 137.

McClintock, C. C., and Hunt, R. G. 1975. "Nonverbal Indicators of Affect and Deception in an Interview Setting," *Journal of Applied Social Psychology* (5:1), pp. 54-67.

McCornack, S. A., and Levine, T. R. 1990. "When Lovers Become Leery: The Relationship between Suspicion and Accuracy in Detecting Deception," *Communications Monographs* (57:3), pp. 219-230.

McHaney, R., George, J. F., and Gupta, M. 2017. "An Exploration of Deception Detection: Are Groups More Effective Than Individuals?," *Communication Research* (0:0), p. 0093650215607627.

Miller, G. R., and Stiff, J. B. 1993. *Deceptive Communication*. Sage Publications, Inc.

O'Sullivan, M., and Ekman, P. 2004. "12 the Wizards of Deception Detection," *The detection of deception in forensic contexts*, p. 269.

Pak, J., and Zhou, L. 2013. "Eye Gazing Behaviors in Online Deception,").

Poole, A., and Ball, L. J. 2006. "Eye Tracking in Hci and Usability Research," *Encyclopedia of human computer interaction* (1), pp. 211-219.

Rayner, K. 1998. "Eye Movements in Reading and Information Processing: 20 Years of Research," *Psychological bulletin* (124:3), p. 372.

Rayner, K., Pollatsek, A., Ashby, J., and Clifton Jr, C. 2012. *Psychology of Reading*. Psychology Press.

Salvucci, D. D. 1999. "Mapping Eye Movements to Cognitive Processes." Carnegie Mellon University.

Strömwall, L. A., Granhag, P. A., and Hartwig, M. 2004. "Practitioners' Beliefs About Deception,").

Sun Park, H., and Levine, T. 2001. "A Probability Model of Accuracy in Deception

Detection Experiments," *Communication Monographs* (68:2), pp. 201-210.

Team, G. D. R. 2006. "A World of Lies," *Journal of Cross-Cultural Psychology* (37:1), pp. 60-74.

Vicianova, M. 2015. "Historical Techniques of Lie Detection," *Europe's Journal of Psychology* (11:3), p. 522.

Vrij, A. 2000. *Detecting Lies and Deceit: The Psychology of Lying and Implications for Professional Practice*. Wiley.

Vrij, A. 2007. "Deception: A Social Lubricant and a Selfish Act," *Social communication*, pp. 309-342.

Vrij, A. 2008. *Detecting Lies and Deceit: Pitfalls and Opportunities*. John Wiley & Sons.

Wang, J. T.-y., Spezio, M., and Camerer, C. F. 2010. "Pinocchio's Pupil: Using Eyetracking and Pupil Dilation to Understand Truth Telling and Deception in Sender-Receiver Games," *The American Economic Review* (100:3), pp. 984-1007.

Wells, G. L., and Windschitl, P. D. 1999. "Stimulus Sampling and Social Psychological Experimentation," *Personality and Social Psychology Bulletin* (25:9), pp. 1115-1125.

Zuckerman, M., DePaulo, B. M., and Rosenthal, R. 1981. "Verbal and Nonverbal Communication of Deception," *Advances in experimental social psychology* (14), pp. 1-59.

## Appendix. IRB Approval

**IOWA STATE UNIVERSITY**  
OF SCIENCE AND TECHNOLOGY

Institutional Review Board  
Office for Responsible Research  
Vice President for Research  
1138 Pearson Hall  
Ames, Iowa 50011-2207  
515 294-4566  
FAX 515 294-4267

**Date:** 3/15/2016

**To:** Akmal Mirsadikov  
3235 Gerdin Business Bldg

**CC:** Dr. Joey F George  
2340 Gerdin Business Bldg

**From:** Office for Responsible Research

**Title:** Determining Indicators of Deception with Eye Tracking

**IRB ID:** 16-069

**Study Review Date:** 3/15/2016

The project referenced above has been declared exempt from the requirements of the human subject protections regulations as described in 45 CFR 46.101(b) because it meets the following federal requirements for exemption:

- (2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey or interview procedures with adults or observation of public behavior where
  - Information obtained is recorded in such a manner that human subjects cannot be identified directly or through identifiers linked to the subjects; or
  - Any disclosure of the human subjects' responses outside the research could not reasonably place the subject at risk of criminal or civil liability or be damaging to their financial standing, employability, or reputation.

The determination of exemption means that:

- **You do not need to submit an application for annual continuing review.**
- **You must carry out the research as described in the IRB application.** Review by IRB staff is required prior to implementing modifications that may change the exempt status of the research. In general, review is required for any modifications to the research procedures (e.g., method of data collection, nature or scope of information to be collected, changes in confidentiality measures, etc.), modifications that result in the inclusion of participants from vulnerable populations, and/or any change that may increase the risk or discomfort to participants. Changes to key personnel must also be approved. The purpose of review is to determine if the project still meets the federal criteria for exemption.

Non-exempt research is subject to many regulatory requirements that must be addressed prior to implementation of the study. Conducting non-exempt research without IRB review and approval may constitute non-compliance with federal regulations and/or academic misconduct according to ISU policy.

**Detailed information about requirements for submission of modifications can be found on the Exempt Study Modification Form.** A Personnel Change Form may be submitted when the only modification involves changes in study staff. If it is determined that exemption is no longer warranted, then an Application for Approval of Research Involving Humans Form will need to be submitted and approved before proceeding with data collection.

Please note that you must submit all research involving human participants for review. **Only the IRB or designees may make the determination of exemption**, even if you conduct a study in the future that is exactly like this study.

Please be aware that **approval from other entities may also be needed.** For example, access to data from private records (e.g. student, medical, or employment records, etc.) that are protected by FERPA, HIPAA, or other confidentiality policies requires permission from the holders of those records. Similarly, for research conducted in institutions other than ISU (e.g., schools, other colleges or universities, medical facilities, companies, etc.), investigators must obtain permission from the institution(s) as required by their policies. **An IRB determination of exemption in no way implies or guarantees that permission from these other entities will be granted.**

Please don't hesitate to contact us if you have questions or concerns at 515-294-4566 or IRB@iastate.edu.

**CHAPTER 3. DETECTING DECEPTION IN TEXT: WHAT YOUR EYES SEE**

Akmal Mirsadikov/Iowa State University

Modified from a manuscript to be submitted to MIS Quarterly

**Abstract**

This study investigates linguistic cues to deception in written real-life, high-stake accounts using eye tracking. Building on the theories of deception and reading, we developed hypotheses on the relationship between the message veracity, linguistic cues to deception, reading patterns, and detection accuracy. The reading behavior of veracity judges varied across honest and deceptive statements. We also found a support for the effect of visual foci on detection performance. In particular, fixation durations on generalizing words, on first-person pronouns, and fixation frequencies on passive voice verbs worsened detection accuracy. Fixation durations on passive voice verbs, third-person pronouns, and fixation frequencies on first-person pronouns, and the total fixation counts improved detection rates.

**Introduction**

Technological breakthroughs in computer mediated communication (CMC) have led to the emergence of a variety of channels for people to communicate with other individuals and groups. The wide adoption of these media has led to a surge in content creation in a variety of formats, e.g., video, audio, image, text, or a mix of these. This content of ever accumulating communication is generally stored and is made accessible through forums, chats rooms, social media and other means. The surge in information content and absence of monitoring of this content by centralized authoritative institutions has raised issues with verifying the credibility and accuracy of such information and hence the risks of acting upon it (Metzger, 2007).

The 2016 US presidential election has turned the spotlight on the issue of “fake news” (Darcy, 2016) and the lack of mechanisms for individuals to verify the accuracy of numerous posts online (Itkowitz, 2016), including those posted by candidates running for official positions. The presence of deceptive information in generated content, both in CMC and in face-to-face communication, raises issues mainly because people are not very good at detecting deception (Bond and DePaulo, 2006). Research on detecting deception in communication has investigated numerous verbal and non-verbal behaviors indicative of deception, referred to as cues to deception. To date, none of the cues has been identified to be always associated with lying (Vrij, 2008; Zuckerman, DePaulo, & Rosenthal, 1981). There are, however, a small number of reliable cues that have been empirically shown to be associated with lying some of the time (DePaulo et al., 2003).

A lot of information today is still presented in text form, and one of the basic ways in which people communicate with each other or obtain information online remains the written form, e.g., e-mails, instant text messaging, SMS texting, online forums and others. Since deception is prevalent in our daily communication and in CMC (DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996; George and Robb, 2008), we can expect text-based deception to be prevalent in many places: in media, social networks, online product reviews with exaggerated negative or positive reviews, online dating profiles, online resumes and other settings.

Because written statements are visual displays, analyzing this display from the point of view of a reader provides a unique opportunity into how the reader perceives the statements. Eye tracking technology enables researchers to look into what readers looked at, how long they looked at it, what word(s) they skipped, if any, and provides information about the scan path of the pattern of eye movements (Duchowski, 2007). In fact, analyzing the eye



movement recordings for studying cognitive processes associated with reading remains the dominant method in research into reading (Rayner, Pollatsek, Ashby, & Clifton Jr, 2012). Eye tracking has received wide adoption in reading research as it allows the inference of moment to moment cognitive processes during reading (Rayner, 1998, 2009).

Research on deception and its detection in written text has identified and empirically validated some linguistic-based cues (Burgoon, Blair, Qin, & Nunamaker, 2003; Fuller, Biros, Burgoon, & Nunamaker, 2013; Hancock, Curry, Goorha, & Woodworth, 2007; Hauch, Blandón-Gitlin, Masip, & Sporer, 2015; Zhou, Burgoon, Nunamaker, & Twitchell, 2004a). These studies used automated methods to derive deception from written text and proposed frameworks of classifying text-based cues into clusters. While automated methods afford a fast and structured approach to detecting deception in text, they have some limitations. The cues to detecting deception in text are limited to those that could be programmed, e.g., count the number of verbs, count the number of pronouns, compute a ratio of adjectives over total number of words, and so on. However, the automated approach cannot identify logical consistency or realism of the statement. Skilled readers can identify words that have different meanings in different contexts or can identify words that do not belong to a certain context. A person reading an account of a murder scene by a witness would expect to see some expressions of shock, fear, or other strong emotions. Observing the eye movement behavior from the point of the viewer provides an objective insight into what draws a reader's attention and what the reader chooses to focus on the most. Research on deception in text using eye tracking is very scant (Cook et al., 2012), and no study, to the best of our knowledge, has applied eye tracking technology to investigate deception in written text generated from real accounts of offenders and witnesses. The eye tracking technology can

help better understand what the participants actually look at when assessing the veracity of the written statement, the pattern of their eye movements, the words they chose to fixate their gaze on, and many other metrics not obtainable in traditional methods of participants' self-reports.

This study seeks to investigate what the veracity judges look at when assessing the veracity of written statements made in high-stakes context with real-life negative consequences. A lot of research into deception involves experiments, where the participants are asked to produce a deceptive statement, or sanctioned lies, with little to non-existent negative consequences. Investigating text-based cues to lies produced in real, high-stakes settings from the point of view of the observers allows us not only to investigate how the message veracity in a written text affects the deception detection accuracy but also to compare our findings with previous research results and, possibly, identify new cues. This exploratory research would also help to understand whether the constructs suggested for automated deception detection methods are applicable for humans reading the text and whether noticing the previously suggested cues improves the accuracy of detecting deception. Thus, the research questions leading this study are:

- What do veracity judges look at when assessing the veracity of a text statement in a high-stakes setting?
- What is the relationship between what veracity judges choose to focus on and their detection accuracy?

To investigate our research questions, we propose using eye tracking technology to analyze what our participants look at when reading written statements produced in a high-stakes domain. It will help us better understand what the study participants look at and how

the things they look at affect their detection success rate. The next section of the paper reviews the literature on reading, eye tracking, and deception, and derives hypotheses. In the methods section we discuss the design of our study and the procedures. Then we present our analysis and findings. We conclude our study with a discussion of our findings.

## **Literature review**

### **Human visual system**

The human sensory system perceives objects and the environment around it. Our ears hear sounds, we sense smell with our nose, and our eyes see things in our surroundings. The process of seeing starts with the light that reflects off an object, travels through our eyes, and falls on the sensory neurons in the retina before it gets processed in the brain (Djamasbi, 2014; Rayner et al., 2012). The retina of human eyes is composed of two types of receptors: rods and cones. Rods are active in low light environments (night vision), and cones are active in well-lit environments (daylight vision) (Duchowski, 2007). While the retina has far more rods than cones, the center of the retina, or fovea, features mainly cones. The cones are specialized for processing detail and acuity, and therefore the fovea displays the external world in much higher resolution than any other part of the retina (Djamasbi, 2014; Rayner et al., 2012). The rods are specialized for detecting movement and registering variability in brightness.

The space in which we can see objects without moving our eyes is referred to as visual field or as a visual perceptual span (Duchowski, 2007; Rayner et al., 2012). Due to visual acuity limitations, not all elements are clearly seen on the visual field. In terms of acuity, the visual field can be divided into three regions: foveal, parafoveal, and peripheral (Rayner et al., 2012). The sharpest and most colorful vision type is foveal, but it only covers about 2% of the visual field, or put differently, it only subtends about 2 degrees of visual

angle around the fixation point (Duchowski, 2007; Rayner et al., 2012). The parafoveal vision area subtends about 10 degrees of visual angle around fixation, and anything else on the visual field beyond the parafoveal area is covered by the peripheral area. Visual acuity drops sharply between these regions; it's highest in the foveal area and lowest in peripheral area. Therefore, visual acuity is associated with only fovea.

While our sensory system may pick up a lot of information about objects around us, our capacity to apprehend all information in detail is limited. Our attention allows us to tune in our mental capacities on selections of sensory inputs so that our mind can process the object of interest more effectively (Duchowski, 2007). When we first look at a particular image stimulus, some areas of this image may attract our attention. At this stage, our visual attention is not fully engaged, and by moving our eyes to the first object of interest, we apply our foveal vision to bring that area under fine detail and higher resolution. While we apply foveal vision to inspect items in fine detail, objects in the periphery of our focus are identified by our parafoveal vision, which in turn dictates the location of the next focus of attention. This “dual” or parallel coworking of foveal and parafoveal visions is at the core of the visual attention process (Duchowski, 2007). While visual attention is an essential mechanism of visual perception, people may divert their attention to objects in a peripheral view without bringing those objects to their visual focus. Thus, eye movements made to see the objects of interest in fine detail reflect the overt visual attention of an individual.

### **Eye tracking**

Eye tracking is a technique that allows the researcher to measure a user's eye movements and determine where and for how long that user looked (Poole and Ball, 2006). This technique allows collecting objective data about the user's eye movements without having to rely on his or her verbal responses. Technological advancement and improvements

in systems recoding eye movements allowed measurements to be more accurate and non-obtrusive, and the decline in cost of these devices allowed for their adoption by researchers in various disciplines (Rayner, 1998). While eye movements can be recorded in many ways, the most widely used method today is the video-based corneal reflection method. It involves the eye tracking device projecting near infrared light into the eye, which creates a corneal reflection, or glint (Duchowski, 2007). The image processing software then captures the information about the positioning of the glint along with the positioning of the center of the pupil. This captured information is then processed and calculated by the software using trigonometric calculations (Djamasbi, 2014), which in turn enables the determination of an accurate position of the viewer's gaze on a stimulus. The eye tracking equipment can only track overt movements of the eyes, and thus one of the main assumptions of the research involving eye trackers and visual attention is that attention is linked to foveal gaze direction (Duchowski, 2007).

We move our eyes across the visible field of view and bring a portion of this field into high resolution to be able to see in better detail the objects in the central direction of our gaze. Our control of eye movements and a choice to focus our gaze onto a certain region is driven generally by our need to divert our attention. Therefore, the data generated by the eye tracking method not only tells about the pattern of eye movements on a certain region but also informs about an individual's visual attention and perception. Most theories in cognition assume attention is required for many cognitive acts; because attention is limited, it affects how much information can be processed at a time (Rayner et al., 2012).

### **Reading**

Reading is defined as the ability to extract visual information from the page and comprehend the meaning of the text (Rayner et al., 2012). Research into the process of

reading and eye movements in reading dates back more than a century. Huey's work (1908), for example, covered a wide range of research on reading to that date, including word recognition, perceptual span, and eye movements in reading. A contemporary of Huey, Emile Javal, noticed the non-smooth pattern of eye movements during reading consisting of series of quick jumps, or saccades, and relatively stable fixations over a region of interest (Rayner et al., 2012). Rayner and his colleagues (2012) wrote an extensive work on the process of reading and discussed in detail how readers go about extracting information from the printed page and comprehending the text. They focused on the cognitive processes associated with reading and analyzed reading from cognitive psychology and information-processing points of view.

The information-processing approach views reading as a highly complex process relying on a number of sub-processes. Rayner and his colleagues (2012) attempted to investigate the complex process of reading by examining its sub-processes. The critics of this approach have argued that isolating processes of reading, such as studies of word recognition, result in tasks that are very different from real reading. In word recognition studies, participants are administered a series of single words for a very short duration of time and are asked to either pronounce the word or identify whether it belongs to a certain category (Rayner et al., 2012).

While the activity of reading may seem straightforward, experts who study the reading process discriminate between different types of reading. Here we focus on silent, skilled reading of English intended to comprehend, which excludes reading involving skimming or skipping. Currently, the dominant technique for studying cognitive processes associated with reading is analyzing eye movement recordings (Rayner et al., 2012). Eye

tracking has been widely adopted in reading research as it allows the researchers to infer moment to moment cognitive processes during reading (Rayner, 1998, 2009). The process of reading requires identifying letters and words in fine detail, and thus we must move our eyes to direct our fovea over the word we want to read.

A cumulative stream of research into reading has identified that, in skilled reading, word identification is a straightforward, automatic process which takes about a quarter of a second on average (Rayner, 1998). The word recognition does not occur serially, i.e., by processing letter-by-letter. Instead, the letters in common short words are processed in parallel, and even letters in long words are not processed sequentially. While the context may affect the speed at which words are processed, it generally has very little effect on how words are processed, i.e., a word in isolation is processed the same way as a word in a text (Rayner et al., 2012). The most cognitively challenging parts of reading are higher order processes that are related to putting the meanings of words together and extracting the meaning of sentences and paragraphs (Rayner et al., 2012).

### **Eye movements in reading**

While most people may believe that eyes move smoothly across the text when reading, research into eye movements suggests a different story. When reading English text, eyes move from left to right on a row while continually coming to rest for periods between 150 and 500 ms. These relative stable gaze movements are referred to as fixations. Eyes also move rapidly between fixations (between 20-35 ms), and these rapid movements are called saccades after the French term for jump (Duchowski, 2007; Rayner et al., 2012). An average length of a saccade is 7-9 character spaces in silent reading. Visual information is extracted and processed only during gaze fixations. No information is processed during saccades, which is referred to as saccadic suppression. Since fixations are associated with information

extraction of a display, researchers are interested in analyzing the durations and frequencies of fixations and how they affect information processing during fixations. The purpose of saccades is to move fixations from one region to another, and thus, researchers are mainly interested in studying the links between the directions, the size of saccades, and what is being processed (Rayner et al., 2012). Because saccades are motor movements that require time to plan and execute, there is a latency period associated with a saccade (Rayner, 1998). The duration of saccade latency is at least 150-175 ms (Abrams and Jonides, 1988).

Thus when reading, the pattern of information extraction consists of an interplay between fixations and saccades: eyes fixate on a region of interest for about 250 ms, followed by a saccade, then fixate on another word, followed by a saccade, and so on. This pattern of information extraction is not unique to reading activity only. Perception of any still display follows a similar pattern, albeit with different durations of fixations and saccades.

There are other types of eye movements unique to reading. One of them is called regressive saccades or regressions. While most saccades move forward (from left to right) when reading, about 10-15% of them move backward, and we, on average, end up making a regression every two seconds when reading (Rayner et al., 2012). Some regressions are more salient, as we want to go back a few words on a text to make sure we understood the line properly, but we are mostly unaware of regressions when they are only subtle moves, i.e., going back a few letter characters. Another eye movement unique to reading is a return sweep, which is a move from the end of the line to the beginning of the next line. Return sweeps start about five to seven character spaces from the end of a line and land on about the third to seventh character space on the next line (Rayner et al., 2012). This results in the



leftmost fixation on the second word of the line. In fact, about 80% of the line falls between the extreme fixations.

Because of the visual acuity discussed earlier, a reader would be expected to bring all words into the foveal vision and process each word in fine detail. While most words are fixated on during reading, many words are skipped, and hence foveal processing of each word is not necessary (Rayner, 1998). Content words (i.e., nouns, verbs, adverbs, and adjectives) are fixated on most of the time (85%), while a little over one third of function words (i.e., prepositions (in, on, at, etc.), auxiliaries (are, was, do, etc.), and pronouns (he, we, etc.) are fixated on (Rayner and Duffy, 1988). Thus, there is a direct relationship between the length of a word and the probability of fixating on it: the longer the word, the higher the chance it will be fixated on (Rayner, 1998). Shorter words are skipped more than longer words (Brysbaert, Drieghe, & Vitu, 2005), but seven- to eight-letter words are fixated on most of the time (Rayner and McConkie, 1976).

There are two major factors that control where we move eyes during reading. Low-level visual information processed by parafoveal vision, such as the position of spaces, dictates where to move the eyes on the next fixation. At the same time, higher-level variables such as the meaning of the text controls how long we fixate on a word. The research has shown that lexical processing drives eyes through the text, evidenced by the duration of fixation on words of varying frequency: if a word is a low-frequency word, the eyes fixate on it longer than on high-frequency word (Rayner et al., 2012).

The duration of fixation on each word varies. Researchers use various metrics for measuring processing time in reading. While single fixation duration is the most frequently used metric, different metrics could be applied for instances with multiple fixations on a

word. For example, researchers can use duration of the first fixation, or gaze duration, which includes all fixation times on a word before the eyes move off, and finally total viewing time, which includes all fixations on a word, including later fixation caused by regressions.

Through the experiments in the eye-contingent display paradigm, researchers monitored the eye movements while changing the visual displays based on where the eyes looked and where they moved to (Rayner, 1998). The perceptual span of a reader to be able to read and comprehend a text normally is 31 letter characters wide, or 15 letter spaces to each side of fixation (McConkie and Rayner, 1975). The perceptual span of a reader is influenced by the limitations of visual acuity. For readers of English, the perceptual span is asymmetric, where the readers depend heavily on the information on the right side of the fixating point to retain the normal reading speed (McConkie and Rayner, 1976). An experiment with a moving window for reading showed that when the window displayed only the fixated word, the speed of reading was only 200 words per minute versus 330 words per minute with no window. (Rayner, Well, Pollatsek, & Bertera, 1982). This is also referred to as preview benefit. Readers get the most preview benefit for the word they want to move next ( $n+1$ ) rather than from  $n+2$  (Angele and Rayner, 2011).

Perceptual span does not imply that words can be identified up to 15 characters to the right of the fixation point. Readers obtain information about the length of a word further to the right of the perceptual span, while they can acquire information about some letters to the right of the fixation. Word length information is used to help the reader to navigate in reading by determining where to fixate next. The word identification span is an area from which words can be identified in a given fixation, and it is shorter (7-8 letter spaces) than the total perceptual span (McConkie and Zola, 1987; Rayner et al., 1982). While readers can identify

words that they don't fixate on, they can rarely identify content words beyond the fixated word. Partial information about a word in a parafoveal area could be encoded on one fixation and used to help in identification of the word on the subsequent fixation.

### **Deception**

Deception is a message “knowingly transmitted by a sender to foster a false belief or conclusion by the receiver” (Buller and Burgoon, 1996). Research on deception detection is built on the premise that the behavior of liars is different from that of truth tellers. When a deceiver delivers a deceptive message, he or she is expected to display some behavior, verbal or non-verbal, which may give away his or her intentions (Zuckerman et al., 1981). These indicators, or cues to deception, have been studied extensively to identify more reliable ones that are strongly associated with lying (DePaulo et al., 2003). To date, however, after almost five decades of studying deception, no cue has been found to be always associated with lying and with lying only (Vrij, 2008).

Deception is prevalent in daily life - we lie about one or two times a day, with most lies being non-trivial in nature (DePaulo et al., 1996; Vrij, 2008). Despite such prevalence of lies in our daily lives, people are poor detectors of deception. In fact, research suggests we can detect about 54% of lies (Bond and DePaulo, 2006), which suggests that our performance is slightly better than chance. Furthermore, empirical evidence suggests that the performance of professional lie catchers (e.g., border patrol officers, police officers, judges, and law enforcement personnel) is no different from that of lay people (Bond and DePaulo, 2008). In investigating cues to deception, researchers mostly conduct experiments, and participants are asked to lie or to detect a lie. One of the criticisms of the research into deception is that most lies that participants make are sanctioned lies, and participants have little to no liability or negative consequences for lying. Nevertheless, this stream of research

helped to identify many cues to deception and investigate those vigorously. One of the reasons we lack research with real scenario, high-stakes lies, where the consequence could be dire and dangerous for the deceiver, is that it is hard to establish the ground truth – knowing with certainty whether the reported statement is true or false.

Theories of deception have suggested many reasons as to why the behavior of liars would be expected to deviate from those telling the truth. A multi-factor framework proposes that telling a lie is associated with different emotions such as guilt, fear, or delight (Ekman, 2009; Zuckerman et al., 1981), cognitive load, and attempted behavioral control (Buller and Burgoon, 1996; Vrij, 2008). Ekman and Friesen's work on leakage theory posits that lying is associated with strong emotions, and that the deceiver must attempt to control his or her behavior not to give away any cues of a deception. At the same time, the deceiver must control the message and make sure it is plausible and does not contradict known information, which adds to the mental load of the deceiver. The deceiver's performance to keep the message intact and control his or her behavior may display inconsistencies and result in inferior behavior, which may betray or "leak" deception (Ekman and Friesen, 1969). Building on the discussion of information management by the deceiver, Buller and Burgoon investigated the interplay between the deceiver and the receiver in Interpersonal Deception Theory, or IDT (Buller and Burgoon, 1996). They stressed the role of interpersonal communication processes and proposed an interactive and dynamic process of deception, whereby the deceiver's message and its delivery is affected by the feedback the deceiver gets from the receiver. IDT was introduced as a theory of interactivity between communicators and thus the temporal and spacial presence of communicators is vital for the perception of immediacy. To achieve the goal of deception, deceivers act strategically and engage in

information management, which is the dimension of IDT most relevant to our discussion of linguistic cues. Information management refers to the manipulation of the content of the deceptive message and its style to foster the message credibility. In managing the deceptive message, the deceiver may alter the message dimensions, such as completeness, veridicality, relevance, and personalization. By controlling the verbal content and the linguistic style, the deceiver may convey uncertainty and vagueness, display reticence by opting to utter minimal information, express nonimmediacy by verbally distancing himself from the message, and indicate insincerity.

DePaulo and her colleagues offered a self-presentation perspective into studying verbal and non-verbal cues to deception (DePaulo et al., 2003). Unlike other theories of deception mentioned above, this view suggests that the behaviors of the deceivers are not a totally unconscious reflection of their inner emotions and states but rather reflect self-presentation goals to appear credible and convey truthful impression. The researchers suggest that truth tellers, too, engage in self-presentation and also attempt to appear credible. The self-presentation perspective, thus, points out the similarities between the truth tellers and deceivers and suggests that they have much in common. The difference between the two, however, is that the deceivers' claim to honesty is illegitimate, and this might lead to two implications. First, the liars may embrace the deceptive self-presentation less than do truth tellers. Second, because they take their credibility less for granted than truth tellers, they experience a greater sense of awareness and deliberateness in their performance. These implications may result in liars telling less compelling tales, being less positive and pleasant, less forthcoming, more tense, and including fewer ordinary imperfection and unusual contents. The majority of the cues analyzed in DePaulo and her colleagues' meta-analysis is

based on research that studied deception in face-to-face communication; however, some verbal cues are also applicable to the text-based deception. Of the cues related to liars being less forthcoming, the meta-analysis suggested that liars' responses are significantly shorter and less detailed. Of the cues where liars tell less compelling tales, they found that liars provide less plausible accounts, which are less consistent and coherent, provide fewer personal experiences and make efforts to distance themselves from their own words or avoid taking responsibility for or claiming ownership of the message. The liars tend to repeat words and phrases. Deceivers are less pleasant than truth tellers and provide more negative statements and complaints. In an attempt to appear more credible and less nervous, liars avoid making corrections to their statements or admitting they have forgotten something.

Research on verbal cues to deception investigated linguistic cues through verbal veracity assessment tools. Two most commonly used such tools are Statement Validity Analysis (SVA) and Reality monitoring (RM). Introduced as a tool to assess the credibility of child witness statements in sexual-abuse cases (Köhnken, 2004), SVA focuses on variables associated with truthfulness of an account. An essential element of SVA, Criteria-Based Content Analysis (CBCA) is based on the hypothesis put forth initially by Undeutsch that statements based on memory of a witness are different in content features from a fabricated statement (Undeutsch, 1989). Undeutsch came up with a list of features of statement content that were less likely to occur in a fabricated statement and called them "credibility criteria." This list was later modified and combined with other features and included 19 reality criteria (Steller and Köhnken, 1989). The CBCA involves two processes: evaluation of a statement with regard to the CBCA criteria and comparing the results of the content analysis to a reference. The main principle of RM is built around differentiating between the memory

characteristics of actually experienced events and imagined events (Vrij, 2008). RM framework argues that the memories of real experiences are likely to contain sensory information (i.e. details of seeing, hearing, smelling, tasting, and touching), contextual information (i.e. spatial and temporal details of the event), and affective information (i.e. details of feelings) while the imagined events are more likely to contain references to cognitive operations. The propositions put forth by CBCA and RM have been tested numerously, achieving 70% accuracy rate for each approach (Vrij, 2015).

The accounts written by deceivers should deviate from those written by truth tellers, manifested in the stylistic and lexical language and reflected through the text-based cues discussed next. The presence of cues, if detected, should draw the attention of veracity judges, which will affect their visual attention, eye movement behavior, and hence their reading behavior. Next, we will discuss some of the linguistic cues rooted in the theoretical frameworks explained above and form study hypotheses that reflect the effect of the linguistic cues on reading behavior and eye movement metrics.

### **Text-Based Cues to Deception**

Prior research investigated text based, or linguistic, cues to deception using mainly computer programs and algorithms. For example, some researchers attempted to automate the detection process by creating tractable constructs clustering linguistic cues and testing models of deception they proposed (Burgoon et al., 2003; Fuller et al., 2013; Zhou et al., 2004a; Zhou, Burgoon, Twitchell, Qin, & Nunamaker Jr, 2004b).

Zhou and her colleagues compared the performance of four classification methods in automated processes to detect deception (Zhou et al., 2004b). Using linguistic data from two experimental studies, they found that compared to truth tellers, deceivers used more modal verbs and made fewer individual references, which resulted in more tentative and non-

specific language. Furthermore, as an indication of cognitive effort, the average length of words for deceivers was shorter than that for truth tellers. Fuller and colleagues (2013) compared two frameworks for detecting deception in text: the one proposed by Zhou et al. (2004a) and their own revised framework. Zhou and her colleagues used an automated linguistic analysis and found nearly 20 significant cues that helped differentiate deceivers from truth tellers (Zhou et al., 2004a). Burgoon and Qin suggested a modified version of the framework suggested by Zhou et al. (2004a) (Burgoon and Qin, 2006). Fuller and her colleagues investigated these text-based cues to deception and their convergence to constructs in a setting with high-stakes consequences for deceiving (Fuller et al., 2013). The confirmatory factor analysis of data suggested that the framework proposed by Zhou and colleagues (2004a) showed a better fit and construct reliability.

In a recent study, Hauch and her colleagues investigated an extensive list of linguistic cues to deception and their detectability with computer programs (Hauch et al., 2015). Based on the theoretical frameworks rooted in cognitive and memory-oriented approaches, social psychology, and the self-presentation perspective, they evaluated linguistic cues from 44 studies. Thirty eight of those linguistic cues were chosen for their meta-analysis and are listed on Appendix A. As the text-based cues in Hauch et al., meta-analysis were used to investigate how computers can detect deception in text, some of them are not applicable for human reading. We focus on 15 linguistic cues (highlighted in the Appendix A) in this study and investigate their effect on ocular-motor measures and detection accuracy.

In this study we focus on four themes, grounded on theoretical frameworks of deception, as an organizing principle to cluster text-based cues to deception listed by prior research. The first theme is built around the premise that lying is cognitively more difficult



than truth telling. The second theme proposes hypotheses based on strong emotions associated with deception, such as guilt and fear. The third theme develops hypotheses based on RM and the information management principle of IDT. The last theme draws from CBCA and RM frameworks to propose hypotheses about the nature of the sender's references to cognitive processes. These themes will help us to cluster linguistic cues to deception around each theme and to analyze the effect of those cues on reading metrics (i.e. fixation duration, fixation frequency, regression count, etc). We start our discussion around the effect of the message veracity on linguistic cues and corresponding ocular-motor measures. Next, we will investigate the relationship between the ocular-motor measures on linguistic cues and deception detection.

### ***Theme 1: Cognitive load***

As suggested by the aforementioned deception theories, lying is associated with increased cognitive load. The multi-factor theory, leakage theory, and IDT suggest that lying involves carrying out multiple tasks simultaneously: creating coherent deceptive message, monitoring the receiver's reactions, adjusting own behavior, remembering details, suppressing the truth, and others. In their meta-analysis of linguistic cues to deception, Hauch and colleagues focused on a cognitive and memory-oriented approach and argued that lying is cognitively taxing because of the difficulty of drawing on episodic memories (Hauch et al., 2015). They found that liars used fewer words and provided less complex stories. The difficulty of formulating lies could be reflected in both its content and delivery. DePaulo and colleagues (2003) demonstrated that liars' accounts are less plausible and coherent. Less plausible stories should result in more frequent revisiting of the read passage on the part of the veracity judge and more time needed to understand the stories, which should be

manifested in a higher number of regressive eye movements, longer reading time, longer fixation durations, and more frequent fixation count. We hence propose:

**Hypothesis 1:** *Deceptive statements should result in more frequent regressive eye movements than honest statements.*

**Hypothesis 2:** *Deceptive statements should take more time to read than honest statements.*

**Hypothesis 3:** *Deceptive statements should result in longer total fixation durations and more frequent fixation counts than honest statements.*

Deceivers have to create a story that did not happen, and it is cognitively hard to construct a sound and compelling story, and thus deceivers tend to keep their stories simple and avoid complexity. Complexity refers to how difficult it is to understand the passage. In a written text, complexity could be displayed through the use of complex words vs. common words (lexical complexity) and the use of compound or complex sentences (syntactic complexity), which include dependent clauses and result in more punctuation (Fuller et al., 2013). The less complex sentence structure of liars should include fewer causation words (e.g., because, hence) and fewer exclusive words (e.g., but, expect, without, and exclude). In their meta-analysis of linguistic cues to deception, Hauch and colleagues operationalized complexity by analyzing the causation and exclusive words. In our discussion of reading behavior, we stated that readers process information by fixating gaze on words. Less frequent presence of causation and exclusive words in deceptive statements should result in corresponding fixation behavior. Complexity is also displayed through the quantity of verbs, words, and sentences. In our study design, because we were interested in the visual attention of the veracity judges, we decided to randomly select written statements of relatively equal

lengths and therefore we will not focus on the word and sentence counts across the statement veracities. We next offer our following hypotheses:

**Hypothesis 4:** *Fixation duration and fixation frequency on causation words in deceptive statements should be less than in honest statements.*

**Hypothesis 5:** *Fixation duration and fixation frequency on exclusive words in deceptive statements should be less than in honest statements.*

Newman and his colleagues (2003) argued that the truth tellers and the liars use language in predictably different ways. They suggested that less complex stories should focus on simpler, concrete verbs because “concrete actions are easier to string together than false evaluations” (p.667). Using the data from five studies, they concluded that deceptive communication was characterized with fewer exclusive words and more motion verbs (e.g., walk, move, go). Hauch and colleagues (2015) also found that honest messages included fewer motion verbs. We hence propose:

**Hypothesis 6:** *Fixation duration and fixation frequency on motion verbs in deceptive statements should be more than in honest statements.*

### **Theme 2: Psychological emotions**

Multi-factor theory suggests that strong emotions, such as guilt and fear, may result in liars psychologically distancing themselves from their own stories. DePaulo and colleagues (2003) found that liars are less-immediate, referring less to themselves, and are more uncertain. In their discussion of linguistic cues and underlying construct definitions, Zhou and colleagues (2004a) suggested that uncertainty refers to the ambiguous and vague nature of statements made to avoid giving direct and unequivocal answers. The accounts of liars, hence, should include less certainty words (e.g., always, clear, never). Psychological

distancing is also characterized by the more frequent use of generalizing terms (e.g., everybody, all, anybody) (Hauch et al., 2015). We hence propose:

**Hypothesis 7:** *Fixation duration and fixation frequency on certainty words in deceptive statements should be less than in honest statements.*

**Hypothesis 8:** *Fixation duration and fixation frequency on generalizing terms in deceptive statements should be more than in honest statements.*

Strong emotions, such as guilt and shame, may result in liars telling less convincing stories. DePaulo et al. (2003) suggested that liars are less immediate verbally and vocally. Nonimmediacy refers to the language used to create a psychological distance between the sender and his or her message (Fuller et al., 2013). The sender will use terms to create spatial and temporal nonimmediacy, which is generally accomplished through the use of passive verbs, avoiding using of first-person pronouns (e.g., I, me, my, we, our, us, etc.) and more frequent use of third-person pronouns (e.g., he, she, her, they, etc.) (Fuller et al., 2013; Hancock et al., 2007; Hauch et al., 2015). In a study of automated detection of deception in text, researchers focused on analyses of linguistic style to create a multivariate profile of deception (Newman, Pennebaker, Berry, & Richards, 2003). The authors argued that truth tellers and liars use language in predictably different ways. Using data from five studies, they concluded that deception communication was characterized with fewer first-person singular pronouns (I, me, myself) and fewer third person pronouns. Hancock and colleagues found that liars used fewer first-person pronouns and more third-person pronouns (2007). Similarly, Hauch et al. (2015) in their meta-analysis found that liars use third-person pronouns (he, she, they, her, their, etc.) more often than truth tellers. We hence propose:

**Hypothesis 9:** *Fixation duration and fixation frequency on passive verbs in deceptive statements should be more than in honest statements.*

**Hypothesis 10:** *Fixation duration and fixation frequency on first-person pronouns in deceptive statements should be less than in honest statements.*

**Hypothesis 11:** *Fixation duration and fixation frequency on third-person pronouns in deceptive statements should be more than in honest statements.*

Prior research has investigated affect, which refers to the information about feelings and represents the hedonic tone of a message (Zhou et al., 2004a). Previous research on emotions of liars found that people felt discomfort and guilt while lying (DePaulo et al., 2003). Strongly felt emotions of guilt or discomfort should be reflected linguistically in deceptive communication by more frequent use of words that reflect negative emotions (e.g., sad, hate, horrible, enemy, worthless). Previous research found that deceptive communication contained more frequent use of negative emotion words (Hauch et al., 2015; Newman et al., 2003). We thus propose:

**Hypothesis 12:** *Fixation duration and fixation frequency on negative emotion words in deceptive statements should be more than in honest statements.*

Deception, as discussed above, is associated with fear. People lie, among other reasons, to avoid punishment and other possible negative consequences. One way to avoid any type of punishment is to deny the wrongdoing or to deny the knowledge of it. DePaulo and colleagues found that liars provide more negative statements and complaints (2003). Similarly, in their meta-analysis of the linguistic cues, Hauch and colleagues (2015) found support for more frequent use of negations (no, never, not) by deceivers. Drawing on our discussion of psychological emotions we draw our next hypothesis:

**Hypothesis 13:** *Fixation duration and fixation frequency on negation words in deceptive statements should be more than in honest statements.*

**Theme 3: Sensory and contextual information**

In accordance with the information management principle of IDT, liars are expected to alter the dimensions of the message, i.e., its content and style. Because of the multi-tasking activity of the deceiver, liars should tend to keep their stories shorter, less detailed, and less specific. Specificity refers to the type and degree of details in a message, which may reflect information about time and space, as well as information about perceptions and sensory experiences (Zhou et al., 2004a). In contrast to real experiences of the truth tellers, senders must construct their accounts from their thoughts and beliefs. According to RM, this should result in deceptive accounts that are less contextually embedded, devoid of sensory experiences, and plain in temporal and spatial references. DePaulo et al (2003) suggested that liars tell less compelling tales than truth tellers, which results in fewer details in a liar's account, less sensory information, and fewer unique words. Truth tellers can refer to their memories of experienced events and provide more detailed stories with contextual and semantic information. Prior research testing the components of the CBCA found that honest accounts included more details and more contextual embeddings (Vrij, 2015). Thus, the messages by deceivers should include fewer details related to the liar's senses (taste, touch, smell, hearing, sight) as well as fewer details about spatio-temporal information (e.g., information about locations, time of events, the sequence of events). We hence propose:

**Hypothesis 14:** *Fixation duration and fixation frequency on sensory words in deceptive statements should be less than in honest statements.*

**Hypothesis 15:** *Fixation duration and fixation frequency on words describing temporal information in deceptive statements should be less than in honest statements.*

**Hypothesis 16:** *Fixation duration and fixation frequency on words describing spatial information in deceptive statements should be less than in honest statements.*

#### **Theme 4: Memory**

In this theme we develop our hypotheses based on RM and CBCA frameworks.

Reality monitoring suggests that memories based on real experiences and those based on imagination differ. This approach predicts that mentioning cognitive operations related to an event are more likely to occur in imagined than in self-experienced events (Hauch et al., 2015; Vrij, 2015). Criteria Based Content Analysis of Statement Validity Assessment, on the other hand, predicts that accounts of subjective mental state, including reports of thought, are more likely to occur in honest statements (Vrij, 2008). Empirical findings of multiple studies report mixed results for the association between references to cognitive processes and deception (Hauch et al., 2015). This criterion has not found support in the meta-analysis by DePaulo et al. (2003). In high-stakes contexts, where the consequences of being found guilty of misconduct are potentially high, liars might tend to provide justifications for their actions using insight words (e.g., think, know, consider, remember) and cognitive reasoning processes (e.g., cause, ought). We hence hypothesize:

**Hypothesis 17:** *Fixation duration and fixation frequency on words describing insight words in deceptive statements should be more than in honest statements.*

**Hypothesis 18:** *Fixation duration and fixation frequency on words describing cognitive words in deceptive statements should be more than in honest statements.*

#### **Omitted linguistic cues**

We tried to draw hypotheses about the linguistic cues to deception around those mentioned in prior research. Not all text-based cues are investigated in our study, mainly because of the study design and the types of stimuli used. For example, quantity is a

construct investigated by most of the prior research on text-based deception. Quantity refers to the number of words in a sentence and the number of sentences in a statement. It is argued that quantity of words used and sentences completed should be less for deceivers who want to remain reticent and avoid providing information that later could be verified. In our study we randomly selected text stimuli of relatively equal length to minimize variance in reading time. Therefore our discussion did not focus on the quantity of words and sentences.

Similarly, another often-mentioned linguistic cue to deception is informality. Informality refers to the use of informal language, manifested in the number of typo errors in a written statement. Because the written statements we will use in our study are transcriptions, we will not discuss informality in this study. Further, some of the cues in Appendix A list both, the umbrella terms and more nuanced elements of those (for example, cues 18 and 28 list both umbrella term cues and more detailed cues, which are part of the umbrella cues). Our study focused on umbrella term linguistic cues.

### **Detection accuracy**

The veracity judges reading the statements will engage in cognitive process of information processing, and any cues noticed during this activity may arouse suspicion in the readers (McCornack and Levine, 1990). The accuracy of detecting deception will thus depend on the cues detected (focused on) and the interpretation of those cues by the readers. We thus hypothesize that:

**Hypothesis 19:** *What veracity judges look at determines their detection success.*

Our earlier discussion on reading metrics suggests that cognitive processing of visual information occurs during fixations only. Whether the reader noticed a linguistic cue or not is manifested in fixation metrics associated with her reading, generated by the eye tracking technology. The data captured by an eye tracker could be analyzed to investigate whether the



linguistic cues highlighted in automated deception detection research are detected when read by humans. This information may help us to identify which one of these cues seem the most salient and dominant in affecting judgment.

## **Research Method**

### **Experimental stimuli**

Written statements obtained from military personnel who completed a “person of interest” statement in criminal investigations were identified and analyzed in a different research study (Fuller et al., 2013). These researchers allowed us to use the research materials used in their study in our experiment. These written statements represent actual official reports written by either subjects or witnesses in investigations of criminal incidents on a U.S. military base. The consequences for offences ranged in their nature from Letter of Reprimand to incarceration to Court Martials. The subjects and witnesses in these investigations attempted to lie to avoid punishment. The base law enforcement personnel established the ground truth for these statements through various means, including subjects admitting to lying, presented evidence to corroborate or to negate statements, and impartial witness testimonies. Fuller and colleagues used 367 written statements applying software tools for linguistic analysis and performed two confirmatory factor analyses to compare the framework suggested by Zhou et al., (2004) and a modified framework, proposed by the authors. They found that the model proposed by Zhou et al. (2004a) had slightly better fit metrics than the revised model.

### **Pilot study**

To investigate the research questions, we conducted a repeated measures, within and between subject experimental design varying experimental condition (deception, truth).

Written statements are the unit of analysis. We recruited participants from a College of

Business whose native language is English. Including only native speakers allows us to reduce variability between participants due their comprehension of English (which may affect the fixation duration on difficult words) and minimize the variability due to reading patterns (natives of languages who read from right to left or from bottom to top may contribute to variation). We recruited participants by verbally announcing the study and by providing an online link to the available time slots from which they chose the desired session time. During the announcement we provided the purpose and the general outline of the study and informed that participants would be paid \$10 for their participation.

We started with a total 104 stimuli statements, ranging from 32 to 490 words per statement. To minimize the variance in reading time due to the lengths of the written statement, we decided to randomly select statements of relatively equal lengths. Ten statements, five honest and five dishonest, were randomly selected, among the range of 110 and 150 words per statement group. Eight participants (2 females), ranging in age from 19 to 35, took part in the pilot study. Each participant of the pilot study was presented with 10 written statements. After signing the informed consent form, each participant was seated in front of a 22" monitor with an eye tracking system mounted underneath the monitor. Before completing the main task, each participant completed the practice task to make sure she was comfortable operating the keyboard and the mouse and to make sure she understood the task well. Each statement was presented to the participants in a randomized order and was followed by two sets of questions. First, the participants had to assess whether the statement was honest or dishonest. Next, they were asked to provide reasons for their assessment. They typed their responses into a form box. After the experiment, they completed a short demographics questionnaire and were debriefed.

After the pilot test, the researchers decided to slightly enhance the study and changed a couple of stimuli statements.

### Main Study

We recruited participants through making verbal announcements in multiple classes taught at the College of Business at a large Midwestern university. The principal investigator explained the purpose of the study and the reasons for why only participants whose native language is English could take part in the study. All the classes where announcements were made received an electronic link for signing up for the study. Those interested in the study could sign up by choosing from the alternative dates and times when the study was available. Those who signed up were automatically reminded of the session's time and location one day before the experiment and were instructed to not wear mascara. Thirty four students took part in the study. Table 3-1 shows participant information. The experiment design was single factor (message veracity) repeated measures, within and between subjects design.

Table 3-1. Participant demographic information

		Frequency	Percent
Gender	Male	29	85.3%
	Female	5	14.7%
		34	100.0%
Age	18-24	32	94.1%
	24-34	2	5.9%
		34	100.0%

Each experimental session was completed in a single event with one participant per session. Each session was scheduled for 60 minutes, and most of them lasted around 30-40 minutes. Each participant began the study by signing the informed consent form. The principal investigator addressed any questions that participants asked. Next, the overall

procedures of the experiment were explained, and the participant was seated in front of a 22” monitor with the eye tracker installed underneath. We used 250 Hz frequency eye tracker model RED 250 from SensoMotoric Instruments (SMI). This eye tracker records eye movement data every 4 milliseconds and allows the capture of eye movements in a remote, contact-free setup with free head movements. This machine is capable of tracking eye movements of participants including those wearing eye glasses or contact lenses.

Each tracking session started with gaze calibration and validation. When necessary, this process involved an adjustment of the participant’s seat forward/backward or upwards/downwards for the best capturing of eye movements. The calibration and validation processes took, when necessary, several rounds to make sure the results were in the desired range. The same researcher ran all experimental sessions and operated the eye tracking software.

Each session started with a practice task. This was designed to make sure the participants were comfortable with instructions, with using the peripherals, and understood the task properly. After the practice task was complete, the screen prompted whether the participants had any questions. When the researchers made sure that participants had no questions and understood the task well, the participants were allowed to continue on the main task. For both practice and main tasks, the participants were presented with written statements they were asked to read and assess for veracity. Each statement was followed by two sets of questions. The first question asked the participants to judge the veracity of the statement, and the follow up question asked them to provide reasons for their assessment. For the first question, they chose their answer on a 7-point Likert scale, with “1” being “completely dishonest” and “7” being “completely honest.” For the second question, they

typed in their responses. All participants were told that they were not restrained on time for either the practice task or the main task. The order in which the statements showed up on the screen was randomized. After the experimental session was complete, the participants completed a short questionnaire on background information. Next, they were debriefed about the study and paid \$10 for participation.

In choosing written statements, we decided to minimize the variance in reading time and therefore randomly choose 12 statements from a subset of relatively equal length statements (see Appendix B). All of the statements were either written statements or accounts of transcribed text of the actual subjects or witnesses who completed “person of interest” statements. Any misspellings in written statements were replaced. Six of those statement were deceptive, and the other 6 were truthful.

### **Measures**

The message veracity is the independent variable with two values (1) honest (coded 1) and (2) deceptive (coded 0). Responses of veracity assessment are the dependent variable. The responses on a 7-point scale were collapsed into three categories: scores of 3 and below were coded as dishonest, 4 was treated as neutral, and scores above 4 was coded as honest. With each of the 34 study participant providing 12 responses, we collected 408 total responses. Of the 408 responses, 28 were treated as neutral and were excluded from the analysis. Thus, our final data included 380 dichotomous assessments of veracity.

Visual behavior of the respondent was recorded and generated by the eye tracking software. The software provided data on total time spent, total fixation duration, and total fixation counts by each participant on each statement. Special features of the software collected data on reading metrics, including the fixation duration and count on each word, lengths of saccadic moves, and the direction of eye moves (regressive vs. progressive).

Reading each statement thoroughly, we identified areas of interest (AOI) based on the discussion of the linguistic cues. For example, each occurrence of the words “no,” “not,” “won’t” were marked as negation. The software generated fixation durations (measured in milliseconds) and fixation counts of participants’ gaze for each of the 16 AOIs we designated.

### **Analysis and Results**

We used repeated measures mixed model in SPSS 24 to analyze the relationship between media and attention foci on linguistic cues to deception. Mixed method does not impose constraints and assumptions of Repeated Measures ANOVA, such as exclusion of missing data. In the pairwise comparison tables shown below, mean differences with positive sign represent the higher values for deceptive statements and negative signs represent higher values of honest statements. The adjustment of statistical tests for multiple comparisons was carried for the tests of hypotheses 1 through 18. Dividing the alpha value of 0.05 by 18 results in the adjusted value of 0.0028 as a new threshold for statistical significance test.

#### **Theme 1: Cognitive load**

##### ***Regression count***

We hypothesized that deceptive statements would be less logical, less coherent, and more ambiguous and thus would require the reader to go back in the text to re-read some portions of the text, which should result in more regressive eye movements. The judges regressed more frequently when they read honest statements (mean=10.86, se=0.933) versus when reading deceptive statements (mean=9.633, se=0.704), and this difference is significant at  $\alpha=0.1$  ( $p=0.083$ ). Hypothesis 1 was not supported (see Table 3-2).

### *Total time*

There was no significant difference in total time spent over assessing the statements across statement veracity. While deceptive statements took longer to assess than honest messages, the mean difference in time for reading and making the assessment was not statistically significant ( $p=0.106$ ). Hypothesis 2 was not supported.

Table 3-2. Pairwise comparisons – mean differences

Hypothesis	Reading metric	Mean Difference	Std. Error	df	Sig
H1	Regressions	-1.233	0.708	213.689	0.083
H2	Total Time	3284.478	2026.599	231.793	0.106
H3	Fixation duration total	3908.489	1694.629	229.93	0.022
H3	Fixation count total	3.496	6.763	244.091	0.606

### *Fixations Duration and Count Total*

We analyzed our data to understand whether the fixation duration of the judges varied across the types of statements. Judges fixated longer on deceptive messages (mean=43022.23 ms, se=2017.443), compared to honest statements (mean=39113.743 ms, se=1915.154), however the mean difference in total fixation duration between honest and deceptive statements was not statistically significant ( $p=0.022$ ). There was no significant difference in total fixation count across the honest and dishonest statements. While the judges fixated more frequently on deceptive statements, this difference was not statistically significant ( $p=0.606$ ). Hypothesis 3 was not supported.

Tables 3-3 and 3-4 show pairwise mean differences in fixation durations and fixation counts, respectively, for all of the linguistic features that were the focus of Hypotheses 4 through 18.

### *Causation*

There was no statistically significant difference between the judges' fixation duration on causation words when reading honest and dishonest statements ( $p=0.107$ ). Similarly, the fixation count on the words describing causation was not significant across the statement types ( $p=0.097$ ). Hypothesis 4 was not supported.

Table 3-3. Pairwise comparisons of fixation duration on AOIs - Mean difference

Hypothesis	Linguistic Cue	Mean Difference	Std. Error	df	Sig
Theme 1: Cognitive load					
H4	Causation	142.782	86.73	45.11	0.107
H5	Exclusive	91.139	71.133	36.834	0.208
H6	Motion verbs	245.305	41.663	159.239	0.000
Theme 2: Psychological emotions					
H7	Certainty	81.366	39.879	66.349	0.045
H8	Generalizing terms	-65.386	29.836	156.343	0.030
H9	Passive voice verbs	155.831	43.422	150.513	0.000
H10	First-person pronouns	96.4	60.027	251.665	0.110
H11	Third-person pronouns	-173.975	51.561	39.161	0.002
H12	Negative emotions	61.7	54.289	36.069	0.263
H13	Negations	-95.552	28.412	54.629	0.001
Theme 3: Sensory and contextual information					
H14	Sensory processes	192.591	36.803	106.848	0.000
H15	Temporal	342.75	98.474	148.888	0.001
H16	Spatial	-369.664	82.779	113.49	0.000
Theme 4: Memory					
H17	Insight	171.14	58.057	130.408	0.004
H18	Cognitive processes	-91.254	54.42	123.189	0.096

Note: mean difference = mean deceptive – mean truthful



Table 3-4. Pairwise comparisons of fixation count on AOIs -Mean difference

Hypothesis	Linguistic cue	Mean Difference	Std. Error	df	Sig
Theme 1: Cognitive load					
H4	Causation	0.496	0.294	48.567	0.097
H5	Exclusive	0.424	0.355	37.111	0.239
H6	Motion verbs	1.028	0.167	176.634	0.000
Theme 2: Psychological emotions					
H7	Certainty	0.501	0.17	111.166	0.004
H8	Generalizing terms	-0.26	0.139	148.5	0.065
H9	Passive voice verbs	0.786	0.198	167.881	0.000
H10	First-person pronouns	0.315	0.227	225.601	0.168
H11	Third-person pronouns	-1.118	0.226	41.031	0.000
H12	Negative emotions	0.305	0.251	37.345	0.231
H13	Negations	-0.487	0.128	182.701	0.000
Theme 3: Sensory and contextual information					
H14	Sensory processes	0.785	0.146	106.703	0.000
H15	Temporal	0.927	0.369	140.475	0.013
H16	Spatial	-1.457	0.35	150.31	0.000
Theme 4: Memory					
H17	Insight	0.568	0.214	130.812	0.009
H18	Cognitive processes	-0.377	0.209	120.914	0.073

Note: mean difference = mean deceptive – mean truthful

### *Exclusive*

There was no statistically significant effect of the message veracity on fixation duration time on words describing exclusive words ( $p=0.208$ ). While judges looked longer at exclusive words in deceptive messages, the difference was not different from what could have been obtained by chance. Similarly, the frequency of fixations on exclusive words did not differ significantly across message types ( $p=0.239$ ). Hypothesis 5 was not supported.

### *Motion verbs*

As predicted, judges fixated longer on words describing simple motions (i.e., go, walk, run) when reading deceptive statements (mean=245.305 ms, se=41.663), compared to

when reading honest statements (mean=122.71 ms, se=21.492). The mean difference in fixation duration on motion verbs was statistically significant ( $p<0.001$ ). Judges fixated more frequently on words describing motion when reading deceptive statements (mean=1.558, se=0.152), compared to when reading honest statements (mean=0.53, se=0.09). This mean difference in fixation count was statistically significant ( $p<0.001$ ). Hypothesis 6 was supported.

## **Theme 2: Psychological emotions**

### *Certainty*

The study participants looked longer at words explaining certainty when reading deceptive messages (mean=225.655 ms, se=39.589) versus truthful statements (mean=144.29 ms, se=23.813), yet this difference was not statistically significant at the adjusted level of alpha ( $p=0.045$ ). Similarly, the judges looked more frequently at certainty words when reading deceptive statements, compared to when reading honest statements, but this difference was not statistically significant ( $p=0.004$ ). Hypothesis 7 was not supported.

### *Generalizing terms*

We did not find a statistically significant association between the message veracity and generalizing terms. The veracity judges fixated longer on words describing generalizing terms in honest statements (mean=193.578 ms, se=29.481) compared to deceptive messages (mean=128.193 ms, se=23.012). This difference in mean fixation duration times was not statistically significant ( $p=0.030$ ). Similarly, the difference in fixation count on generalizing terms, was not significant at  $a=0.0028$  ( $p=0.065$ ). Hypothesis 8 was not supported.

### *Passive voice verbs*

As predicted, the judges' gaze fixated longer on passive verbs in deceptive statements (mean=330.662 ms, se=34.309), compared to that of when reading honest statements

(mean=174.831 ms, se=26.626). This difference is statistically significant ( $p<0.001$ ).

Similarly, the judges looked more frequently at passive verbs in deceptive statements, compared to when reading honest statements ( $p<0.001$ ). Hypothesis 9 was supported.

### ***First-person pronouns***

Judges reading deceptive statements fixated longer on the first noun pronouns versus when reading honest statements, but this difference was not statistically significant ( $p=0.11$ ). Similarly, judges fixated more frequently on the first person pronouns in deceptive messages than when reading honest statements, but the mean difference in the frequency of fixations on the words was not statistically significant ( $p=0.168$ ). Hypothesis 10 was not supported.

### ***Third-person pronouns***

The judges fixated longer on the words describing second person pronouns (i.e., she, he, they) in honest statements (mean=343.302 ms, se=44.661), compared to those in dishonest statements (mean=169.326 ms, se=25.768), and this difference was statistically significant ( $p=0.002$ ). Similarly, the judges fixated more frequently on second person pronouns when reading honest statements ( $p<0.001$ ). The direction of the relationship was in opposite direction than predicted. Hypothesis 11 was not supported.

### ***Negative emotions***

The judges looked longer at the words describing negative emotions in deceptive messages than when reading honest statements, but this difference was not significantly different ( $p=0.263$ ). Similarly, the difference in fixation count on words describing negative emotions was not statistically significant across the two types of statements ( $p=0.231$ ). Hypothesis 12 was not supported.

### *Negations*

There was a statistically significant effect of the message veracity on the fixation duration on negative words. Interestingly, however, the judges looked longer at negative words in honest statements (mean =163.663 ms, se=25.748) than in deceptive statements (mean =68.111 ms, se=16.705). The mean difference of fixation durations is statistically significant ( $p=0.001$ ). Similarly, the judges looked more often at the negation words in honest statements than in dishonest statements ( $p<0.001$ ). Hypothesis 13 was not supported.

### **Theme 3: Sensory and contextual information**

#### *Sensory processes*

When reading deceptive statements, the judges looked longer at the words describing sensory processes (mean=275.689 ms, se=33.720) versus when reading honest statements (mean=83.098 ms, se=18.261). The mean difference (mean diff=192.591 ms, se=36.803) is statistically significant ( $p<0.001$ ). Similarly, there is a statistically significant effect of the message veracity on how many times judges fixated on the words that express the sender's senses ( $p<0.001$ ). Judges tended to fixate more frequently on the words expressing senses in deceptive statements. The relationship between the words describing sensory perceptions and message veracity was not in predicted direction. Hypothesis 14 was not supported.

#### *Temporal*

The veracity judges looked longer at the words describing the temporal dimension of the context in dishonest messages (mean=657.188 ms, se=98.849), compared to fixation durations in honest statements (mean=314.438 ms, se=52.804). This difference in mean fixation duration is statistically significant across the types of messages ( $p=0.001$ ). The predicted direction of the effect of message veracity was different, however. The judges looked more frequently at words describing temporal aspect of the context in deceptive

statements compared to that in truthful statements, but this difference is not statistically significant ( $p=0.013$ ). Hypothesis 15 was not supported.

### *Spatial*

The judges looked longer at words describing spatial dimensions of the context in honest statements (mean=704.195 ms, se=78.175) than in dishonest statements (mean=335.53 ms, se=42.051). This difference is statistically significant ( $p<0.001$ ). The judges looked more frequently at words providing spatial details of the honest statements (mean=3.282, se=0.327) versus dishonest statements (mean=1.825, se=0.203). This difference was statistically significant ( $p<0.001$ ). Hypothesis 16 was supported.

## **Theme 4: Memory**

### *Insight*

Judges fixated their gaze longer on words referring to insight in deceptive messages (mean=362.535 ms, se=52.068) compared to when reading honest statements (mean=191.395 ms, se=25.682). This difference, however was not statistically significant at the adjusted alpha level ( $p=0.004$ ). Veracity judges looked more frequently at insight words in deceptive messages (mean=1.356, se=0.19) than in honest messages (mean=0.788, se=0.1). Similarly, this mean difference in fixation count across statement types was not significant ( $p=0.009$ ). Hypothesis 17 was not supported.

### *Cognitive processes*

The judges fixated longer on words describing cognitive processes in truthful statements compared to those when reading deceptive statements. This difference, however, while statistically significant at  $\alpha=0.1$ , was not significant at  $\alpha=0.05$  ( $p=0.096$ ). A similar level of statistical significance is found when comparing the fixation count across honest and

dishonest statements. The judges looked more frequently at the cognitive words in honest statements than in lies ( $p=0.073$ ). Hypothesis 18 was not supported.

### **Detection accuracy: All cues**

The judges correctly identified the veracity of 235 statements out of 380, which resulted in 61.8% total accuracy rate. This is a much better performance rate compared to the pilot study results of 47.1% total accuracy rate. The total hit rate (lies correctly identified as lies) was 52.8%, and the true negative rate (truthful statements correctly identified as truthful) was 71.1%.

We used repeated measures logistic regression to analyze the association between the eye movements (i.e., regressions, fixation duration, and fixation count) on linguistic cues and detection accuracy because the dependent variable, detection accuracy, is a dichotomous variable. The GENLIN command on SPSS was used with a binomial distribution and logit as the link function. Some of the linguistic cues were not present on all 12 statements. For example, exclusive words, or words describing causation were present in only 3 and 2 statements, respectively. Combining these cues with other variables will result in fewer observations and the binary model not being estimable.

We ran logistic regression for each of the cues individually. The results of this analysis are split into two tables. Table 3-5 shows parameter estimates for logistic regression for the duration of each independent variable. Table 3-6 shows parameter estimates for logistic regression for the fixation count on each of the independent variables.

Table 3-5. Parameter estimates of logistic regression. Each estimate is analyzed individually.

<b>Parameter</b>	<b>B</b>	<b>Std. Error</b>	<b>Wald Chi-Sqr</b>	<b>df</b>	<b>Sig.</b>	<b>Exp(B)</b>
Total Time	-0.007	0.003	6.134	1	0.013	0.993
Total Fixations Duration	-0.008	0.004	5.957	1	0.015	0.992
Causation	-0.212	0.583	0.132	1	0.717	0.809
Exclusive	-2.702	1.473	3.363	1	0.067	0.067
Motion verbs	-0.341	0.201	2.870	1	0.090	0.711
Certainty	-0.215	0.240	0.797	1	0.372	0.807
Generalizing terms	0.247	0.278	0.787	1	0.375	1.280
Passive voice verbs	0.096	0.107	0.807	1	0.369	1.101
First-person pronouns	-0.184	0.109	2.878	1	0.090	0.832
Third-person pronouns	-0.151	0.122	1.527	1	0.217	0.860
Negative emotions	-0.317	0.652	0.237	1	0.626	0.728
Negations	0.528	0.303	3.038	1	0.081	1.695
Sensory processes	-1.210	0.462	6.874	1	0.009	0.298
Temporal	-0.135	0.063	4.536	1	0.033	0.874
Spatial	-0.048	0.077	0.384	1	0.535	0.954
Insight	-0.091	0.290	0.098	1	0.754	0.913
Cognitive processes	-0.182	0.599	0.092	1	0.761	0.834

Table 3-6. Parameter estimates of logistic regression. Each estimate is analyzed individually.

<b>Parameter</b>	<b>B</b>	<b>Std. Error</b>	<b>Wald Chi-Sqr</b>	<b>df</b>	<b>Sig.</b>	<b>Exp(B)</b>
Total Fixations Count	-0.002	0.001	2.798	1	0.094	0.998
Regressions	-0.004	0.010	0.158	1	0.691	0.996
Causation	-0.106	0.158	0.450	1	0.502	0.899
Exclusive	-0.654	0.444	2.178	1	0.140	0.520
Motion verbs	-0.082	0.057	2.113	1	0.146	0.921
Certainty	-0.122	0.081	2.236	1	0.135	0.885
Generalizing terms	0.035	0.057	0.375	1	0.540	1.035
Passive voice verbs	0.018	0.031	0.352	1	0.553	1.018
First-person pronouns	-0.036	0.029	1.503	1	0.220	0.965
Third person pronouns	-0.023	0.035	0.423	1	0.516	0.977
Negative emotions	0.029	0.133	0.048	1	0.827	1.029
Negations	0.192	0.091	4.396	1	0.036	1.211
Sensory processes	-0.271	0.123	4.878	1	0.027	0.763
Temporal	-0.026	0.017	2.327	1	0.127	0.974

Continued from Table 3-6

Parameter	B	Std. Error	Wald Chi-Sqr	df	Sig.	Exp(B)
Spatial	0.000	0.021	0.000	1	0.993	1.000
Insight	-0.055	0.068	0.646	1	0.421	0.947
Cognitive processes	0.017	0.145	0.014	1	0.905	1.017

The results of the logistic regression (Table 3-5) suggest that fixation duration on words describing sensory processes ( $X^2(1, N=192) = 6.874, p=0.009$ ) and words describing temporal information ( $X^2(1, N=380) = 4.536, p=0.033$ ) were statistically significant. Increasing fixation duration on either of these words decreases detection accuracy by the factors of 0.298 and 0.874, respectively. The results also found that total time ( $X^2(1, N=380) = 6.134, p=0.013$ ) and total fixation duration ( $X^2(1, N=380) = 5.957, p=0.015$ ) were statistically significant. More time spent reading the statements decreased the detection accuracy by a factor of 0.993, and total fixations duration lowered deception detection success by a factor of 0.992. The results of logistic regression on Table 3-6 show that fixation frequency on negations ( $X^2(1, N=380) = 4.396, p=0.036$ ) and words describing sensory processes ( $X^2(1, N=192) = 4.878, p=0.027$ ) were statistically significant. While looking more frequently on words describing sensory information decreased detection accuracy by a factor of 0.763, fixation counts on negations improved detection accuracy by a factor of 1.211.

#### Detection accuracy: By theme

Next, we sought to analyze the relationship between the ocular-motor measures on linguistic cues based on our four proposed themes and deception detection. For each of the four themes, we ran logistic regression analyses, predicting the detection accuracy. The



models used fixation duration on linguistic cues to deception and the fixation counts on linguistic cues as independent variables.

The linguistic cues in the first theme, cognitive load, were developed around the claim that lying is cognitively more difficult than truth telling. The variables included in the first theme were motion verbs, causation, exclusive, total fixation duration, total fixations count, regressions, and total time. The logistic regression in this model omitted two linguistic cues – causation and exclusive words – as those cues were present in very few stimuli statements. In Table 3-7, the variables regressions and total fixations count represent the number of regressive moves and total fixation counts on each statement by each participant. Total time represents the total time each participant spent assessing each statement. Other variables represent fixation metrics on specific AOIs.

Table 3-7. Parameter estimates of logistic regression for theme 1 variables

<b>Parameter</b>	<b>B</b>	<b>Std. Error</b>	<b>Wald Chi-Sqr</b>	<b>df</b>	<b>Sig.</b>	<b>Exp(B)</b>
(Intercept)	0.939	0.2382	15.553	1	0.000	2.558
Motion verbs fix dur	0.501	0.6879	0.529	1	0.467	1.65
Motion verbs fix freq	-0.149	0.1839	0.657	1	0.418	0.862
Total fixations duration	-0.029	0.0188	2.429	1	0.119	0.971
Total fixations count	0.011	0.0041	7.796	1	0.005	1.011
Regressions	0.001	0.0161	0.002	1	0.961	1.001
Total Time	-0.024	0.0131	3.394	1	0.065	0.976

The results of the logistic regression (Table 3-7) show that fixation duration on total fixation count ( $X^2(1, N=315) = 7.796, p=0.005$ ) was statistically significant. The model suggests that the total fixation frequency improves the detection accuracy by a factor of 1.011. The effect of other model variables was not statistically significant.

The linguistic cues discussed around the second theme, psychological emotions, were based on the emotions associated with deception and subsequent psychological distancing that liars create from their accounts. The variables included were certainty, generalizing terms, passive voice verbs, first-person pronouns, third-person pronouns, negative emotions, and negations.

Table 3-8. Parameter estimates of logistic regression for fixation duration of theme 2 variables

Parameter	B	Std. Error	Wald Chi-Sqr	df	Sig.	Exp(B)
(Intercept)	3.319	1.212	7.495	1	0.006	27.630
Certainty fix dur	-5.503	7.888	0.487	1	0.485	0.004
Generalizing terms fix dur	-12.075	5.336	5.120	1	0.024	0.000
Passive voice verbs fix dur	4.225	1.829	5.337	1	0.021	68.387
First-person pronouns fix dur	-9.803	3.595	7.436	1	0.006	0.000
Third-person pronouns fix dur	7.051	3.514	4.025	1	0.045	1153.573
Negative emotions fix dur	-1.999	5.874	0.116	1	0.734	0.135
Negations fix dur	-4.214	3.221	1.712	1	0.191	0.015
Certainty fix freq	-2.073	1.614	1.650	1	0.199	0.126
Generalizing terms fix freq	0.167	0.861	0.037	1	0.847	1.181
Passive voice verbs fix freq	-0.986	0.453	4.738	1	0.030	0.373
First-person pronouns fix freq	1.970	0.799	6.076	1	0.014	7.174
Third-person pronouns fix freq	-0.080	0.284	0.079	1	0.779	0.923
Negative emotions fix freq	0.194	1.451	0.018	1	0.894	1.214
Negations fix freq	2.303	1.214	3.600	1	0.058	10.003

The results of the analysis (Table 3-8) suggest that the fixation duration on words describing generalizing terms ( $X^2(1, N=66) = 5.120, p=0.024$ ), on passive voice verbs ( $X^2(1, N=66) = 5.337, p=0.021$ ), on first-person pronouns ( $X^2(1, N=66) = 7.436, p=0.006$ ), and on third-person pronouns ( $X^2(1, N=66) = 4.025, p=0.045$ ) were statistically significant. Both, the fixation duration on generalizing terms and first-person pronouns seem to worsen detection performance, but the factor of effect is very trivial. Fixation duration on passive voice verbs and third-person pronouns, on the other hand, improved detection accuracy by

the factors of 68.38, and 1153, respectively. Regarding the effect of fixation frequency, Table 3-8 suggests that fixation frequencies on passive voice verbs ( $X^2(1, N=66) = 4.738, p=0.03$ ) and on first-person pronouns ( $X^2(1, N=66) = 6.076, p=0.014$ ) were statistically significant. While fixation frequency on passive voice verbs worsened detection accuracy by a factor of 0.373, fixation counts on first-person pronouns improved detection accuracy by a factor of 7.174.

The third theme, sensory and contextual information, drew arguments from the RM framework and the strategic information management approach put forth by IDT and claimed that the accounts of deceivers would contain less details related to the sender's senses and fewer spatio-temporal contextual details. It featured variables sensory processes, temporal, and spatial.

Table 3-9. Parameter estimates of logistic regression for fixation duration of theme 3 variables

Parameter	B	Std. Error	Wald Chi-Sqr	df	Sig.	Exp(B)
(Intercept)	0.628	0.294	4.580	1	0.032	1.874
Sensory processes fix dur	-1.444	0.855	2.851	1	0.091	0.236
Temporal fix dur	0.175	0.307	0.326	1	0.568	1.192
Spatial fix dur	-0.048	0.479	0.010	1	0.919	0.953
Sensory processes fix freq	0.032	0.233	0.018	1	0.892	1.032
Temporal fix freq	0.000	0.094	0.000	1	0.999	1.000
Spatial fix freq	0.013	0.130	0.010	1	0.922	1.013

The results of the analysis suggest that the effect of none of the fixation metrics on linguistic cues to deception included in the theme 3 were statistically significant.

In the fourth theme, memory, we proposed the relationship between deception and references to cognitive processes based on CBCA and RM frameworks. The variables included were insight and cognitive processes.

Table 3-10. Parameter estimates of logistic regression for fixation duration of theme 4 variables

<b>Parameter</b>	<b>B</b>	<b>Std. Error</b>	<b>Wald Chi-Sqr</b>	<b>df</b>	<b>Sig.</b>	<b>Exp(B)</b>
(Intercept)	1.629	0.481	11.474	1	0.001	5.100
Insight fix dur	-0.353	1.588	0.049	1	0.824	0.702
Cognitive processes fix dur	-5.250	3.461	2.300	1	0.129	0.005
Insight fix freq	0.131	0.443	0.087	1	0.768	1.140
Cognitive processes fix freq	1.012	0.856	1.398	1	0.237	2.752

As suggested in Table 3-10, fixation durations and frequencies on words expressing cognitive processes or insight were not significantly associated with deception detection success. Hypothesis 19 was partially supported. supported. The results of our hypotheses tests are summarized in Table 3-11.

### Discussion

The primary objective of this study was to understand the effect of a written statement's veracity on the reading behavior of a veracity judge and subsequent effect of the visual foci on detection accuracy. Based on the discussion of the theories of deception, reading behavior, and linguistic cues, we predicted that the eye movements and reading behaviors of judges reading deceptive messages should be different from those reading truthful messages. Of the 18 tests we ran, ocular-motor measures on 7 of the tests suggested an effect of the message veracity on reading patterns. We have elaborated on this difference by proposing more detailed hypotheses grounded on theoretical frameworks. Addressing the first research question and applying eye tracking technology, we investigated what judges looked at when reading, how long they fixated on particular words, how often they fixated their gaze, and other metrics.

Table 3-11. Results of hypothesis testing.

Hypotheses	Support?	Finding
<b>Theme 1: Cognitive load</b>		
H1: Deceptive statements should result in more frequent regressive eye movements than honest statements.	No	Message veracity not associated with regressive eye movements.
H2: Deceptive statements should take more time to read than honest statements.	No	Message veracity not associated with total time reading statements.
H3: Deceptive statements should result in longer total fixation durations and more frequent fixation counts than honest statements.	No	Message veracity not associated with total fixation duration or total fixations counts.
H4: Fixation duration and fixation frequency on causation words in deceptive statements should be less than in honest statements.	No	Message veracity not associated with fixation duration and fixation frequency on causation words.
H5: Fixation duration and fixation frequency on exclusive words in deceptive statements should be less than in honest statements.	No	Message veracity not associated with fixation duration and fixation frequency on exclusive words.
H6: Fixation duration and fixation frequency on motion words in deceptive statements should be more than in honest statements.	Yes	Deceptive messages associated with longer fixation duration and higher fixation frequency on motion words.
<b>Theme 2: Psychological emotions</b>		
H7: Fixation duration and fixation frequency on certainty words in deceptive statements should be less than in honest statements.	No	Message veracity not associated fixation duration and fixation frequency on certainty words.
H8: Fixation duration and fixation frequency on generalizing terms in deceptive statements should be more than in honest statements.	No	Message veracity not associated with fixation duration and fixation frequency on generalizing terms.
H9: Fixation duration and fixation frequency on passive verbs in deceptive statements should be more than in honest statements.	Yes	Deceptive messages associated with longer fixation duration and higher fixation frequency on passive verbs.
H10: Fixation duration and fixation frequency on first-person pronouns in deceptive statements should be less than in honest statements.	No	Message veracity not associated with fixation duration and fixation frequency on first-person pronouns.
H11: Fixation duration and fixation frequency on third-person pronouns in deceptive statements should be more than in honest statements.	No	Deceptive messages associated with shorter fixation duration and lower fixation frequency on third-person pronouns. The effect not in predicted direction.
H12: Fixation duration and fixation frequency on negative emotion words in deceptive statements should be more than in honest statements.	No	Message veracity not associated with fixation duration and fixation frequency on negative emotions.
H13: Fixation duration and fixation frequency on negation words in deceptive statements should be more than in honest statements.	No	Deceptive messages associated with shorter fixation duration and lower fixation frequency on negation words. The effect not in predicted direction.

Continued from Table 3-11

Hypotheses	Support?	Finding
<b>Theme 3: Sensory and contextual information</b>		
H14: Fixation duration and fixation frequency on sensory words in deceptive statements should be less than in honest statements.	No	Deceptive messages associated with longer fixation duration and higher fixation frequency on sensory processes. The effect not in predicted direction.
H15: Fixation duration and fixation frequency on words describing temporal information in deceptive statements should be less than in honest statements.	No	Deceptive messages associated with longer fixation duration and higher fixation frequency on temporal information. The effect not in predicted direction.
H16: Fixation duration and fixation frequency on words describing spatial information in deceptive statements should be less than in honest statements.	Yes	Deceptive messages associated with shorter fixation duration and lower fixation frequency on spatial information.
<b>Theme 4: Memory</b>		
H17: Fixation duration and fixation frequency on words describing insight words in deceptive statements should be more than in honest statements.	No	Message veracity not associated with fixation duration and fixation frequency on insight words.
H18: Fixation duration and fixation frequency on words describing cognitive words in deceptive statements should be more than in honest statements.	No	Message veracity not associated with fixation duration and fixation frequency on words referring to cognitive processes.
<b>Detection accuracy</b>		
H19: What veracity judges look at determines their detection success.	Partial	Five linguistic cues associated with detection success.

Only three of our 18 hypotheses were fully supported and one hypothesis was supported partially. As predicted, judges fixated their gaze longer and more frequently on verbs describing motion (i.e. walk, move, go, run) in deceptive statements (H6). Deceivers have to create a story that did not happen. It is cognitively hard to construct a sound and compelling story; thus deceivers tend to keep their stories simple. This results in the use of simple and concrete motion verbs (Newman et al., 2003). Hauch et al., (2015) found that truth tellers' accounts included fewer motion verbs. While prior research on text-based cues focused on proportion on linguistic cues across honest and dishonest accounts, we made inferences from proportions to gaze behavior: the more frequent occurrence of a certain cue should result in more frequent fixation on such cues. Fixation duration and frequency on passive voice verbs was higher when reading deceptive statements (H9). In their meta-

analysis, Hauch and colleagues did not find support for more frequent use of passive voice verbs in deceptive accounts. The study by Zhou and colleagues (2004a) found that deceptive communication was more nonimmediate, characterized by more frequent use of passive voice verbs and generalizing terms. Veracity judges fixated less frequently and shorter on words describing spatial dimensions of the statements (H16). RM predicts that accounts of real experienced events will have more spatial information. DePaulo et al., (2003) found significant effect of message veracity on details provided: deceptive stories included less details. Other studies investigating linguistic cues to deception using automated methods did not find a strong effect on spatial information (Hauch et al., 2015; Zhou et al., 2004a).

The direction of 4 of the statistically significant linguistic cues was not as predicted. These cues are third-person pronouns, negations words, sensory processes, and temporal information. The judges fixated longer and more frequently on the third-person pronouns (H11) in honest statements. As pointed out by Fuller et al., (2013), who created these stimuli materials, respondents to person of interest reports are instructed to be very specific about who they are talking about. This requirement may result in fewer use of third-person pronouns and more frequent reference to specific names. Contrary to our predictions, judges fixated longer and more frequently on negations in honest statements (H13). Prior research has mixed findings on the relationship between deception and negations: meta-analyses found a relationship between negations and deception (Depaulo et al., 2003; Hauch et al., 2015), Newman et al., (2003) found no significant effect, while Hancock et al., found that motivated liars tended to produce fewer negations (Hancock et al., 2007). Negative words such as “no” and “not” are short relative to other content words, and it is plausible that such words are skipped in reading. Human reading processes differ from automated text analysis,

and fixation on shorter words could be explored further by future research. Fixations on sensory words were longer and more frequent in deceptive statements (H14). Hauch and colleagues found no effect on overall sensory words but found that truth tellers referred to their hearing more frequently (Hauch et al., 2015). DePaolo's meta-analysis had found no significant effect of sensory information. Prior research showed that variance in the context results in larger discrepancy in the predictive power of the linguistic cues. Contrary to our prediction effect, deception was associated with longer fixation duration and higher fixation frequency on temporal information (H15). Prior studies found no significant effect of deception on information referencing temporal dimensions an account (DePaulo et al., 2003; Hauch et al., 2015; Newman et al., 2003). It is interesting how spatial dimensions of contextual embedding were significant in one direction and temporal dimensions in an opposite dimension. Deceptive accounts are not always completely deceptive: some information may be truthful (i.e., half-truths, omissions). Such accounts may result in detailed descriptions of spatial and temporal dimensions of an event and thus may result in mixed findings in different research studies.

Eleven of the hypothesis were not supported. Message veracity was not associated with regressive eye movements (H1) and total reading time (H2). Prior research found that honest statements were longer and more elaborate (Hauch et al., 2015). One explanation for finding no effect of regressive moves and total time could originate in the nature of the task. Requests to assess a statement for veracity heightens suspicion of veracity judges, which results in more deliberate processes of revisiting a written account affecting the number of regressive moves and time spent reading the statement. Message veracity was not associated with total fixation duration and total fixation counts (H3). We chose written statements of



relatively equal length. Despite the size equality of the statements, we had predicted that liars' accounts should result in longer fixation duration. While the total fixation duration of judges reading deceptive statements was longer than judges reading honest statements, it was not significant at the adjusted significance test level. DePaulo et al., (2003) found that liars' accounts are less plausible and lack logical structure. This discrepancy from the prior empirical findings may have resulted, as suggested earlier, from the nature of the task, which heightens the suspicion levels in judges and results in more detailed reading of all statements. There was no effect of message veracity on fixation behavior on causation words (H4), on exclusive words (H5), certainty words (H7), on generalizing terms (H8), on first-person pronouns (H10), on negative emotion words (H12), on insight words (H17) and words describing cognitive process (H18). Prior research has found mixed results on the effect of memory on use of cognitive and insight words in written accounts (Hauch et al., 2015). While the fixation behavior on insight words was in the direction predicted by RM, the relationship was not statistically significant at the adjusted significance level. We hypothesized that the judges would look shorter and less frequently at words describing certainty in dishonest statements (H7). Similar to our results, Hauch and colleagues, found no significant effect of message type on certainty words (2015). Given that the statements were produced in high-stake conditions with real-life negative consequence for the people who produced them, it is possible that those people tried to present themselves in such manner, trying to appear more certain. More use of certainty words might have resulted in similar use of certainty in both honest and deceptive statements and hence no significant difference detected. Our findings of non-significant effect of message veracity on fixation behavior on generalizing terms (H8) is similar to the findings of DePaulo and colleagues, who found no

significant association between deception and generalizing terms. Fuller and colleagues found significant association between deceptive communication and generalizing terms. One explanation we can offer is the instrumentalization of this this cue. Hancock suggested that liars would be wary of using distinction markers, and words such as “everybody” or “all” add more precision than, for example, the word “some.” Future research may look into more precise distinction between generalizing terms.

We predicted that the veracity assessments of the judges would be affected by the areas they focus on. We found partial support for our hypothesis (H19). Specifically, we found that total fixation duration on generalizing terms, on first-person pronouns, and fixation frequency on passive voice verbs worsened detection accuracy. Zhou and colleagues (2004a) found that deceptive accounts exhibited greater nonimmediacy manifested in more frequent use of generalizing terms, fewer self and more frequent group references, and more use of passive voice verbs. As prior research has shown, people are worse at detecting deception than honesty. We found that fixation durations on passive voice verbs, third-person pronouns, and fixation frequency on first-person pronouns improves the detection accuracy. Newman and colleagues found that third-person pronouns were used less in deceptive communication (2003). A meta-analysis by Hauch and colleagues found that the third-person pronouns were associated with deceptive accounts. It is interesting how fixation duration and fixation frequency on passive voice verbs and first-person pronouns have varying relationship with the message veracity. While both, fixation duration and frequency are important metrics of different gaze patterns, they are strongly associated with each other. In our study we combined self and group references into one group of first-person pronouns and it is possible that the mix of those could lead to mixed effects. Future research should

investigate the effect of these pronouns in separation. Finally, we found that total fixation frequency on the whole statement improved detection accuracy. The more the judges fixated on words, the better became the odds of detecting deception. Fixation frequencies and durations on other linguistic cues to deception that were found to have an association with deception in prior research were not significantly associated with detection success rate in this study. These discrepancies with prior research could stem from the type of the statements our participants analyzed. Most of the research findings on indicators of deception were studies in labs, where students are recruited and asked to either lie or detect a lie. Indicators of deception created in low-stake environments may result in inconsistent results (DePaulo et al., 2008, Vrij, 2008). Our study, using high-stake statements as experimental stimuli, offers a valuable insight the realistic cues to deception.

Breaking down supported hypotheses by the four themes, we get at least one supported hypothesis for all the themes, but the 4th. The mixed propositions and empirical findings of the relationship between the memory, deception, and associated linguistic cues related to insight words and cognitive processes suggest splitting insight words and words referencing cognitive processes and investigating them separately. The themes provide an important framework for understanding the linguistic cues in written text and in predicting their direction of effect. Written text is viewed as a communication form that transmits fewer cues to deception compared to face-to-face communication (e.g. (George et al., 2018). The discussions and empirical evidence from the themes, however, suggests that certain cues get “leaked” out even in the written form of communication. When assessing the veracity of a speaker, judges tend to focus more on nonverbal cues rather than the content of information (Vrij, 2008). Since written text provides no behavioral cues, what indicators do readers

“seek” in text to confirm or to reject their apriori beliefs? The linguistic cues clustered around the four themes and the eye tracking data help us partially address this question.

### **Implications and future research**

Our findings have implications for both research and practice. From the research perspective, our work makes contributions to the literature on linguistic cues to deception. We proposed an exploratory insight into reading of written deceptive accounts from the perspective of a veracity judge. This study demonstrated how deception is not only manifested in linguistic cues but also in the pattern of decoding it by a human. Prior research has pointed out the importance of investigating “high-stakes” deception to better understand the impact of motivation on deception. The consequences for getting caught deceiving for the army personnel who produced the written statements was very high. By investigating linguistic cues to deception in a high-stake domain, we identified significant associations between message veracity and a reader’s foci on certain linguistic cues and then subsequent impact of fixation foci on detection accuracy. We proposed four themes, based on theories of deception, which helped us to propose hypotheses on the effect of message veracity on the linguistic cues. Our findings provide support for theories of deception. Moreover, the direction of effect of some of the cues, provides an insight about the effect of lies generated from real-life, high-stake settings.

One of the goals of research investigating linguistic cues with computer programs is to be able to compare the machine generated performance with that of humans. Unlike the computer programs that generally analyze the frequency of words, human readers take semantic context into account when reading. In our study, veracity judges’ total accuracy rate was 61.5%, which while higher than the average deception detection rate of 54.4%, was lower than results reported by the computer programs (Hancock et al., 2007; Newman et al.,

2003). Future research should look into the effect of semantic context in affecting detection accuracy.

From the practical perspective, our findings help us better understand written communication and deception detection. Our findings can be utilized in multiple settings: in an organizational setting, employers can be informed of the potential inviable job candidates by more carefully analyzing their applications, in e-commerce environment consumers may be better informed about the quality of a product or a vendor by differentiating between honest and deceptive online reviews, in cyber security environment, users can be better informed of the potential malicious software threats and help their organizations avoid costly expenses of dealing with the aftermath of cyber breaches. Understanding the nature of deceptive information and processes of detecting it could help organizations avoid threats and lower cyber risks. By identifying and empirically testing linguistic cues and their relationship to detection success, we help people to better detect deception.

### **Limitations**

We used student groups to assess the veracity of the written statements. While the recruitment of student judges may be seen as a limitation, students, as do other humans, engage in both deceiving and detection activities. Second, the participants in this study were expected to make veracity judgments. When asked to assess a statement for veracity, the suspicion in judges may get triggered, which may affect their reading patterns. Also, raised suspicion levels are associated with better detection performance and thus generalizations from this study need to take this into account.

Our study was limited to analyzing linguistic cues to deception in English language. Moreover, the statements were collected from a US base, which might affect the style and certain terminology used. For example, most statements indicated very price timing of the

events described. Army personnel has very specific daily routine and thus are expected to be aware of timing. The daily schedules of civilians vary vastly, thus making temporal perceptions less aware. The domain of our text stimuli is very specific and the generalizability of our findings needs to be further studied and explored.

In testing our hypotheses we used written stimuli generated by another research. Some of the linguistic cues discussed by prior research were not available in our stimuli materials or were present in very limited amount. Future research may look into the cues that were missing from our study (positive affect, tentative words, and others).

We operationalized reading of the linguistic cues by the frequency and duration of gaze fixations devoted to certain areas of interest. In our discussion of reading behavior we pointed out that a reader must bring a word into her foveal vision to be able to process it. Over gaze behavior is not always associated with cognitive processing: a person may fixate on a word, but may drift her attention and thinking to something else. Since we cannot measure this covert mechanism of attention, we make an assumption that the participants' foveal focus is aligned with their attention and thus measure only overt attention.

### References

- Abrams, R. A., & Jonides, J. (1988). Programming saccadic eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, 14(3), 428.
- Angele, B., & Rayner, K. (2011). Parafoveal processing of word  $n+2$  during reading: Do the preceding words matter? *Journal of Experimental Psychology: Human Perception and Performance*, 37(4), 1210.
- Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and social psychology Review*, 10(3), 214-234.
- Bond, C. F., & DePaulo, B. M. (2008). Individual differences in judging deception: accuracy and bias. *Psychological bulletin*, 134(4), 477.
- Brysbart, M., Drieghe, D., & Vitu, F. (2005). Word skipping: Implications for theories of eye movement control in reading. *Cognitive processes in eye guidance*, 53-77.

- Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication theory*, 6(3), 203-242.
- Burgoon, J., Blair, J., Qin, T., & Nunamaker, J. (2003). Detecting deception through linguistic analysis. *Intelligence and security informatics*, 958-958.
- Burgoon, J. K., & Qin, T. (2006). The dynamic nature of deceptive verbal communication. *Journal of Language and Social Psychology*, 25(1), 76-96.
- Cook, A. E., Hacker, D. J., Webb, A. K., Osher, D., Kristjansson, S. D., Woltz, D. J., & Kircher, J. C. (2012). Lyin'eyes: Ocular-motor measures of reading reveal deception. *Journal of Experimental Psychology: Applied*, 18(3), 301.
- Darcy, O. (2016, December 29, 2016). The left's emerging 'fake news' problem. *Business Insider*. Retrieved from <http://www.businessinsider.com/fake-news-left-liberal-problem-2016-12>
- DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of personality and social psychology*, 70(5), 979.
- DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological bulletin*, 129(1), 74.
- Djamasbi, S. (2014). Eye tracking and web experience. *AIS Transactions on Human-Computer Interaction*, 6(2), 37-54.
- Duchowski, A. (2007). *Eye tracking methodology: Theory and practice (Vol. 373)*: Springer Science & Business Media.
- Ekman, P. (2009). *Telling lies: Clues to deceit in the marketplace, politics, and marriage (revised edition)*: WW Norton & Company.
- Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry*, 32(1), 88-106.
- Fuller, C. M., Biros, D. P., Burgoon, J., & Nunamaker, J. (2013). An examination and validation of linguistic constructs for studying high-stakes deception. *Group Decision and Negotiation*, 22(1), 117-134.
- George, J. F., Gupta, M., Giordano, G., Mills, A. M., Tennant, V. M., & Lewis, C. C. (2018). THE EFFECTS OF COMMUNICATION MEDIA AND CULTURE ON DECEPTION DETECTION ACCURACY. *MIS quarterly*, 42(2), 551-A559. doi:10.25300/MISQ/2018/13215
- George, J. F., & Robb, A. (2008). Deception and computer-mediated communication in daily

life. *Communication Reports*, 21(2), 92-103.

Hancock, J. T., Curry, L. E., Goorha, S., & Woodworth, M. (2007). On lying and being lied to: A linguistic analysis of deception in computer-mediated communication. *Discourse Processes*, 45(1), 1-23.

Hauch, V., Blandón-Gitlin, I., Masip, J., & Sporer, S. L. (2015). Are Computers Effective Lie Detectors? A Meta-Analysis of Linguistic Cues to Deception. *Personality and social psychology Review*, 19(4), 307-342. doi:10.1177/1088868314556539

Huey, E. B. (1908). *The psychology and pedagogy of reading*: The Macmillan Company.

Itkowitz, C. (2016, November 18, 2016). Fake news on Facebook is a real problem. These college students came up with a fix in 36 hours. *The Washington Post*. Retrieved from [https://www.washingtonpost.com/news/inspired-life/wp/2016/11/18/fake-news-on-facebook-is-a-real-problem-these-college-students-came-up-with-a-fix/?utm\\_term=.965308e16557](https://www.washingtonpost.com/news/inspired-life/wp/2016/11/18/fake-news-on-facebook-is-a-real-problem-these-college-students-came-up-with-a-fix/?utm_term=.965308e16557)

Köhnken, G. (2004). Statement validity analysis and the detection of the truth. *The detection of deception in forensic contexts*, 41-63.

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Attention, Perception, & Psychophysics*, 17(6), 578-586.

McConkie, G. W., & Rayner, K. (1976). Asymmetry of the perceptual span in reading. *Bulletin of the psychonomic society*, 8(5), 365-368.

McConkie, G. W., & Zola, D. (1987). Visual attention during eye fixations while reading.

McCornack, S. A., & Levine, T. R. (1990). When lovers become leery: The relationship between suspicion and accuracy in detecting deception. *Communications Monographs*, 57(3), 219-230.

Metzger, M. J. (2007). Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology*, 58(13), 2078-2091.

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin*, 29(5), 665-675.

Poole, A., & Ball, L. J. (2006). Eye tracking in HCI and usability research. *Encyclopedia of human computer interaction*, 1, 211-219.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, 124(3), 372.



Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The quarterly journal of experimental psychology*, 62(8), 1457-1506.

Rayner, K., & Duffy, S. A. (1988). On-line comprehension processes and eye movements in reading. *Reading research: Advances in theory and practice*, 13-66.

Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements? *Vision research*, 16(8), 829-837.

Rayner, K., Pollatsek, A., Ashby, J., & Clifton Jr, C. (2012). *Psychology of reading*: Psychology Press.

Rayner, K., Well, A. D., Pollatsek, A., & Bertera, J. H. (1982). The availability of useful information to the right of fixation in reading. *Attention, Perception, & Psychophysics*, 31(6), 537-550.

Steller, M., & Koehnken, G. (1989). Criteria-based statement analysis *Psychological methods in criminal investigation and evidence* (pp. 217-245). New York, NY, US: Springer Publishing Co.

Undeutsch, U. (1989). The development of statement reality analysis *Credibility assessment* (pp. 101-119): Springer.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*: John Wiley & Sons.

Vrij, A. (2015). Verbal Lie Detection tools: Statement validity analysis, reality monitoring and scientific content analysis. *Detecting Deception: Current Challenges and Cognitive Approaches*, John Wiley & Sons.

Zhou, L., Burgoon, J. K., Nunamaker, J. F., & Twitchell, D. (2004a). Automating linguistics-based cues for detecting deception in text-based asynchronous computer-mediated communications. *Group Decision and Negotiation*, 13(1), 81-106.

Zhou, L., Burgoon, J. K., Twitchell, D. P., Qin, T., & Nunamaker Jr, J. F. (2004b). A comparison of classification methods for predicting deception in computer-mediated communication. *Journal of Management Information Systems*, 20(4), 139-166.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. *Advances in experimental social psychology*, 14, 1-59.

## Appendix A. Definitions of Linguistic Cues to Deception (Hauch et al., 2015)

	Linguistic cue	Final operational definition
01	Word quantity // word count // number of words // productivity	Total number of words
02	Content word diversity // diversity // content diversity	Total number of different content words divided by total number of content words, where content words express lexical meaning
03	Type-token ratio // unique words // lexical diversity // different words	% of distinct words divided by total number of words
04	Six-letter words // percentage words longer than six letters	% of words that are longer than six letters
05	Average word length (AWL; complexity) //	Total number of letters divided by the total number of words
06	Verb quantity // verb count	Total number of verbs
07	Sentence quantity // number of sentences	Total number of sentences
08	Average sentence length (complexity measure) // words per sentence	Total number of words divided by total numbers of sentences
09	<b>Causation</b>	% of words that try to assign a cause to whatever the person is describing (e.g., because, effect, hence)
10	<b>Exclusive</b>	% of words that make a distinction what is in a category and what is not (e.g., without, except, but)
11	Writing errors // typographical error ratio (informality) // typo ratio // misspelled words	% of writing errors or misspelled words divided by number of words
12	Tentative	% of tentative words (e.g., maybe, perhaps, see)
13	Modal verbs // uncertainty // discrepancy	% of modal verbs or auxiliary verbs or words expressing uncertainty (e.g., should, would, could)
14	<b>Certainty</b>	% of words that express certainty (e.g., always, never)
15	Emotions // emotional / affective processes // affect (ratio) // positive and negative affect	% of words that express any type of emotions/affects (e.g., happy, ugly, bitter)
16	Pleasantness and unpleasantness	% of words that express pleasantness/unpleasantness
17	<b>Negations // less positive tone // spontaneous negations // negation connectives</b>	% of words that express negations (e.g., no, never, not)
18	<b>Negative emotions // negative affect // anger // anxiety, fear // sadness</b>	% of words that express negative emotion/affect (e.g., hate, worthless, enemy) AND anger (e.g., hate, kill, annoyed) AND anxiety (e.g., worried, fearful, nervous) AND sadness (e.g., crying, grief, sad)
18.1	Negative emotions (only) // negative affect	% of words that express negative emotion/affect (e.g., hate, worthless, enemy)
18.2	Anger	% of words that express anger (e.g., hate, kill, annoyed)
18.3	Anxiety	% of words that express anxiety (e.g., worried, fearful, nervous)
18.4	Sadness	% of words that express sadness (e.g., crying, grief, sad)
19	Positive emotions and feelings // positive emotions // positive affects // positive feelings	% of words that express positive emotion/affect (e.g., happy, pretty, good) AND positive feelings (e.g., joy, love)
19.1	Positive emotions (only) // positive affect	% of words that express positive emotion/affect (e.g., happy, pretty, good)
19.2	Positive feelings (only)	% of words that express positive feelings (e.g., joy, love)
20	Total pronouns // personal pronouns	% of all personal (e.g., I, our, they) or total pronouns (e.g., that, somebody, the)
21	First-person singular	% of first-person singular pronouns (e.g., I, my, me)
22	First-person plural	% of first-person plural pronouns (e.g., we, us, our)

	Linguistic cue	Final operational definition
23	<b>Total first-person</b>	% of first-person singular and first-person plural pronouns (e.g., I, we, me)
24	Total second-person	% of second-person pronouns (e.g., you, you'll)
25	<b>Total third-person // other references // third person singular // third-person plural</b>	% of third-person pronouns (e.g., she, their, them)
26	<b>Passive voice verbs // verbal nonimmediacy</b>	% of passive voice verbs (e.g., "it was searched for")
27	<b>Generalizing terms // leveling terms</b>	% of generalizing terms (e.g., everybody, all, anybody)
28	<b>Sensory-perceptual processes // perceptual processes/information // perceptions and sense // sensory ratio // see // hear // feel</b>	% of words that express sensory-perceptual processes (e.g., taste, touch, feel) AND visual (e.g., view, saw, seen) AND haptical (e.g., feels, touch) AND aural (e.g., listen, hearing) sensory-perceptual processes
28.1	Sensory-perceptual processes (only) // perceptual processes // perceptual information // perceptions and sense // sensory ratio	% of words that express sensory-perceptual processes (e.g., taste, touch, feel)
28.2	Seeing	% of words that express visual sensory-perceptual processes (e.g., view, saw, seen)
28.3	Feeling	% of words that express tactile sensory-perceptual processes (e.g., feels, touch)
28.4	Hearing	% of words that express aural sensory-perceptual processes (e.g., listen, hearing)
29	<b>Time // temporal ratio // temporal specificity // temporal cohesion</b>	% of temporal words (e.g., hour, day, o'clock)
30	<b>Space // spatial terms // spatial ratio // spatial specificity // spatial cohesion</b>	% of spatial words (e.g., around, over, up)
31	Temporal-spatial terms // temporal and spatial details total // spatio-temporal information // space and time	% of temporal (e.g., hour, day, o'clock) AND spatial words (e.g., around, over, up)
32	Prepositions	% of prepositions (e.g., on, to, from)
33	Numbers	% of numbers (e.g., first, one, thousand)
34	Quantifier	% of quantifier (e.g., all, bit, few, less)
35	Modifiers (adverbs and adjectives) // rate of adjectives and adverbs (specificity and expressiveness)	% of modifier: adverbs and adjectives (e.g., here, much, few, very)
36	<b>Motion verbs // motion terms</b>	% of words that describe movements (e.g., walk, move, go)
37	<b>Cognitive processes // all connectives</b>	% of words related to cognitive processes (e.g., cause, know, ought)
38	<b>Insight</b>	% of words related to a person's insight (e.g., think, know, consider)

Linguistic cues in bold were used in our analysis

**Appendix B. Stimuli Statements****Statement #1**

I arrived at about a little after 10. At Jesse alcohol was already there - it was vodka, and beer. There was also Sunny Delight. There were people in the household... kids and teenagers... I did not see adults... About the alcohol... I did drink a little vodka and Sunny D, I do not know exactly who purchased... I did hear the mother bought it. I also heard someone paid an adult to purchase the beer and vodka from a store. People I saw - Jim, John, Jesse. I did recognize a girl from school... I think its Julie... She left early. I dropped off John about 12. I went back to Jesse's house. My friend picked me up and I stayed at her house. I had already planned at staying at her house. I did see people drinking in the kitchen where the alcohol was.

**Statement #2**

I went to put fluid in my zippo lighter next to my wall locker. Afterwards there was some fluid on the floor from where I had spilled it. I light the fluid to evaporate it off of the floor and it got out of control and started to singe the carpet. When I noticed black smoke I stopped the flame out and by that that time the smoke detector went out. I proceeded to run to the day room and call the Tinker Fire Department to tell them a false alarm was going off and I needed my smoke detector turned off. The Fire Department came, fixed the detector and a little later the SP came to see the damage.

**Statement #3**

On NOV 8th there was a going away party planed for Jesse and Jim. I am at this time a resident of the house and in the planning part it was agreed that there would be no underaged drinking! Due to me not having eaten all day while I started drinking early, before

the planned barbeque and the fact that I do not drink much anyway, I [last] was in my room passed out before anyone underage arrived, to the best of my recollection. Myself and the other residents; John, and Justin (Army Res.) did make it a point to talk to as many personnel as possible again that there would be no underage drinking before the party started. This all was talked about days before the party.

#### **Statement #4**

On this date, 11 June 1999 at 12:28, Jesse interviewed me regarding inappropriate material on a government computer. I explained to Jesse that I had been to a site regarding Rocky Point that I had seen pornographic material at. I also stated that after seeing this material on one part of the site, I was sure to avoid that area. I also explained to Jesse that members of the fire dept would jump onto someone's logon that had been left on when receiving a call and set inappropriate pictures on that person's background. I explained that this has happened to me on occasion, and that computers regularly get left logged on. I stated that I had never intentionally searched for pornographic material on a government computer, nor had I ever witnessed other members on the fire dept doing so.

#### **Statement #5**

On December 29, '94 around 4:00 in the morning I was intoxicated when I showed up to the armorer. I got my M4 from the armorer because I was supposed to shoot. When I got the M4 I walked up to the clearing barrel. When I was there Sgt. Jesse came up to me and asked me if I was drinking. I told him no. After that I walked outside and then Sgt Jill asked me if I was drinking then I told her yes. Then she took my M4. After that they called Flight Chief. When they showed up they started to give me a test. After the test they then brought me to the front desk and gave me a breath test.

**Statement #6**

I, Jack, walked into the BX on Sept. 13th shopping with my mother and aunt Jennifer. After shopping around and picking out a couple of shirts I proceeded to look for a buggy to put stuff into. After deciding to carry my merchandise instead I carried the shirts on my shoulder and I placed a bottle of Aqua Digio (cologne) in my cargo pocket so I could have my hands freed. Then I went to look for a wallet and found one I liked so I put it into my pocket also planning to pay for before I left but the things in my pocket I forgot about.

**Statement #7**

10 Sep 2005-Everyday for the past month whenever I am in my parking lot or around that area there is a group of Navy guys in their dorms screaming and hollering at me. They also whistle at me. They say things like, "come over here, we won't bite," "you're hot," and tonight they said "Hey baby! I want to fuck you!" I have not said anything to them because I didn't want them to go any farther. The group of guys is always the same and there are usually 3-8 guys out there. I feel very uncomfortable walking in the parking lot by myself when they are out there for that is the time when they usually do it. Until tonight and earlier today when I was with Jesse (organization). Tonight they also said something to him. But I did not hear what was said.

**Statement #8**

I, Jane on November 2, 2005 saw a Caucasian female with brownish blonde hair and glasses come through my register (3) with purchases and a bike. At the time of purchase around 14:13 she had the bike and I wasn't aware that it was a stolen item to which I failed to ask if she was going to purchase it. As well as the woman didn't tell me that she wanted to buy the bike. Once done she told me that she had to go to customer service and that she

needed to go to the i.d. stand. Once she left my register I was soon told that the bike was stolen.

### **Statement #9**

I started playing spades with a woman named Jennifer from Houston, Texas. My spades name is Screenname. One time I mentioned that I had to go and tame my dragon. She replied, "Can I help you with that?" I said sure (kinda flirting with her). She asked if I could send a picture. I said maybe at work, because I didn't want my wife to know. I have never done that before, and for some dumb reason I did. I had her email address loaded on my computer under contacts. So I clicked on her name. How it got sent to XXXXXX I don't know. I accept full responsibility.

### **Statement #10**

On Dec. 27 2005 my husband and I got into an argument about the house not being clean, and the fact that he is off all day and I work 8 or 9 hrs a day. We were in the kitchen I was washing a pan he was unloading the dish washer. One of the pans had water on it and it splashed water on me. Since I was washing a pan I splashed him back. After that he hit me in the back with a small cake pan when I turned around. I sprayed him with the water and he hit me in the face with the cake pan. I called a friend to come get him and locked myself in the bedroom until they got here.

### **Statement #11**

I, Jack, on 15 June was interviewed by Jesse and Jim about pornography on a computer under my logon name. I have in the past left on computers while I still logged on and occasionally still make that mistake. I did not go to pornographic websites on any computer on base. I also do not know who visited any of these websites on the computer I previously used. I am sorry that this has happened and will make sure it will never happen

again. From this I have learned to make sure I am properly and fully logged out of a computer before I leave.

### **Statement #12**

On 20 July 2006, I, Jack, was shadowing Jesse as part of the XXXXX program of USAFA. During my time there I was instructed on the basic responsibilities of the Honor Guard. I witnessed the firing line, flag folding, as well as, basic discipline. I saw the bearing line, push-ups, and a thing called the floppy cock. The floppy cock was where a certain individual (an airman from VA) was told to hold what looked like a giant hot glue stick in his hands. When told to do this the airman was reluctant to perform the asked task from Jesse. He wasn't forced to perform the task in my opinion. There seemed to be an atmosphere of joking and brotherhood, nothing out of the ordinary. In my opinion if asked the same thing at the Academy I would do it and be done with it. There would be no damage physically or mentally to myself.



## Appendix C. IRB Approval

IOWA STATE UNIVERSITY  
OF SCIENCE AND TECHNOLOGY

Institutional Review Board  
Office for Responsible Research  
Vice President for Research  
2420 Lincoln Way, Suite 202  
Ames, Iowa 50014  
515 294-4566

**Date:** 10/5/2017

**To:** Akmal Mirsadikov  
3235 Gerdin Business Bldg

**CC:** Dr. Joey F George  
2340 Gerdin Business Bldg

**From:** Office for Responsible Research

**Title:** Determining Indicators of Deception in Written Text with Eye Tracking

**IRB ID:** 17-410

**Approval Date:** 10/5/2017      **Date for Continuing Review:** 10/4/2019

**Submission Type:** New      **Review Type:** Expedited

The project referenced above has received approval from the Institutional Review Board (IRB) at Iowa State University according to the dates shown above. Please refer to the IRB ID number shown above in all correspondence regarding this study.

To ensure compliance with federal regulations (45 CFR 46 & 21 CFR 56), please be sure to:

- **Use only the approved study materials** in your research, including the recruitment materials and informed consent documents that have the IRB approval stamp.
- **Retain signed informed consent documents for 3 years after the close of the study**, when documented consent is required.
- **Obtain IRB approval prior to implementing any changes** to the study by submitting a Modification Form for Non-Exempt Research or Amendment for Personnel Changes form, as necessary.
- **Immediately inform the IRB of (1) all serious and/or unexpected adverse experiences** involving risks to subjects or others; and (2) any other unanticipated problems involving risks to subjects or others.
- **Stop all research activity if IRB approval lapses**, unless continuation is necessary to prevent harm to research participants. Research activity can resume once IRB approval is reestablished.
- **Complete a new continuing review form** at least three to four weeks prior to the date for continuing review as noted above to provide sufficient time for the IRB to review and approve continuation of the study. We will send a courtesy reminder as this date approaches.

Please be aware that IRB approval means that you have met the requirements of federal regulations and ISU policies governing human subjects research. **Approval from other entities may also be needed.** For example, access to data from private records (e.g. student, medical, or employment records, etc.) that are protected by FERPA, HIPAA, or other confidentiality policies requires permission from the holders of those records. Similarly, for research conducted in institutions other than ISU (e.g., schools, other colleges or universities, medical facilities, companies, etc.), investigators must obtain permission from the institution(s) as required by their policies. **IRB approval in no way implies or guarantees that permission from these other entities will be granted.**

Upon completion of the project, please submit a Project Closure Form to the Office for Responsible Research, 202 Kingland, to officially close the project.

Please don't hesitate to contact us if you have questions or concerns at 515-294-4566 or [IRB@iastate.edu](mailto:IRB@iastate.edu).

## CHAPTER 4. CONCLUSION

The goal of this study was to explore the relationships between media, message veracity, and deception detection. We investigated the relationship between them through the visual attention of the veracity judges employing eye tracking technology. Specifically, we sought to understand how media modes and message veracity influenced the visual foci of the veracity judges and how these foci affected deception detection. The research model in Figure 4-1 guided our work.

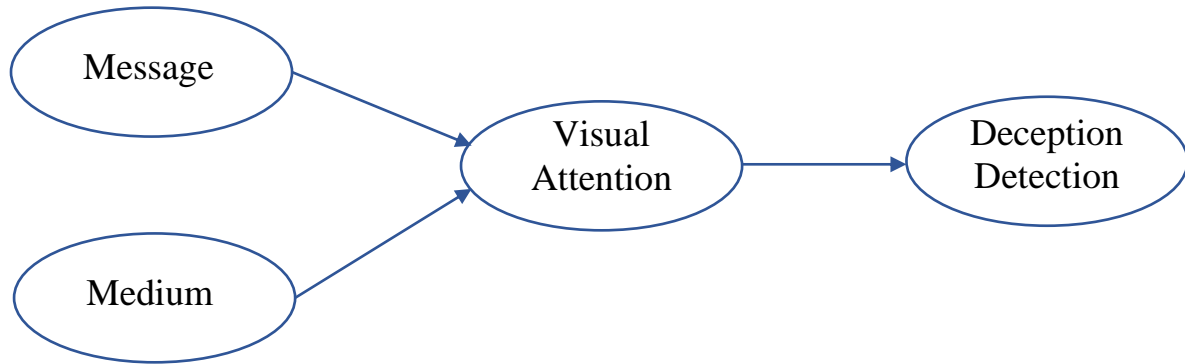


Figure 4-1. Research model

This dissertation is comprised of two studies, each addressing specific research questions. The first research study was guided by three research questions. We sought to understand how, if at all, media played a role in drawing the visual attention of veracity judges. We also investigated what the judges look at when presented with honest and dishonest messages. Finally, we sought to understand the link between what the judges looked at and how successful they were in correctly classifying the messages. In the first research study, presenting honest and dishonest video stimuli and varying the audio availability of the video stimuli, we found that both media and the nature of the message had an effect on where people tended to focus. We also found very slight support for the association between the visual foci and detection success.

We investigated the gaze behavior of the judges on three areas, namely, the senders' eyes, mouth, and torso. Deceptive messages resulted in longer and more frequent fixations on

the sender's mouth and torso. There was no significant difference in gaze behavior on the senders' eyes. We also found support for the effect of media on the visual foci of the judges: judges looked longer at the mouth of the speakers when the sound was absent from the image. Moreover, when presented with full audio-visual stimuli, the judges fixated more frequently on the eyes of the speakers. These findings provide support for the associations between the message type, media, and visual attention presented in Figure 1. We further found that gaze fixation frequencies on the sender's eyes and mouth significantly diminished the deception detection success of the veracity judges. Last, we found no significant support for the mediating role of the visual attention. Thus, in the first study we were able to explore the relationships of the whole research model.

In the second study we sought to understand deception in written text and how veracity judges were able to discriminate between honest and dishonest written statements. The second study was guided by two research questions, namely: (1) what veracity judges focused on when presented with honest and deceptive statements, and (2) how the visual elements of written statements influenced their deception detection success. To address the research questions, we used real-life written statements produced in high-stake settings with potentially severe consequences. We identified 15 linguistic cues grounded in theories of deception and analyzed how the judges processed them when reading. As predicted, the judges fixated longer when reading deceptive statements. Furthermore, they fixated longer and more frequently on verbs describing motion, on passive voice verbs, on words describing spatial information of an event (spatial context), and on word referencing insight. Further, our findings suggest that fixation durations and counts on words describing senses, on generalizing terms, and on certainty words worsened detection accuracy. Fixation durations

and frequency on third person pronouns, fixation frequency on negations, and total fixation counts improved deception detection rates.

Empirical evidence from these two studies provide support for the associations put forth by the research model. Understanding the process of deception detection through the point of regard of a veracity judge sheds light not only on the present indicators of deception, but rather on those that are more noticeable than other indicators.