REAL-WORLD, HIGH-STAKES DECEPTIVE SPEECH: THEORETICAL VALIDATION AND AN EXAMINATION OF ITS POTENTIAL FOR DETECTION AUTOMATION

By

JOSEPH YORK THOMAS

Bachelor of Science in Computer Science
Texas State University
San Marcos, Texas
1993

Master of Arts in Information Resource Management
Webster University
St. Louis, Missouri
2003

Submitted to Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 2014

REAL-WORLD, HIGH-STAKES DECEPTIVE SPEECH:

THEORETICAL VALIDATION AND AN

EXAMINATION OF ITS POTENTIAL FOR

DETECTION AUTOMATION

Dissertation Approved:

Ramesh Sharda

Dissertation Adviser

David P. Biros

Mark Weiser

Ronald Thrasher

Outside Committee Member

ACKNOWLEDGMENTS

I would like to thank my dissertation advisor, Dr. Ramesh Sharda, for his sage guidance and encouragement. I would also like to thank my committee members, Dr. Mark Weiser and Dr. Ronald Thrasher for their support and valuable suggestions throughout this study. Many thanks go out to Dr. Carla Goad and Dr. Dale Fuqua not only for their time and patients but for steering me through the murky waters of statistics and research methods. This dissertation would not have been possible without the help of two special people, Dr. David Biros, my research advisor, and Detective Maureen Wall of the Madison Police Department, the source of my data. Thank you Dr. Biros for your guidance and for introducing me to such a rich research topic, I look forward to our future collaborations with vehemence. Almost as important as the advice is the data (some may argue even more so). My deepest gratitude goes out to Det. Wall for not only giving me access to your data but for painstakingly reviewing transcripts and providing your expert opinion.

A special thank you goes out to Grand Master Roy and Janette Kurban. At a very young age you instilled in me a passion for excellence and a never quit attitude, the ripples of which can be seen here, 30 years later. Finally, I am extremely thankful that God has blessed me with an understanding and supportive wife that seems to have been tailor-made for a military officer trying to earn a PhD.

Name: JOSEPH YORK THOMAS

Date of Degree: DECEMBER, 2014

Title of Study: REAL-WORLD, HIGH-STAKES DECEPTIVE SPEECH: THEORETICAL VALIDATION AND AN EXAMINATION OF ITS POTENTIAL FOR DETECTION AUTOMATION

Major Field: BUSINESS ADMINISTRATION, MANAGEMENT INFORMACTION SYSTEMS

Abstract: The study of deception and the theories which have been developed have relied heavily on laboratory experiments, in controlled environments, utilizing American college students, participating in mock scenarios. The goal of this study was to validate previous deception research in a real-world high-stakes environment. An additional focus of this study was the development of procedures to process data (e.g. video or audio recordings) from real-world environments in such a manner that behavioral measures can be extracted and analyzed. This study utilized previously confirmed speech cues and constructs to deception in an attempt to validate a leading deception theory, Interpersonal Deception Theory (IDT). Several measures and constructs, utilized and validated in existing research, were explored and validated in this study. The data analyzed came from an adjudicated real-world high-stakes criminal case in which the subject was sentenced in federal court to 470 years in prison for creating child pornography, rape, sexual exploitation of children, child sexual assault and kidnapping; a crime spree that spanned over a five years and four states. The results did validate IDT with mixed results on individual measures and their constructs. The exploratory nature of the study, the volume of data, and the numerous methods of analysis used generated many possibilities for future research.

TABLE OF CONTENTS

Chapter **Page**

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figures **Page**

CHAPTER 1

INTRODUCTION

Deception is a ubiquitous form of communication (Handcock, Woodworth, & Goorha, 2010). Several studies have looked at how frequently people lie and they indicate that on average, people lie almost daily (Camden, Motley, & Wilson, 1984; DePaulo & Kashy, 1998; DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996; Hample, 1980; Lippard,1988). In fact deception is a major characteristic of the most common communication channels; 14% of people self-reported deceiving in emails, 37% in phone calls, and 27% in face-to-face interactions (Hancock, 2007). In a study by Turner, Edgley, and Olmstead (1975) they asked 130 subjects to keep a record of whether they were truthful or not in their communications throughout their day and found that over 61% of their subjects' conversations lacked verisimilitude.

Just as prevalent as deception is the desire to detect deceit in human communication. Despite the desire, humans are rather poor at detecting deception, scoring no better than chance (Vrij, Edward, Roberts, & Bull, 2000). Current research on methods to assist humans at deception detection have been to a great extent focused on controlled settings where the researcher manipulates the study. More to the point, it is impossible to ethically replicate an environment in which subjects truly have significant potential for loss or suffering and must deceive in order to evade high-stake loss. It is not possible to simulate real risk where deceivers have a vested interest in the

deception because they have something significant to gain or lose. A second aspect of deception not possible in a lab is contextual complexity. It is exceedingly difficult to replicate the duration and complexity typical of high-stakes dyadic interactions, such as those which take place during law enforcement interviews. To be of practical use current methods and theories need to be validated in the real-world where it is not always possible to control the stimuli and intrude on a subject's freewill to deceive.

In general studies of deception inevitably focus on behavioral indicators or cues to deception. These cues can be categorized into two groups, verbal and nonverbal. Verbal cues include written and spoken words and focus on the linguistic characteristics of speech (Burgoon, Buller, and Woodall, 1996). Nonverbal cues include physical ways of communication, including paralinguistic (e.g. tone of the voice), haptic (e.g. touch), and kinetic (e.g. body motion) (Burgoon, et al., 1996; Schuller, Steidl, Batliner, Burkhardt, Devillers, Müller, & Narayanan, 2013).

Speech is a very rich communication channel, one that contains both verbal and nonverbal characteristics. Speech is also an interactive communication channel with many back and forth exchanges giving both parties multiple opportunities to influence each other in a short period of time (Buller & Burgoon, 1996).

This study is an exploration of real-world, high-stakes (RWHS) deceptive behavior manifested in human speech, and analyzed by objective measures. It is worth saying that this study is not a laboratory experiment, with controlled settings, in a closed environment. Though several statistical tools were employed and every opportunity to follow sound methodology was practiced, their use was not to prove or disprove

hypotheses but to explore the data and examine propositions based on theory. The impetus for this study came after a lengthy literature review on deception detection and has three tenets: (1) the state of existing theories on deception crave for validation, (2) outside the lab in a RWHS setting, (3) where typical dyadic interactions are long and more complex than those studied in a controlled setting. These tenets are the research gaps identified and where it is believed the most stands to be gained by exploration.

Human Performance

The formal study of deception detection and its cues has been covered in numerous cross discipline studies and the consensus across the board is that humans are poor detectors of deceit (Bond & DePaulo, 2006; Kraut, 1980; Vrij, Edward, Roberts, & Bull, 2000). The most common approach to studying deception detection is to look for signs in the form of behavioral changes; these signs are typically referred to as cues. In perhaps the most comprehensive meta-analysis of deception detection cues and their accuracy, Bond and DePaulo (2006), looked at 206 studies with 24,483 judgments and found a mean accuracy of 53.4%. To be more colloquial, humans might as well flip a coin when it comes to detecting deception. However, humans are not just inaccurate detectors of deceit but poor judges of what cues are indicators of deception and are often affected by multiple biases (Vrij, 2000). Human bias toward unreliable deceptive cues hampers our ability to perceive deception and can further decrease accuracy below chance (Vrij, 2000).

Deception researchers have confirmed that the general population has a "truth bias" or regularly assume others are truthful in typical interpersonal encounters (Buller & Hunsaker, 1995; Buller, Strzyzewski, & Hunsaker, 1991; Clark, H. & Clark, E., 1977; Kalbfleisch, 1992; O'Sullivan, Ekman, & Friesen, 1988; Riggio, Tucker, & Throckmorton, 1987). This over willingness to trust may further hinder our deception detection ability. In contrast, law enforcement personnel are significantly more likely to respond "deceitful" rather than "truthful" in a suspect interview session (Meissner & Kassin 2002). Meissner and Kassin call this "investigator bias". Law enforcement personnel who conduct interviews are trained in techniques on how to spot deceptive cues; some law enforcement personnel such as FBI, CIA, Secret Service, and Homeland Security agents receive more advanced training in human behavior. However, even with advanced training human performance at detecting deception is not significantly increased (Akehurst, Bull, Vrij, & Koehnken, 2004; Biros, Hass, Wiers, Twitchell, Adkins, Burgoon, & Nunamaker, 2005; Vrij, 2000). It has also been pointed out that law enforcement training on deception detection cues could be wrong and biased towards visual cues (Kassin & Fong, 1999; Mann, Vrij, & Bull, 2004; Vrij, Edward, & Bull, 2001). In an interesting study by Levine, Freeley, McCornack, Hughes, and Harms (2005) they compared valid deception detection training to bogus training and no training. They found that the effects of both valid training and bogus training to be generally small and that valid training was not much better than bogus training. They also suggest that even bogus training could improve accuracy over no training. This last point would connote that the content of the training may be completely irrelevant, a piquant point considering that content was based on current theories of deception

detection. These same theories are also based mostly on laboratory settings, have had few validations in real-world environments, and where subjects had very little at stake (DePaulo et al., 2003).

Real-World High-Stakes

Previous research has found that cues to deception differ when the motivation differs from low to high as is often the case when comparing mock-scenarios to real-world events (Zuckerman & Driver, 1987). RWHS deceptive environments are those were the parties are personally invested in the exchange and highly motivated to deceive or detect, depending on their role. The world has seen a dramatic increase in real-world high-stakes environments since Sept 11, 2001. From 2002 to 2011, terrorist attacks around the world increased 460% (Institute for Economics & Peace, Global Terrorism Index [GTI], 2012). In 2011 alone 91% of terrorist attacks were successful resulting in 7473 deaths and 13,961 injuries (GTI, 2012). Increases in high-stakes environments caused by threats of terrorism has resulted in more screening at airports, border crossings, and other key social and environmental choke points, driving a compelling reason for better deception detection tool development (DePaulo et al. 2003; Elkins, Burgoon, & Nunamaker, 2012; Frank & Ekman 1997). According to the Department of Homeland Security (DHS), "To a greater degree than at any point in history, individuals and small groups…have the ability to engage the world with far-reaching effects, including those that are disruptive and destructive" (DHS, Quadrennial Homeland Security Review Report, 2010). At the federal level there is also a push to mandate the electronic

recording of all interviews involving felony custodial cases (Sullivan, 2010, 2008). This and many other trends toward information storage and data warehousing will increase the available data sets that need to be analyzed for their credibility as well as increasing the need for tools to guide investigators and lawyers during follow-up questioning. However, real-world deception detection research must overcome the wicked problem of establishing ground truth (Iacono, 1991).

Ground Truth

Ground truth is a verified or indisputable fact, for example adhering to evidentiary guidelines used in a court of law. In a laboratory setting, establishing ground truth is a matter of experimental design, fully controlled by the researcher. This same control is not possible in the real-world and to attempt to subject people to real stressors that would lead up to deceptive communication would be unethical and most likely illegal (e.g. ask a student to steal a computer from the schools lab and then monitor them during police interviews). In addition, random assignment of participants to treatment groups is not possible in field studies. These issues and other RWHS deception detection research is covered in greater detail in Chapter 2. What follows is an introduction to the principle theory the current research is based on Interpersonal Deception Theory by Buller and Burgoon (1996) and the role suspicion plays in deception detection.

Interpersonal Deception Theory & Suspicion

According to Interpersonal Deception Theory (IDT) (Buller and Burgoon, 1996), the counterpart to senders' deception is receivers' suspicion. IDT suggests that deception is a dyadic interaction and as the deception takes place, receivers may become suspicious of the senders attempts to deceive and may adapt their behavior because of it. For example, they may choose to conceal their suspicion by quickly moving on to another topic or admit their suspicion and confront the sender to gauge their reaction. Buller engaged in two studies that suggest that when people suspect someone is lying to them, they attempt to conceal their suspicion by altering their behavior (Buller, Strzyzewski & Comstock, 1991; Burgoon, Buller, Dillman, & Walther, 1995). Whether or not a receiver attempts to conceal their suspicion, their behavior or lack thereof, may affect the senders behavior (i.e., deceivers may try even harder to be convincing if they believe the receiver is suspicious).

At this point it is not clear whether suspicion plays a consistent role in deception. Therefore it is a goal of this study to better understand the impact suspicion has in a RWHS setting. In addition to suspicion, IDT suggests the dyadic interaction between sender and receiver may change over time. The following section will introduce the topic of deception over time.

Deception Over Time

The subject of deception itself is difficult to study because of the complex nature of human interactions and the dynamic interchange that takes place over time. One facet

of complexity in human interaction over time is strategic vs. nonstrategic behavior (Buller & Burgoon, 1994; Zuckerman and Driver, 1987).

Humans are strategic in their interactions and attempt to monitor multiple channels of communication with the intent of adapting their behavior based on how well they perceive their message is accepted (Buller & Burgoon, 1994). For example, a car salesman may change sales strategy and back off a potential customer based on their initial interaction. During a deceptive dyadic communication, deceivers often take into account not only what and how they convey their message but how the receiver's behavior changes. Deceivers then attempt to modify their behavior, both strategically and non-strategically, in response to perceived suspicion from the receiver in order to achieve their communication goal (Buller & Burgoon, 1994; Zuckerman and Driver, 1987).

In addition to the impact strategic and nonstrategic stratagems have on deceptive behavior, human behavior is also influenced to a large degree by whether the liar (sender) is telling a prepared lie or a spontaneous lie (Cody, Marston, Foster, 1984; Cody & O'Hair, 1983). When comparing a prepared lie to a spontaneous lie, the prepared liar should be less aroused, have more control, and should not find lying as mentally difficult (Zuckerman & Driver, 1987). In general, spontaneous lies contain more deceptive cues than prepared lies and spontaneous liars make a less credible impression (deTurck & Miller, 1990; Littlepage & Pineault, 1985). Additionally, some researchers have noted an increase in the amount of time it takes for an individual to respond to a question (response latency) in situations involving spontaneous lies (DePaulo, Lindsay, Malone, Muhlenbruck, Charlton, & Cooper, 2003). It stands to reason that as the duration of a

dyadic communication increases, like in extensive law enforcement interviews, it becomes more and more difficult to simultaneously maintain credibility and deception. The following section introduces interpersonal speech, one of the more promising communication channels from which to study deception detection over time.

Case for Speech Communication Channel

Almost every possible communication channel has been studied with regards to deception detection. The nonverbal channel, which includes paralinguistic measures, conveys an incredible amount of information, some of which can be used for deception detection (Buller et al. 1996; Vrij 2000). Linguistic-based cues also show promise as a deception detection diagnostic (Fuller, Biros, Burgoon, & Nunamaker, 2013, Fuller, Marett, & Twitchell, 2012; Fuller, Biros, & Delen, 2011; Ott, Choi, Cardie, & Hancock, 2011; Tower, Jensen, Dunbar, & Elkins, 2013; Zhou, Burgoon, Nunamaker, & Twitchell, 2004). However, relying on any single channel may be more disadvantageous than relying on multiple channels or on the full spectrum of a particular channel, especially when those channels are closely related as with the linguistic and paralinguistic sub-channels of speech.

Audio recordings are commonly used in law enforcement and high-stakes settings for their ease of use and their admissibility in a court of law. As stated earlier, audio is also a strong source of deceptive cues (DePaulo et al. 2003). It is a one-to-many communication channel, namely text-based (linguistic), vocal-based (paralinguistic), and content cues (Buller et al. 1996; Zuckerman et al., 1981). Content cue analysis is beyond

the scope of this work leaving linguistic and paralinguistic channels for analysis. The author believes studying both these communication modes simultaneously will reveal useful information towards deception detection research. This study hopes to find a synergistic effect within the speech channel showing that more cues and more types of cues to deception can be extracted when examining multi-dimensional channels vs. single channels alone. The following section briefly addresses current deception detection methods and their drawbacks; each is covered in greater detail in Chapter 2.

Concerns With Current Methods

Current methods to detect deception all have drawbacks which need to be addressed with additional research. Deception detection methods can be split into two categories, invasive and non-invasive. Of the invasive technologies currently available to help identify and measure deceit, the polygraph is the most well-known. In a summary of laboratory tests, Vrij reports that the polygraph is about 82% accurate at identifying deceivers (Vrij, 2000). Although it is not admissible in a court of law in the United States, the polygraph is useful during investigations for identifying potential suspects or to narrow down possible leads. However, polygraph exams have several strong limiters namely a willing subject, an invasive exam, and the need for a trained examiner. The polygraph exam itself can evoke fear and apprehension in its subjects making it a controversial investigative tool.

The newest invasive method to detect deception utilizes functional magnetic resonance imaging (fMRI) to map blood flow in the brain during structured questioning.

Though initial findings are promising, fMRI shares the same restrictions as the polygraph. Additional limiters to their general use are their sheer size and cost to operate, the fact that subjects cannot move at all, and they cannot be used on people with claustrophobia or metallic implants.

One attempt at a non-invasive, unobtrusive, deception detection technology that would have been easily automated but which failed to stand up to academic rigor was voice stress analysis (VSA). VSA was introduced in the 1970's as a possible replacement of the polygraph (Rice, 1978). VSA is a technique that analyzes the voice pitch changes as a measure of arousal. The machines built around VSA are fundamentally designed to detect stress, not lies. The accuracy of voice stress analyzers is reported to range from chance to about equal to that of the polygraph (Gamer, Rill, Vossel, & Godert, 2006; Vrij, 2005). Despite its initial promise, the voice stress analyzer has failed to gain scientific acceptance (Ford, 2006; Hollien & Harnsberger, 2006; Hopkins, Benincasa, Ratley, & Grieco, 2005). It should be noted that VSA only looks at pitch while the current study examined a much broader range of vocal dimensions.

A common method used in deception detection research involves structured questioning. In fact the preponderance of deception detection research looks at yes/no or short answer lies, like those used during a polygraph exam (DePaulo, Lindsay, Malone, Muhlenbruck, Charlton, & Cooper, 2003). The yes/no response to structured questioning has its value but is not characteristic of RWHS questioning where interviewers typically ask for more details and subjects are more verbose in their responses. Another predominant method in deception detection research looks at mean scores over entire interactions, in essence detecting general deceptive states not pinpointing lies (DePaulo,

et al., 2003). This course granularity of deceptive episodes is of little practical value in the real world, what is needed is to identify the needle (lie) in the hay stack (interview). This is similar to what Meissner and Kassin (2002) call "response bias" or the threshold of evidence necessary for deception detection.

Having given an introduction to the field of deception detection in a RWHS setting what follows is the research questions which attempt to address the gaps in existing research.

Research Question & Contributions

This study's research method takes the form of a case study in a RWHS setting in which the communication channel of interest is speech and its linguistic and paralinguistic dimensions. Based on the review of literature and the theoretical underpinning of IDT, a conceptual framework was developed for understanding these deceptive cues in RWHS environments. From this conceptual framework a research model and series of propositions were developed and then used in an attempt to validate a leading deception theory by examining the following research question (RQ):

RQ: *Are speech cues to deceptive behavior moderated over time by receiver suspicion during dyadic interactions in a real-world high-stakes setting?*

This study aimed to address three areas of deception research in hopes to contribute to: (1) procedures for extracting deceptive cues from real-world environments,

(2) identification of temporal speech patterns, and (3) granularity of deceptive speech. Expanding on each area it is hoped the following will be gained by the study:

- Develop effective procedures for extracting speech-based behavioral cues in RWHS settings. Given the complex and shifting nature of real-world settings is it even conceivable to extract speech behavioral characteristics of high enough quality to make reliable measurements?
- Given an extraction procedure that produces usable data, can the veracity of individual messages be accurately determined using speech-based cues? In other words, how fine is the granularity of deceptive speech behavior? Having a better understanding of the size of a deceptive epoch may help steer other deception detection studies.
- A major contribution hoped to be made with this study is to determine if speech-based cues to deception (both linguistic and paralinguistic) change over time in a RWHS environment and what do these patterns look like.
- Finally, add to the growing body of knowledge trying to answer the question whether deception detection can be automated using information systems tools, techniques, and procedures.

In addition to the research question above, this study aimed to better the understanding of how people deceive when using vocal-base communication methods in RWHS settings. General research questions in the context of a RWHS interview-style dyadic communication that follow from the primary research question are:

- Which speech-based cues distinguish truthful from deceptive messages?

- Can the veracity of individual messages within a larger body of dyadic interactions be accurately determined using speech-based cues?
- Are there deceptive behavioral patterns?
- If there are patterns what do they look like?
- Do cues change over time with respect to sender deception and receivers suspicion?

An additional focus of this study was how the collection, measurement, processing, and management of deceptive cue data is extracted from RWHS environments. The author wished to begin to understand if deception detection can be automated using existing information systems tools and techniques. Finally, it is expected to develop a procedure for extracting speech-based cues from RWHS raw data sources such as video and audio recordings.

Throughout this dissertation the term "deceiver" will also be referred to as the "sender" (e.g. suspect, criminal, etc.) while the person attempting to determine the truth will be referred to as the "receiver" (e.g. researcher, law enforcement investigator, detective, etc.). Chapter 2 gives a brief overview of deception detection literature and the speech channel as a promising source for cues in a RWHS context. Chapter 3 will define the propositions developed based on the literature review, discuss the methodology and measurements used in this research, as well as further justification for utilizing the case study methodology in this context. Chapter 3 will also describe the procedural model used to process the data for analysis. Chapter 4 will cover the results of analysis as well

as findings, and Chapter 5 will conclude the dissertation with a discussion on limitations and recommend future research.

CHAPTER 2

REVIEW OF LITERATURE

The impetus for this study came after a lengthy preliminary literature review on deception detection. During the initial review it became clear that deception detection research contains several gaps, namely: (1) existing theories on deception crave validation (2) that this validation needs to occur outside the lab in a RWHS setting (3) and where typical dyadic interactions are longer and more complex than those studied in a controlled setting. These tenets are where it is believed the most stands to be gained and became the focus of the formal literature review.

Deception & Deception Detection

Deception is defined as "a message knowingly transmitted by a sender to foster a false belief or conclusion by the receiver" (Buller et al. 1996). For as long as people have been lying, people have been trying to detect lies (Ford, 2006). However, as will soon be seen, people's desires far exceed their abilities. With the advancement of new methods and new technology humans are continuing to pursue an automated system to assist them with accurately recognizing deceptive behavioral patterns.

Human Deception Detection Accuracy

When asked, most people believe they can detect when someone is lying to them, but they are misguided. Empirical research shows that human accuracy at detecting deception is little better than chance (Aamodt & Custer, 2006; Bond & DePaulo, 2006; Vrij, 2000). The typical experimental procedure for determining accuracy usually consists of exposing subjects (i.e. receivers) to statements made by potential deceivers (i.e. senders). The channel of communication varies from written statements (Fuller, 2008; Fuller, Biros, & Wilson, 2009) to audio and video (Meservy, 2007) to face-to-face (Jensen, Burgoon & Nunamaker, 2010). It is also true that few studies include interactions between sender and receiver (Buller & Burgoon, 1996). Subjects typically indicate their judgment of trust or deception on a paper form and may be asked to explain what led them to their decision (Masip, Garrido, & Herrero, 2009).

Several meta-analysis studies have been done on deception detection accuracy and found the mean to be 56-57% (Kraut, 1980; Vrij, 2000). Another meta-analysis was performed in which they examined the accuracy of professionals trained to detect lies, for example, police officers, detectives, judges, and psychologists (Aamodt and Custer, 2006). Their analysis reported a mean accuracy of 55.5% compared to 54.2% for students and regular citizens. One explanation for this counter intuitive result may be that, even with professional training, people often rely on misleading cues (Vrij, 2000). In a more recent and perhaps the most extensive meta-study to date, they found a mean accuracy of 53.4% (Bond and DePaulo, 2006). The preponderance of evidence suggests not only that unaided humans are poor at detecting deception but that training has no

significant impact on deception detection accuracy rates (Biros et al. 2002; Levine, Feeley, McCornack, Hughes, & Harms, 2005).

Another possible reason humans are poor at detecting deception is because many of the behaviors perceived as deceptive, are not. For example, many people believe liars gaze less and truth tellers gaze more. However, both Riggio and Friedman (1983) and DePaulo, Stone, and Lassiter, (1985) found that it is liars who tend to gaze more. As already stated, research shows that the average person can detect deception about equal to the flip of a coin. However, other researchers suggest humans are even less accurate than chance (Feeley & Young, 1998; Levine, Park, & McCornack, 1999). When accuracy rates for truthful and deceptive messages are examined separately, people can detect truths just above 50% and deception well below 50% (Levine, et al., 1999).

Despite what the average person's abilities are at detecting deception there is something to be learned from those at the ends of the spectrum. Just as some people are better at deceiving others, some people are more skillful at detecting deception. One aspect of human communication that impacts deception detection accuracy is involvement (Petty & Cacioppo, 1986). Forrest and Feldman's (2000) study suggests that people who are highly involved in a task and focus mainly on verbal cues, are less accurate at detecting deception than people who were not involved and focus on nonverbal cues. In other words, the less involved someone is in an investigation the better at detecting deception they may be. Another characteristic of humans that may improve deception detection is self-monitoring. High self-monitors tend to pick up on deceptive behaviors that other people miss (Brandt, Miller, & Hocking, 1980).

It could be argued that information systems are not involved in the interaction and by a large margin better at self-monitoring because they can measure quantitatively much more than humans are capable. In order to improve their deception detection accuracy humans should exploit the advantages of information systems impartiality and their ability to self-monitor.

Deception Automation

People often communicate simultaneously across multiple channels such as email, video, telephone, and face-to-face. Because of the multi-channel and often digital nature of communication it may be helpful to examine deceptive cues from the lens of management information systems (MIS) research. The amount of information that passes between two people, conscious and unconscious, that must be processed in order to identify deceit is incredible. The typical conversation speed is approximately 150-160 words per minute (wpm) (Williams, 1998) while humans can consciously think linguistically at approximately 400 wpm (Wong, 2014). This difference would seem to beg the question, why did humans evolve to process verbal communication faster than others can speak? One possible answer is that humans need multimodal communication. Multimodal communication has been shown to have a synergistic effect on a message (D'Mello, & Graesser, 2010). Several qualities of speech that communicate nonverbally are tone, volume, pitch, cadence, inflection, rhythm, emotion, accents, and slang, to name a few. Add a kinesthetic communication layer and it becomes obvious, humans process an incredible about of information in simple conversations, let alone a deceptive one.

It is because of the large amount of information that must be considered and analyzed that information systems tools and methods show great promise for improving deception detection by assisting human receivers. As suggested by Carlson and George (2004) electronically encoded media may stand the best chance for automation in support of deception detection and by extension the best channel for law enforcement interview tool development (Bazin, 2010). They based their study on the theories that describe aspects of this change, namely media richness and media synchronicity (Dennis, Fuller, & Valacich, 2008). Carlson and George further state that during this encoding that some of the variety of cues may be lost. Therefore, cue identification is crucial to deception detection, here again technological advancement in information processing and management shows promise. As real-time information processing power permeates to the desktop and handheld levels, new methodologies will need to be developed to take advantage of available communication channels for analysis.

Real-World High-Stakes Deception

There is an abundance of deception detection studies involving mock lies, but for research in high-stakes environments, such as interviews during a criminal investigation, it is lacking (Ben-Shakhar & Elaad, 2003; DePaulo et al., 2003; Frank & Ekman, 1997; Fuller, Marett, & Twitchell, 2012; Porter & Brinke, 2010; ten-Brinke & Porter, 2012). This has driven a strong need for more field studies in deception detection research (Frank & Feeley, 2003). Vrij and his colleagues have performed a number of field studies involving police officers (Mann, Vrij & Bull, 2002; Vrij, 2000, 2005; Vrij & Mann, 2001a, 2001b). They were the first and only (Porter & Brinke, 2010) to code the

nonverbal behavior of criminal deceivers during high-stakes deception. Vrij and Mann (2001a) performed a case study to examine the paralinguistic behavior of a murderer during his police interview prior to and during his confession. When the murderer lied he paused longer, spoke slower, and made more non-ah speech disturbances. Mann, Vrij, and Bull (2002) found similar behavior after examining videotaped police interviews of sixteen convicted criminals; they paused longer and blinked less when lying as compared to their baseline behavior. With few exceptions, the call for more field studies in deception detection research has gone unanswered in the past decade.

As society increasingly depends on on-line communication cyber deception and crime increases as well. The severity of these crimes including human trafficking, child pornography, and fraud has spurred a high interest in methods to detect on-line deception (Hancock, 2007; Toma, Hancock, & Ellison, 2008; Whitty & Carville, 2008). One area that has developed in large part to increases in on-line communication is linguistic analysis. By examining linguistic characteristics, researchers are beginning to understand deceptive behavior in cyberspace interactions. As far as real-world high-stakes linguistic analysis research there are few studies to date (e.g. Fuller, 2008; Fuller, et al, 2009). However, researchers have found during controlled studies that instant messaging deceivers tend to initiate conversations more often, write longer messages, and take shorter breaks between sent messages, than truth tellers (Hancock, Curry, Goorha, & Woodworth, 2008). Deceivers also use fewer self-oriented pronouns, believed to be an attempt to distance themselves from the deception (Hancock, Curry, Goorha, & Woodworth, 2008; Zhou, 2005). An increase in real-world high-stakes cyber deception

detection research is becoming increasingly crucial to maintain trust in digital communication.

An issue with the current state of deception detection research that separates it from real-world settings is the fact that a vast majority of current studies utilize university students instructed to lie in mock scenarios (DePaulo et al., 2003; Vrij, & Mann, 2001b).

Student Subjects & Mock Lies

A principal deception detection meta-analysis of 120 studies showed 101 used student subjects. Only four of these studies (3%) involved situations where the subjects were not given instructions to lie but chose to do so on their own (DePaulo et al., 2003). There is evidence that behavior differs between those who choose to lie and those directed to lie by an experimenter (Feeley & deTurck, 1998). For example, those who chose to lie compared to those instructed to lie made fewer speech errors and hesitations, and fewer references to others. Therefore, studies utilizing real-world samples of subjects who either chose to be deceptive or not may contribute more deeply to the understanding of deception than those studies utilizing mock lie scenarios, as well as provide more generalizable findings. The over use of student subjects and subsequent lack of generalizability to RWHS situations is of utmost importance especially where deception detection is most critical (e.g. law enforcement, military operations, and domestic defense).

In addition to the issues of generalizability caused by using a single class of subjects, the use of mock scenarios has been questioned as to whether they can offer

better understanding into how deception occurs in a real-world setting (Pollina, Dollins, Senter, Krapohl, & Ryan, 2004). Another criticism of mock lies is on the lack of motivation; participants do not chose to lie hence have little or no vested interested in whether or not they get caught (Miller & Stiff, 1993). A lack of personal involvement in the lie is another critique of laboratory studies (Koper & Sahlman, 1991). The author acknowledges the value of laboratory studies but believes there is a need for more research examining real-world high-stakes environments.

In examining RWHS deception it is important to consider the characteristics of a real-world setting that separate them from the controlled setting of the lab. Three such characteristics often examined by researchers are: motivation, duration, and context. Considering the motivation behind deceptive attempts is one of the more common ways to categorize lies (Goffman, 1974). Certainly, there are some times when people are more motivated to lie successfully than others. A fisherman may not care if his friends find out that the trophy fish that got away was really a fingerling. A shoplifter, however, might have more at stake if his larceny was discovered. Several studies have concluded that telling high-stakes lies motivates people to succeed at their deception and ironically, makes them more detectable (DePaulo, Kirkendol, Tang, & O'Brien, 1988; DePaulo, Lanier, & Davis, 1983; DePaulo, Lane & DePaulo, 1999; LeMay, & Epstein, 1991; Forrest & Feldman, 2000; Frank & Ekman, 1997).

Research has also shown the duration and content of a lie can influence how successful a person can be at deception. Longer lies, for instance, are more difficult to tell than short ones (Kraut, 1978). The idea that longer lies are more complex and difficult to maintain than short and simple lies seems to be common sense. In the meta-

analysis done by DePaulo et al., (2003) they predicted that if deceivers were required to sustain their deception for greater lengths of time, then cues to deception would be clearer and more numerous. Their findings supported their hypothesis; duration did moderate the size of the effect. When interactions were sustained for greater amounts of time, deceptive responses were shorter than truthful ones.

The simplest division of deception may be into benign and exploitive lies. Benign lies, commonly called "white lies" tend to have low perceived stakes, short life expectancies, and are often told for the benefit of others. Exploitive lies, which are motivated by the selfish interests of the deceiver range from simple lies (e.g. calling in sick when you are just hung-over) to extremely high stakes lies (e.g. being charged with espionage, murder, or rape) (Phillips, Meek, & Vendemia, 2011). For this study, benign lies are not the focus for two reasons; first they are extremely difficult to detect and more importantly, the return on investment outside of academic psychosocial and theoretical context, is nearly nothing. Exploitive lies that are also high-stakes on the other hand hold the most potential for detection and the benefit is far more valuable to society (Utz, 2005). One aspect of current deception detection research is the inevitable focus on identifying and measuring behavioral cues to deception which will be explored next.

Cues to Deception

Deception in face-to-face conversations has been studied for many years from the lens of many different fields (DePaulo, & Kashy, 1998; deTurck, & Miller, 1990; Ekman, 1992; Jensen, Burgoon, & Nunamaker, 2010; Littlepage, & Pineault, 1979; Mann, Vrij,

& Bull, 2004). Several meta-analyses exist that attempt to summarize the large body of studies in deception and deception detection. A meta-analysis is a summary of several studies and attempts to resolve inconsistencies in research. Three previous meta-analyses (i.e., DePaulo, Stone, & Lassiter, 1985; Kraut, 1980; and Zuckerman & Driver, 1987) examined cues associated with deception in a large number of studies. For example, Zuckerman and Driver (1987) found in their meta-analysis that negative statements, verbal immediacy and discrepancies in the narrative were the most powerful indicators of deception. DePaulo et al.'s meta-analysis (2003) of 158 cues to deception revealed several correlates of deception including the number of details in the participant's message and how uncertain the participant seems. Based on these reviews, the following cues in Table 1 were found to be associated with deception:

Table 1, Cues Associated with Deception

Cue	Description (Liars…)	Reference
Blinks	blinked more often	DePaulo et al., 1985; Kraut, 1980; Zuckerman & Driver, 1987
Adaptors	moved their hands more when giving responses	DePaulo et al., 1985; Kraut, 1980; Zuckerman & Driver, 1987
Speech Errors	made more errors when speaking	DePaulo et al., 1985; Kraut, 1980; Zuckerman & Driver, 1987
Duration	were more brief	DePaulo et al., 1985; Kraut, 1980; Zuckerman & Driver, 1987
Pupil Dilation	pupils are more dilated	DePaulo et al., 1985; Zuckerman & Driver, 1987
Irrelevant Information	include less relevant information	DePaulo et al., 1985; Zuckerman & Driver, 1987
Negative Statements	responses contain more negative expressions	DePaulo et al., 1985; Zuckerman & Driver, 1987
Shrugs	shrug more	DePaulo et al., 1985
Immediacy	exhibit less involvement	DePaulo et al., 1985; Zuckerman & Driver, 1987

Cue	Description (Liars…)	Reference
Pitch	vocal pitch is more anxious	DePaulo et al., 1985; Zuckerman & Driver, 1987
Hesitations	hesitate more	DePaulo et al., 1985; Zuckerman & Driver, 1987
Leveling	use more leveling terms (e.g., overgeneralized statements)	DePaulo et al., 1985; Zuckerman & Driver, 1987
Message Discrepancy	messages contain more discrepancies	Zuckerman & Driver, 1987

The behaviors listed in Table 1 contain both verbal and nonverbal cues. Because audio recording are a strong source of deceptive cues (DePaulo et al, 2003), they are capable of capturing many dimensions of verbal and nonverbal behavior simultaneously, and hold the potential for automation, the following section will discuss speech cues to deception more closely.

Speech Cues

Deception researchers have long been interested in speech cues (Anolli & Ciceri, 1997; Fay & Middleton, 1941; Howard, 2011; Podlesny & Raskin, 1977; Zuckerman, DePaulo et al., 1981). The relationship between verbal characteristics and nonverbal characteristics is an intricate one. These two communication modes are interdependent with temporal referencing and the interaction of personal and situational objectives (Bruneau, 1980).

Linguistic-Based Deception Cues

Linguistic-based and paralinguistic-based channels are promising sources of deceptive cues in interview style communication (DePaulo, et al., 2003). Zuckerman et al. (1981) found that transcripts of verbal content were higher than any other single channel at detecting deception. They also found that tone of voice was among the better channels for detecting deception, above that of facial expressions. Given that speakers generally control what they say there must be leakage in the audio channel comprised of how they speak and other not so obvious aspects of what they are saying (e.g. word complexity patterns).

Examining how someone speaks leads to paralinguistic information which includes vocal characteristics of time, intensity, frequency, and fluency as well as nonverbal or linguistic based characteristics including for example syntax choice and unintentional word choices. Other aspects of what is spoken and their potential to carry deceptive cues have been studied by many researchers (Burgoon & Qin, 2006; Moffitt & Burns, 2009; Sporer, 1997; Vrij, 2008; Zhou, Burgoon, Twitchell, Tiantian, & Nunamaker 2004; Zhou, Twitchell, Tiantian, Burgoon, & Nunamaker, 2003). Zhou, Burgoon and their colleagues found eight categories of linguistic cues that were above chance at detecting deception. One study looking at these linguistic categories in a RWHS setting was performed by Fuller (2008). Fuller, Biros, and Wilson (2009) revised these categories using data from a RWHS environment, namely *Quantity*, *Specificity*, *Uncertainty*, *Clarity*, *Immediacy*, *Affect*, and *Cognitive Processing*.

Fuller et al.'s (2009) study looked at 370 written suspect statements given during law enforcement interviews following RWHS criminal cases. The seven constructs and the number of measures for each examined by Fuller are listed in Table 2.

Table 2, Deception Detection Lexical Constructs

Construct	# of Measures	Construct	# of Measures
Quantity	3	Immediacy	3
Specificity	5	Affect	6
Uncertainty	5	Cognitive Processing	3
Clarity	5	**Total:**	**30**

Two factors already discussed but which were not possible to examine in the Fuller data set are (1) the impact of the dyadic interaction and (2) the behavioral changes over time. These constructs and their measures are discussed in greater detail in Chapter 3.

Paralinguistic-Based Deception Cues

According to Dictionary.com (2011), paralinguistics is "the study of vocal features that accompany speech and contribute to communication but are not generally considered to be part of the language system, as vocal quality, loudness, and tempo". In other words, if linguistics is the *what* in human speech, paralinguistics is the *how* they speak. Four paralinguistic domains frequently examined when analyzing speech for deception are: time, intensity, frequency and fluency. What follows is a summary of research in each domain and their findings.

Time

Speech cues related to time are duration and the speed of verbalization. Frequent measurements of time when considering deceptive speech are: total length of the interaction, response length, response latency, tempo of speech, and the rate of change of speech.

Researchers found that the length of interaction generally decreases with deception as subjects attempt to distance themselves from the lie (Feeley & deTurck, 1998; Rockwell, Buller, & Burgoon, 1997). Researchers reported mixed results in regards to response length. DePaulo et al. (2003) found that the total response length to interviewer questions was not significantly different in deceptive versus truthful exchanges while Anolli and Ciceri (1997) reported longer more expressive responses. Response length seems rather content specific (e.g. yes/no questions vs. "describe in your own words" questions). DePaulo et al, (2003) also reported that deceivers take up a smaller proportion of talk time than truth tellers.

Another measure of time is response latency or the amount of time between the end of a question and the beginning of the response. Though some reports contend that response latency is not significantly correlated with deception (Feeley & deTurck, 1998), DePaulo et al., (2003) reported an increase in latency when deceivers did not have time to prepare their deception. The final measure of time frequently examined by deception detection researchers is tempo or the rate an individual speaks as well as the change in their rate. The meta-analysis by DePaulo et al. (2003) reported somewhat mixed results;

the rate of speaking when a transgression was committed was significantly correlated with deception (see p. 101) but the rate of change was not.

Intensity

Cues dealing with the loudness or amplitude of speech are grouped under *Intensity*. In general, researchers have reported mix results regarding intensity. Mehrabian (1971) reported a decrease in amplitude while Buller and Aune (1987) reported an increase in amplitude of deceiver's speech. DePaulo et al.'s meta-analysis suggested that it was not significant (DePaulo, et al., 2003). However, variation in amplitude has been reported to increase during deception (Rockwell, Buller et al., 1997).

Frequency

Of all speech cues, perhaps frequency and pitch have been the most studied as possible indicators of deception. Researchers have found a significant increase in frequency during deception (DePaulo, et al., 2003; Ekman, Friesen et al., 1976). As a point of clarification, researchers often use the terms *pitch* and *frequency* interchangeably, this is not completely appropriate. *Frequency* describes a physical phenomenon while *pitch* describes a perceptual phenomenon. An increase in frequency variation has also been found to be correlated with deception (Vrij, 1995).

Fluency

Four measures of speech fluency in deception detection research are filled pauses, non-ah disturbances, silent pauses, and interruptions; all of which show mixed results with a slight decrease in fluency by researchers (DePaulo, et al., 2003). Non-ah speech

disturbances include speech disturbances other than “um”, “er”, and “ah”. Some researchers found that non-ah speech disturbances decrease during deception (Bond, Kahler, & Paolicelli, 1985) but the meta-analysis by DePaulo et al. (2003) did not find a significant relationship with deception. The filled pauses include the total amount of “um”s, “er”s, and “ah”s in a subject’s speech while filled pause length refers to the length of each such filled pause. Feeley & deTurck (1998) report an increase in filled pauses while Bond, et al. (1985) report no correlation with deceptive speech. Some researchers have noted that deceivers commit more speech errors and hesitations (Feeley & deTurck, 1998), while others have found no relationship between deception and filled pauses (Sporer & Schwandt, 2006). The total number of silent pauses and their duration have also had mixed results (DePaulo, et al., 2003; deTurck & Miller, 1985). Finally, interruptions, the overlaps of speech between a subject and an interviewer have also shown mixed results. Some researchers report a decrease in interruptions during deception (Buller, Comstock et al., 1989) while others have not found a significant correlation with deception (DePaulo, et al., 2003).

In general, current literature on paralinguistic-based cues to deception report mixed results. This may be as a result of low RWHS data sets examined. The five paralinguistic constructs and the number of measures for each are in Table 3.

Table 3, Deception Detection Paralinguistic Constructs

Construct	# of Measures	Construct	# of Measures
Fluency	6	Duration	4
Tempo	2	Intensity	2
Frequency	3	**Total:**	**17**

Discovery of cues to deception is an initial step, what follows are several methods developed to identify and measure deception itself.

Existing Deception Detection Methods

Many methods have been developed to detect deception with varying levels of success. This review discusses the advantages and disadvantages of several of the most well-known methods. Several of these methods are based on theory while others were developed by trial and error. A discussion of leading theories of deception follows the discussion on existing methods.

Polygraph

The most well-known method of deception detection is the polygraph, a device that takes various cardiac, skin conductivity, and respiratory measures to detect deception. It is based on the idea that these physiological measures are directly linked to the conditions that are brought on by deception attempts (Vrij, 2000). The two main questioning techniques that are used during a polygraph are the Comparative Question Test (CQT) and the Guilty Knowledge Test (GKT). The CQT establishes a baseline behavioral score based on a series of irrelevant control questions then compares the baseline scores to scores on crime-specific questions. However, it has often been criticized as subjective, non-scientific, and unreliable (Ben-Shakhar & Elaad, 2003). CQT is reported to be accurate on deceivers from 83% to 89%, but only 53% to 78% for truth tellers (Bull, Baron, Gudjonsson, Hampson, Rippon, Vrij, 2004). Regarding CQT

Iacono (2001) concluded that, "Although the CQT may be useful as an investigative aid and tool to induce confessions, it does not pass muster as a scientifically credible test."

The GKT compares the physiological responses to multiple choice questions about a crime, where one of the choices contains details about the crime only the criminal would know. For example, "what type of weapon killed Mr. Green: (a) gun, (b) knife, (c) rope, or a (d) lead pipe?" The GKT enjoys a more objective approach than CQT (Ben-Shakhar & Elaad, 2003) however it has major problems. One problem with the GKT is that there may be many reasons for a subject to choose the "guilty" choice and with say four choices, 25% of the innocent people may get false positive scores. Accuracy rates for the GKT range from 42% to 76% for liars and between 94% and 98% for truth-tellers (Bull et al., 2004). The general conclusion from the scientific community is that the polygraph is pseudoscience (Iacono 2001). Observation of behavioral cues is another method used as a means of deception detection.

Behavioral Analysis Interview

Behavioral Analysis Interview (BAI) is an interview technique developed by Inbau, Reid, Buckley, and Jayne (2001). BAI is a non-accusatory structured questioning technique designed to provoke verbal and nonverbal behavior from subjects which indicate deception. Numerous studies have shown that deceivers act differently than truth-tellers (DePaulo et al., 2003; Ekman & Friesen, 1969; Zuckerman, DePaulo, Rosenthal, & Leonard, 1981). However, most people are mistaken in their beliefs about which behaviors are associated with deception (Vrij, 2000). As an example of this

behavior, Inbau et al. (2001) expect liars to be more nervous and less helpful. However, this is the exact opposite of what other deception literature predicts (Vrij, Mann, and Fisher, 2006). Vrij et al. (2006) conducted the first empirical test of the BAI technique where grounded truth was established. Their results did agree with Inbau et al. (2001) that the technique lead to differences between liars and truth tellers, however, their results were consistent with existing deception literature and in the exact opposite direction of what Inbau et al. (2001) predicted. For example Inbau et al. (2001) expected liars to be more nervous and cross their legs more often, but Vrij et al. (2006) found that liars sat more still than truth tellers, which is what other deception research has stated (DePaulo et al., 2003). This technique shows promise but more empirical testing needs to be accomplished.

Voice Stress Analysis

The voice stress analysis (VSA) was introduced in the 1970's as a possible replacement of the polygraph (Rice, 1978). The VSA measures psychophysiological responses of the subject's voice. VSA procedures attempt to rely upon microtremors in the voice as indicators of deception. These machines are fundamentally designed to detect stress, not lies. The accuracy of voice stress analyzers is reported to range from chance to about equal to that of the polygraph (Gamer, Rill, Vossel, & Godert, 2006; Vrij, 2005). However, other studies of VSA have yielded accuracy rates not significantly different than chance (Cestaro, 1996; Janniro & Cestaro, 1996). There are commercial VSA systems available that claim to distinguish truth from lie but independent reports

fail to confirm these claims (Haddad & Ratley, 2002; Hollien, 2006). Despite its initial promise, VSA has failed to gain acceptance in the scientific community (Ford, 2006; Hollien & Harnsberger, 2006; Hopkins, Benincasa, Ratley, & Grieco, 2005). If that were not enough, a Department of Justice funded research showed "little validity" in the technique as well (Damphousse, Pointon, Upchurch, & Moore, 2007).

One small advantage VSA has over many other deception detection methods is its adaptability to be automated. However, VSA is highly dependent on the quality of the audio, so in an uncontrolled environment like a busy police station or an airport terminal it may be of little value.

Though the current study examines voice characteristics, it is not VSA. VSA only looks at frequencies; the current study examines 41 dimensions of human speech and how it is produced (e.g. duration, intensity, pitch, temporal, linguistics, etc.).

fMRI

A newer method of deception detection that has shown great promise is the analysis of brain activity and blood flow utilizing a functional magnetic resonance imaging (fMRI) machine.

Recently fMRI reached deception detection accuracy of 100% when subjects do not employ countermeasures (Ganis, Rosenfeld, Meixner, Kievit, & Schendan, 2011). fMRI measures the hemodynamic response, or changes in blood flows, that are related to brain activity. This neuroimaging technique can display what is called blood-oxygen-

level dependence (BOLD) (Ogawa, Lee, Nayak, & Glynn, 1990). When brain cells are active they consume more energy which results in localized increases in blood flow to that area and subsequent concentrations of deoxyhemoglobin. It is the deoxyhemoglobin that is the contrast in the MRI. It has been suggested that BOLD fMRI may be sensitive to differences between truth and deception as seen in the activation of different regions of the brain (Langleben, 2008). Researchers have noticed differences between the brain activity of truth-tellers and deceivers (Ganis, Kosslyn, Stose, Thompson, & Yurgelun-Todd, 2003; Johnson, Barnhardt, & Zhu, 2004). Additionally, researchers using fMRI have localized the regions of the brain that were most informative in terms of deception detection, namely the right prefrontal regions and the bilateral posterior cortex (Davatzikos, Ruparel, Fan, Shen, Acharyya, Loughead, Gur, & Langleben, 2005). Most physiological methods of deception detection require the use of invasive sensors either attached to the body or requiring immobilization. A subject must be willing and cooperative or the devices used to measure physiological reactions cannot accurately function. For example, fMRI seems to work around the need for a willing subject by measuring blood flow in the brain, something people presumably cannot control. However, the fMRI machine requires the subject to be completely still; even talking during a scan can distort the images (Ogawa et al., 1990). Researchers have adapted their methodology to utilize stimuli that do not require the subject to move their head (e.g. audio instructions, images, handheld input controllers).

One criticism of fMRI use in deception detection is the temporal differences between electrical activity and blood flow in the brain; brain electrical activity is on the order of milliseconds, while changes in blood flow are on the order of seconds. Because

there is no normative BOLD fMRI baseline, studies must rely on the difference in brain response between a target and control stimuli (pictures of a crime scene vs. random unrelated pictures). This is referred to as "cognitive subtraction", a method of isolation of cognitive processes by comparing BOLD responses to different stimuli (Aguirre, 2010). The use of cognitive subtraction makes the selection of appropriate target and stimuli critical (Aguirre & D'Esposito, 1999).

fMRI is drawing a lot of attention not only for its reported accuracy but because it is based on several valid physiological and methodological assumptions; for example the fact that cerebral blood flow and neuronal activation are coupled (Huettel, Song, & McCarthy, 2004) and Ogawa's BOLD discovery (Ogawa et.al., 1990). However, it is these same apparent strengths which are drawing the most criticism namely methodological design flaws and the physiological differences between individuals (Langleben, 2008). Other researchers have pointed to problems with replication, large individual brain differences and unclear brain regions associated with truth-telling (Spence, 2008).

Statement Validity Analysis and Criteria-Based Content Analysis

Statement Validity Analysis (SVA) is a technique for analyzing deception which focuses on verbal content. SVA was originally developed for determining the veracity of the testimony of children in sexual abuse cases, but has since been adapted for and applied to other types of cases including those with adult subjects. SVA involves an examiner searching for linguistic cues and gaps in a subject's statements and using

follow-up questions to uncover discrepancies. Proponents of SVA say it has proven highly effective as a police interrogation technique; however critics argue that it encourages a lie bias; assuming a subject is deceptive then trying to affirm guilt before questioning even begins (Richard, 2008).

One common method of SVA is Criterion-Based Content Analysis (CBCA). CBCA is based on the Undeutsch-Hypothesis which states that "A statement derived from a memory of an actual experience differs in content and quality from a statement based on invention or fantasy" (Vrij, Roberts, & Bull, 2000). CBCA takes place during a structured interview where an interviewer trained in CBCA scores the subject's responses according to predefined criteria such as logical structure. The results of past studies report a wide range of accuracies from 55% to 90% (Vrij, Mann, & Fisher, 2006).

The SVA technique has been criticized for its lack of theoretical foundation (Sporer, 1997; Vrij, 2000). For example, CBCA has no formal rules for determining whether a statement is truthful or deceptive (Vrij, 2000). CBCA has also been shown to have a truth bias; it works better at detecting truths than lies (Vrij, 2000). Perhaps because of the complex training required, CBCA has been shown to have problems with inter-rater reliability (Godert, 2005).

SVA and CBCA are not suitable candidates for automation because they require trained interviewers and coders to conduct and score the interviews. Despite the adaptability of this technique, the subjective nature of CBCA, the amount of training required, and the delayed results limits its potential for automation (Kulaney, 1982). To

help guide development and understanding of further deception detection methods research should be grounded in theory.

Theories of Deception

There are many psycho-social and communication theories which delve into deception and credibility assessment as well as interdisciplinary theories that have been applied to the subject. In order to cover the leading theories a rough taxonomy is attempted by first separating pure deception detection theories and then covering related theories and perspectives. Pure deception detection theories are further dichotomized into strategic and non-strategic theories.

The majority of deception theories focus on non-strategic cues that are triggered by non-rational, uncontrollable behavior, and in which the sender is typically not aware they are occurring (Buller & Burgoon, 1994). The principle of non-strategic cue development is straight forward; deceptive actions can have an emotional or cognitive impact on an individual which may create outwardly detectable behavior not consciously under the sender's control. However, deception is also thought to be governed by strategic actions. A long stream of research suggests that individuals voluntarily adapt behavioral patterns in response to a receiver's perceived acceptance of previous messages (Buller & Burgoon, 1996; Buller, Strzyzewski & Comstock, 1991; Burgoon, Buller, Dillman, & Walther, 1995; Burgoon & Qin, 2006; Clark & Clark 1977). Senders not only plan and calculate their deceptive behavior in the initial interaction but continuously adapting their behavior during the dyadic interaction.

Ekman's Cues to Deceit

The current thinking about deception detection has largely evolved from Ekman and Friesen's (1969) ideas of deceptive cues. Paul Ekman describes two kinds of cues to deceit: leakage and deception cues (Ekman, 1985; Ekman & Friesen, 1969). Leakage describes the mistakes deceivers make when attempting deceit. Deceivers try to control behavioral displays so as not to give themselves away, but cues associated with deception leak out anyway, often through communication channels that are more difficult to control. Deception cues are other behavioral changes that reveal that deception is taking place such as fear, guilt, duping delight, and other performance related changes. According to this view, there are emotional correlates of deception and emotional expression is not entirely under conscious control. These emotions may become observable and measurable behavior. The practice of deceiving can cause a variety of emotions such as guilt or fear (DePaulo et al., 2003; 1992; Vrij, 2000). For example if someone feels ashamed of their actions they may lie to avoid guilt, or if the stakes are high enough they may fear getting caught and paying the price. Ekman (1985) also argued that liars may also experience what he calls "duping delight." Duping delight refers to those who experience excitement associated with the challenge of getting away with a lie.

These emotions might manifest in measurable physiological conditions such as increased pulse, breathing rate, voice pitch, etc. Ekman calls the deceptive cues from these emotions "feeling cues" (Ekman, 1992). The flip side of "feeling cues", Ekman calls "thinking cues", those measurable behavioral cues that result from cognitive aspects of deception. Similar to feeling cues, thinking cues may manifest in slower speech when

a lie is complex or long and may appear rehearsed when deceivers over prepare their story (DePaulo et al., 2003). Table 4 summarizes Ekman's thinking and feeling cues associated with deception.

Table 4, Ekman's Thinking & Feeling Cues

Category / Component	Predicted Cognitions / Behaviors / Cues
Thinking cues	Those who over prepare seem rehearsed Those who think carefully may speak more slowly
Feeling Cues:	
Fear	Higher pitch Faster and louder speech More/longer pauses Speech errors Indirect speech
Guilt (sadness)	Lower pitch Softer and slower pitch Downward gazing
Duping Delight	Higher pitch Faster and louder speech Use of more illustrators
Performance related thoughts and feelings	Thoughts about whether or not the performance is succeeding Feelings about the performance Feelings about any fabricated or discreditable portions of the performance

(DePaulo, et al., 2003; Ekman, 1992, 1985)

If deceivers have control over the various communication channels and can manage the messages they send as well as the reactions of the receivers then it may be possible for them to suppress their unintended behavioral expressions (Buller & Burgoon, 1994). Other researchers argue that when this complex process is attempted some cues may still be leaked (DePaulo, et al., 2003; Ekman, 1992, 1985). The Ekman and Friesen (1969) perspective was expanded on by Zuckerman, DePaulo, Rosenthal, and Leonard, (1981) into what they call the Four Factor Theory (also referred to as the Four Factor Model).

Four Factor Theory

Zuckerman and Driver's (1987) Four Factor Theory (FFT) is another theory with potential application in RWHS environments. The Four Factor Theory was initially proposed by Zuckerman and his colleagues in a series of papers (Zuckerman et al., 1981; Zuckerman & Driver, 1985, 1987) in an attempt to explain the underlying processes governing deceptive behavior. The model tries to tell why people behave differently when lying than when telling the truth. FFT assumes that people are more aroused or anxious when telling lies than when telling the truth. The authors defined four factors involved in deception that can influence behavior: arousal, emotion, attempted control, and cognitive load. According to this theory, deceivers will try to control their behavior to prevent disclosure of deception which ironically reveals cues to deception such as behavior that appears planned, rehearsed or lacking in spontaneity.

FFT, like Ekman's feeling cues to deceit, suggests that deceit may cause physiological arousal. As an example of the arousal factor, an individual might be physiologically aroused due to intense questioning about a particular topic and as a result, voice pitch is elevated, speech errors increase and response times are altered away from their norm. Several other nonverbal behaviors such as pupil dilation and eye blinks are also suggested to change with deception but are not a focus of this study.

This theory is also similar to Ekman's thinking cues to deception (Ekman, 1985). The FFT suggests a cognitive component to deception as well. It is believed that deception is more difficult than telling the truth because our brains are taxed more when fabricating and maintaining cohesive facts and timelines then when telling the truth (Vrij,

Fisher, Mann, & Leal, 2006). Other researchers have considered that the amount and type of thinking may be impacted during deceptive interactions. Creating a credible lie often demands that multiple layers to a story be woven together in a logical manner (Zuckerman et al., 1981). These deceptive tapestries can be difficult to convincingly express while juggling fact and fiction without leaking deceptive cues (Anolli & Ciceri, 1997). Research by Schacter and Buckner (1998) as well as more recent fMRI studies discussed above suggest that the mental processes for retrieving an experienced memory are different than fabricating a false one.

Finally, FFT asserts that lying requires people to think a lot harder than telling the truth does. Researchers have hypothesized that liars, compared to people telling the truth, would take a longer time to respond, pause more when speaking, and deliver messages with few details (Vrij, Edward, Roberts, & Bull, 2000; Zuckerman & Driver, 1985). In other words, a deceiver's linguistic pattern should be simpler than a truth teller's. Some research has been skeptical about certain assumptions contained in FFT (Feeley & Young, 1998). For example, McCornack (1997) argued that in some cases telling lies may be less cognitively difficult or less arousing than telling the truth. For example, a wife asks her husband, "Do you like my new hairdo?" and he hates it but instead of taxing his mind and behavior to come up with a nice way to tell the truth he simply tell her he loves it.

It is common practice in law enforcement interviews to "dig deeper" during questioning which would require a more elaborate deception as questioning goes on. This complexity increases cognitive load which may lead to identifiable changes in the behavior of the subject such as more frequent hesitations, and a decrease in frequency of

illustrators. Critics of FFT often note that some of the behaviors associated with cognitive complexity may also be related to arousal and that it may not be possible to isolate exact causal antecedents (Vrij, Fisher, Mann, & Leal, 2006). Critics of FFT also suggest that cognitive complexity may be low when deception is built from partial truths or past experiences (Vrij, et al., 2006). It follows that when periods of high-stake stress are longer, as in the context of law enforcement interviews, cognitive load increases and may impact deceptive cues (Vrij et al. 2006).

The four factors of the FFT model, with the exception of attempted control are primarily nonstrategic in nature. That is, an individual displays these behavioral cues independent of any long-term plan of action to achieve a particular goal. The attempted control of behavior is strategic in intent, but the indicators associated with this factor are inadvertently displayed.

Table 5 summarizes the FFT model and provides predicted behaviors or cues for each factor with speech sourced cues in bold.

Table 5, Four Factor Theory Summary

Factor	Predicted Behavior / Cues
Arousal	Greater pupil dilation Increased blinking **More frequent speech disturbances** **Higher Pitch**
Feelings while lying	**Fear or guilt** Fidget more **Sound more unpleasant** **Distance themselves – more evasive**

Factor	Predicted Behavior / Cues
	maintain less eye contact
Greater cognitive effort	**Longer response latencies** **More speech hesitations** Greater pupil dilation **Fewer illustrators**
Attempted Control	**Less spontaneous behavior** Verbal and nonverbal discrepancies

(DePaulo et al., 2003; Zuckerman et al., 1981)

The claim made by FFT that deceivers will try to control their behavior to prevent disclosure of deception is also supported by previous research conducted by Ekman and Friesen (1974, 1969). According to their sending capacity hypothesis, when people tell lies they try to control their behaviors but tend to pay more attention to some communication channels than others. However, most people tend to focus on the behaviors they believe communicate the most information, like the facial movements but tend to forget about those behaviors that communicate little information, such as our legs and feet (Ekman & Friesen, 1974, 1969). This is the basis for Ekman's leakage cues mentioned above. For example, one study found that people who watched liars' heads and faces were less accurate at detecting deception than people who watched liars' bodies (Ekman & Friesen, 1974). In a summary of more than 30 studies in which judges tried to detect others' deception from either single channels (i.e., only the face, body, tone of voice, or words) or from particular channel combinations, DePaulo, Stone, and Lassiter (1982) found that in all conditions where judges relied on facial cues, detection accuracy was lower. This study also concluded that when judges paid attention to what liars were saying, they were more accurate at deception detection than when verbal channels were

unavailable. This last point is why this study is focusing on the verbal channel to examine its potential in applied situations.

Critics point out that pathological liars and sociopaths will not have the same behavioral responses to lying that the general population does. This raised a question that is outside the scope of this study but interesting none the less; are those who enjoy duping delight sociopaths? However, these disorders are atypical and can be diagnosed by other means and more importantly, are not the population FFT attempts to describe.

Information Manipulation Theory

Another way of approaching deception was proposed by McCornack (1992) in his Information Manipulation Theory (IMT). IMT proposes that deceptive messages function because they covertly violate the principles governing conversational exchanges namely the maxims of quality, quantity, relation and manner. These conversational maxims were originally proposed by Grice (1989) as guidelines for effective and efficient use of language. In order to deceive, a sender can manipulate any combination of the four maxims. IMT is also based on the proposition that it is assumed others are truthful, informative, relevant, and clear. Deception occurs when speakers alter the amount of information that should be provided (quantity), the veracity of the information presented (quality), the relevance of information provided, or the clarity of information provided.

Quality manipulations would be represented by deliberate distortions or fabrication of information. The quantity of the information may be manipulated simply by altering the amount of information that is presented (i.e. holding back the full truth;

half-truths). The relation maxim would be violated when liars manipulate the relevance of information such as answering a direct question with an indirect, off topic answer. Finally, the manner maxim of IMT can be manipulated by being ambiguous or not clear such as when using words with multiple meanings. For example, in the sentence, "I cannot tell you how much I miss my mother-in-law's visits", it is not clear whether the subject enjoys the visits so much they cannot express their feelings or that they hate the visits so much they do not want to say anything. Here, information is conveyed in an ambiguous fashion or with a lack of clarity. Deception is made more complex and difficult to detect when people alter several or all of the maxims at the same time, though this becomes increasingly more difficult to perform. IMT therefore suggests that deception can take on a very large number of forms (McCornack, Levine, Solowczuk, Torres, & Campbell, 1992). However, if it were possible, identifying a particular person's pattern of manipulations of the four maxims could improve detection accuracy. All four maximums are present in speech with only the relation maximum being beyond the scope of this study. Based on IMT an examination of quality, quantity, and clarity characteristics of speech should be examined for deceptive cue and patterns.

Self-Presentational Perspective

DePaulo (1992) described the Self-Presentational Perspective (SPP) of nonverbal communication and attempted to bridge the conceptual void between spoken and non-spoken communication research. It was not until 2003 that DePaulo et al. (2003) applied the SPP to the study of deception.

DePaulo's SPP advocates that liars and truth tellers both need to appear honest but that only truth tellers have grounds for their claim, while a liar's claims of honesty are false. The deceiver therefore must attempt to regulate their nonverbal behavior in order to appear honest. SPP suggests that people are often not very successful in their attempts and that their behavior can give away their deception (DePaulo, 1992).

SPP suggests that liars are less forthcoming, less compelling, less positive/pleasant, more tense, and include less unusual content than truth tellers. However, DePaulo et al.'s SPP is largely based on the pretext that most lies that are told are *everyday lies* (Sporer & Schwandt, 2006). Despite any overlap in the predictions of the self-presentational perspective with prior theories of deception, SPP may be most applicable when examining white lies and therefore is of less interest in RWHS deceptive environments.

Media Richness Theory

Another promising theory coming from the MIS field that lends strong support for the current research is Media Richness Theory (MRT) by Daft and Lengel (1986). MRT defines media richness as the ability of information communicated on the medium to reduce equivocality. It tries to explain that richer, more personal means of communication are generally more effective for communication of equivocal issues than leaner, less rich media. It suggests that the richer a medium the more information that is communicated. It assumes that resolving ambiguity and reducing uncertainty are the main goals of communication. In deceptive communication this assumption is reversed

for the deceiver. The goal of the sender is to appear as if this assumption is valid while the reverse assumption is their true objective. MRT states that media richness is a characteristic of a communications channel that affects the ability of that channel to support messages with varying levels of cogitative and affective content (Carlson 1995;; Daft and Lengel 1986; Zigurs and Buckland 1998). The broad levels of media channels in order of richness are face to face, video, audio, and then text messages. More recent research has failed to support the full breadth of the original MRT (Dennis & Kinney 1998; Dennis & Valacich,. 1999). However, MRT does suggest that the selection of speech as a rich medium to explore for deceptive cues is a sound one because speech carries multiple channels of communication simultaneously (i.e. it is a very rich medium which can be used to reduce equivocality or manipulated to possibly increase deception).

Signal Detection Theory

Though Signal Detection Theory (SDT) is not a deception detection or credibility assessment theory per se, it is a supporting theory to this research because the voice can be treated as a signal for data analyze. SDT developed by Green and Swets (1966), defines two sets of probabilities in a test, in which two possible state types must be discriminated. In the context of deception detection, the two possible stimuli types are deceptive and truthful intent. If the actual intent is deception and the output judgment is suspicion, the trial is a "hit". If the actual intent is truthful and the output is judged suspicion, it is a "false alarm". If the actual intent is deceptive but the judgment is one of

trust, it is a "miss". Finally, if the actual intent is trustworthy and the judgment is one of trust, it is a “correct decision” as shown in Table 6.

Table 6, Possible Judgments from SDT

		Judgment	
		Suspicion	**Trust**
Actual Intent	**Deception**	**Hit**	**Miss**
	Truthful	**False Alarm**	**Correct Decision**

According to SDT, the output of such a binary test is based on the value of a decision variable, which in the context of deception detection is the trust/suspicion judgment level. The threshold value of the decision variable is called the criterion. For humans, the selection of a criterion is not only related to the value of actual stimuli but also related to their psychological characteristics. In other words, the criterion is a function of perceived stimuli which in the context of deception detection are the behavioral deviations. The SDT calculation methods described in Stanislaw & Todorov (1999) can be used to study the distribution of the values of the suspicion level variable across the behavioral profile deviations to determine the appropriate criterion for the final decision making.

Interpersonal Deception Theory

One of the more promising theoretical foundations to examine deception detection during interview-style communication is Interpersonal Deception Theory (IDT) (Buller &

Burgoon, 1996, 1994; Burgoon & Buller, 1994; White & Burgoon, 2001). IDT is by far the most encompassing theory on deception research with 18 propositions and 42 hypotheses. IDT models deception as an interactive dyadic communication between sender and receiver in a back and forth nature where each is simultaneously encoding and decoding messages over time then adapting their behavior to meet their goal, deception (sender) or detection (receiver).

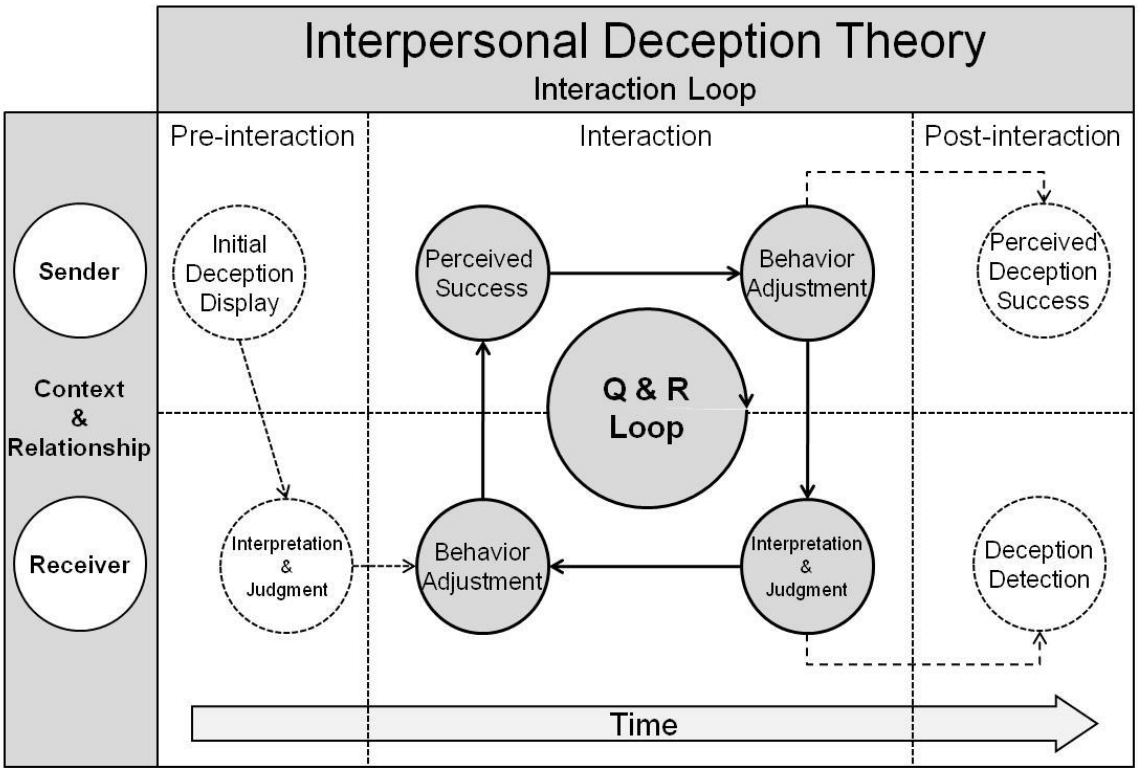


Figure 1, Interactive Deception Model (Adapted)

Figure 1 illustrates the model proposed by IDT. According to this model, both the sender and receiver of deception bring to an interaction their expectations, goals, familiarity, etc. During the interaction the sender will begin his or her deceit with certain strategies but will modify those strategies throughout the interaction based on perceived deception success.

The receiver, on the other hand, begins with some level of suspicion, which changes throughout the interaction based on credibility judgments they make on the senders behavior. Although it is suggested that suspicion plays a role in both senders' and receivers' behavior (Buller & Burgoon, 1996), another issue is the concern with whether suspicion affects cue intensity. Specifically, when people are more suspicious are the cues of deception more or less pronounced in the sender? Some scholars have found that suspicious subjects are no more accurate at detecting deceptive cues than naive subjects (Mattson, 1994). Bond and Fahey (1987) argued and Hubbell, Mitchell and Gee (2001) found that this could be because suspicious people are *lie bias*; more likely to interpret ambiguous information as lies rather than truths. This is particularly problematic for those pre-disposed to lie bias (e.g. law enforcement). However, McCornack and Levine (1990) argued that accuracy may depend on the level of suspicion. Specifically, they found that moderate levels of suspicion led to greater accuracy when judging deception.

In either case both parties will likely inadvertently reveal linguistic and paralinguistic cues of their psychological state. In the end, both sender and receiver are able to evaluate their success at deceiving and detecting deceit, respectively.

IDT is also the leading theory of deception that views deception as strategic communication. According to the authors of IDT, communication includes both strategic and nonstrategic behaviors. Strategic behavior refers to large-scale plans and intentions, not necessarily to specific routines or tactics. Non-strategic behaviors reflect unintentional, unconscious behavior also labeled leakage (Ekman & Friesen, 1969). During the interaction, both parties may alter their strategies as they perceive the

effectiveness of their discourse. IDT proposes that unintentional behavioral leakage will occur from both sender and receiver during the interaction which may reveal their deception or suspicion respectively. Specifically, IDT argues that deceivers' strategic attempts to appear honest are often poorly executed and nonstrategic behaviors may be beyond the liar's control.

One extrapolation from IDT that can be made is that IDT supports the development of unbiased tools for deception detection (Burgoon, Nunamaker, George, & Biros, 2007). For example, compare human familiarity and computer information systems on bias. The theory argues that familiarity may lead to both truth bias and lie bias. It may seem counterintuitive but the better a subject is known, the less effective one is at detecting their lies (McCornack & Parks, 1986). Specifically, familiarity increases confidence which leads to truth bias (others are perceived as honest). Stiff, Kim, and Ramesh (1992) found that truth bias was positively associated with familiarity and negatively associated with deception detection accuracy. In other words, in relationships based on trust, a truth bias is likely. McCornack and Levine (1990) also found the negative of this to be true; in adversarial relationships (e.g. law enforcement & suspects), a lie bias (believing others are dishonest) becomes more likely. In either case familiarity may lead to biases which make one less accurate at detecting deception. However, information systems tools have no familiarity and can be built to avoid biases typical of humans.

IDT & Law Enforcement Interview Process

IDT is also a good theoretical match to the law enforcement interview process where interviews are typically longer and more numerous than lab studies will allow. These longer interactions may intensify the impact of repeated questioning, making it harder to lie while giving law enforcement personnel more opportunity to identify the deception. It is also possible with longer interactions that patterns of behavior will appear which can be examined for changes that correlate with deception.

This interview process is depicted in Figure 2 below. Law enforcement interviews are typically recorded with audio or video equipment for post-interview analog analysis and possible legal evidence; not for real-time analysis. Any behavior changes must be captured by the investigator during questioning in order to have a direct impact on the outcome of that interview session. Any behavior changes caught after the fact during analog analysis may be of much less value because the sender is given time for post-interview processing. Real-time feedback would give the interviewer a serious advantage over the sender because they could dynamically change their questioning technique, giving the sender little time to adapt.

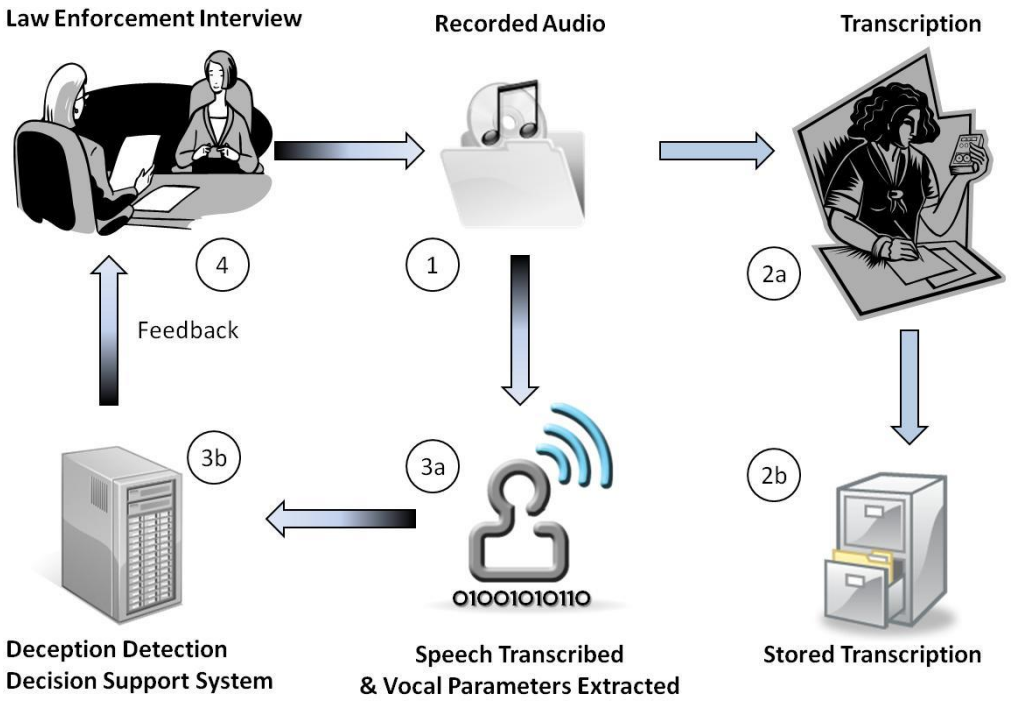


Figure 2, Deception Detection Feedback Loop

In reference to Figure 2 above, (1) during the law enforcement interview audio recordings are made. (2a) These recordings are later transcribed by a third party or by law enforcement personnel themselves for (2b) further case analysis or evidence in a court of law. (3a) If significant cues to deception can be extracted from speech (3b) and processed real-time by decision support systems, the potential exists to provide an almost immediate feedback loop to the law enforcement investigation team. (4) This close to real-time feedback could significantly assist law enforcement questioning by identifying possible deception attempts or reactions to key questions.

IDT & Theory Validation

Theories on deception need validation in RWHS settings. This study focused on IDT, one of the most referenced theories on deception detection and attempted to validate it in a real-world high-stakes case study (validation of other deception detection theories is left for future research).

To address the major gaps in current deception detection research and in an attempt to validate portions of IDT within these gaps, the following modifications are made to the previous IDT model (see Figure 3).

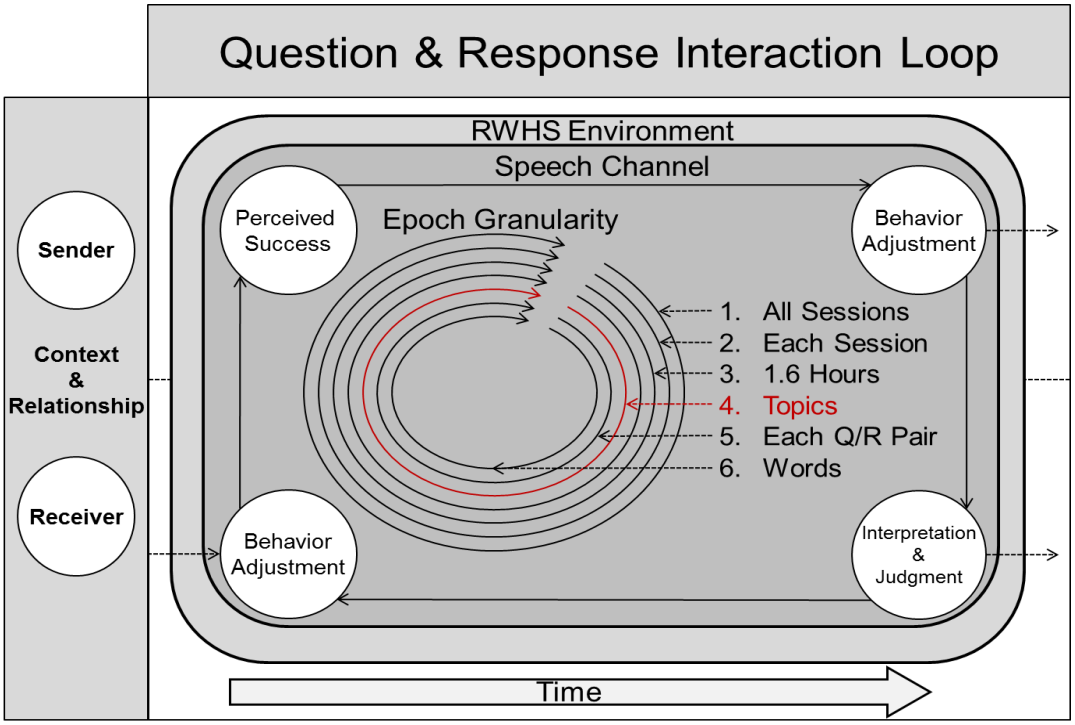


Figure 3, Overlay of Law Enforcement Interview Process on IDT

IDT is centered on the interactive process so the current study focused on the dyadic interaction; leaving the pre- and post-interaction phases to later research.

The interactive portion of the model is encapsulated in a RWHS environment. Within this environment the communication channel of focus is speech, a highly interactive and multi-dimensional channel that lends itself to automation of cue processing. To address the aspect of change over time suggested by IDT the question/response loop has been broken down into epochs of decreasing duration. With the exception of the third level these follow an implicit hierarchy of six different sizes from largest to smallest, specifically: 1.) all sessions between sender and receiver, 2.) individual sessions, 3.) 1.6 hour sections, 4.) question/response pairs by topic, 5.) individual question/response pair, and 6.) individual words. The third level was chosen as a point of reference base on work by Kassin, Leo, Meissner, Richman, Colwell, Leach, and Fon, (2007) in which 631 police investigators were questioned on the average duration of questioning sessions. The average duration of a RWHS questioning session was 1.6 hours. This provides the opportunity to make comparisons within and between epochs. Table 7, Epoch Descriptions, describes each epoch level and the general levels of evaluation which can be performed.

Table 7, Epoch Descriptions

Epoch Length	Description	Evaluation Point
All Sessions	All interactions between receiver and sender regardless of breaks	E_1; {Truthful vs. Deception}
Each Session	Each continuous interaction separated by breaks without communication	Same as above plus patterns: Between: $E_1 \rightarrow E_2 \rightarrow E_3 \ldots$ Within: ΔE_1 ; ΔE_2 ; $\Delta E_3 \ldots$
1.6 Hours	~ 96 minutes	Same as above up to n: $E_1 \rightarrow E_2 \rightarrow E_3 \rightarrow E_4 \ldots E_n$

Epoch Length	Description	Evaluation Point
Topics	Groups of question/response pairs separated by topic	Same as above.
Each Q/R Pair	Each question/response pair	Same as above.
Words	Individual words	Same as above however several linguistic cues will be null.

"All Sessions" is defined as all recorded interactions between receiver and sender regardless of time or frequency of session interruptions. Because this level encompasses a single mean for each indicator across the entire data set, comparisons and tests for patterns is very limited even though this is the level at which many deception detection studies base their comparisons. Because of these limitation this level of granularity was left for further study when other RWHS data sets can be gathered. "Each Session" was defined as a continuous interaction separated by breaks without communication. These breaks could be days or minutes in duration (e.g. a break for lunch, restroom, or sleep). Due to the very large data set this study focused on the sub-epochs in a single session leaving multiple session analysis for future study. As discussed "1.6 Hours" level was chosen as a reference point matching RWHS durations typical of law enforcement interviews. This level may prove to be less valuable in the current case study but will make comparisons to other studies and data sets more compatible. "Topic" level is defined as groups of question/response pairs separated by changes in content topic. During dyadic interactions the questions/response can go both ways (i.e. sender asks for clarification; receiver asks follow up questions). This level of epoch granularity may seem to go against potential automation. Identification and separation of content is very difficult even with advanced computing and artificial intelligence (Pudota, N., Dattolo,

A., Baruzzo, A., Ferrara, F., & Tasso, C., 2010). However, implementation would greatly simplify topic separation; the receiver or an accomplice simply indicates a change in topic electronically during interviews. "Each Q/R Pair" level is defined as each question/response pair regardless of topic. This level is important to examine because it is common for deceivers to mix deception and truth to appear more believable. "Word" level is simply defined as each individual word. The lowest level possible from a linguistic point of view would be the combination of sounds used to make words, however their examination is beyond the scope of this study. It is worth noting that many studies on audio compare levels of granularity down to 1/30 second. This is done because the duration of time matches frame rates in video, making comparisons between audio and video simpler. This level of granularity is outside the scope of this study because it does not consider video; however this could prove to be productive in a future study.

In addition to the modifications proposed to the IDT model, the context and relationship within which the interactions take place are restricted to a RWHS context where the relationship between sender and receiver is one in which the receiver is authorized and directs questions toward the sender; typical of a law enforcement interview.

Research Question & Propositions

Based on the above literature review and to a greater extent the theoretical basis of IDT, the primary research question ("*Are speech cues to deceptive behavior moderated*

over time by receiver suspicion during dyadic interactions in a real-world high-stakes setting?") is expanded to include the following seven propositions:

P1: *The difference in speech cues to deceptive behavior will be moderated over time by the level of suspicion.*

P1A: *The magnitude of speech cues to deceptive behavior will be positively related to an increase in level of suspicion.*

P1B: *The level of moderation by suspicion will be strongest when epochs are at the topic level.*

P2: *Speech cues during deceptive behavior will form identifiable patterns that differ from speech cue patterns during truthful behavior.*

P2A: *Patterns of speech cues during deceptive behavior will be most detectable when epochs are at the topic level.*

P3: *As the granularity of epochs decreases (epochs gets smaller), the number of speech cues to deceptive behavior will increase to an apex then decrease.*

P3A: *Speech cues will have the highest correlation with deceptive behavior when epochs are at the topic level.*

Based on this literature review we know that the majority of studies on deception detection, and by extension the theories used to explain deceptive behavior, took place in controlled environments with structured questions. Furthermore, the population studied

consisted mostly of college age students from universities who participated either for minor monetary incentive or as a requirement in a course. These studies were also designed around specific mock scenarios with low-stakes for the deceiver if they were "caught".

What the current study is going to illustrate is whether the measurements and constructs common in many deception detection research (i.e. linguistic and paralinguistic) can hold up under a real-world, high-stakes environment. Linguistic and paralinguistic measurements and the leading deception detection theory, IDT were used to help focus the study onto a manageable subset of deception research.

This chapter described existing deception detection methods, leading deception theories, and the measurements that are common to those theories. In addition, the need for more RWHS deception research was summarized.

The following chapter will detail the methodology used to answer the propositions above based on the literature review. Chapter 3 will also explain the development of a set of constructs for studying deception detection in a RWHS setting given the new model of IDT proposed.

CHAPTER 3

METHODOLOGY & ANALTICAL APPROACH

This chapter presents the case study data, how it was gathered and prepared for analysis, what measures were examined, and the analytical approach taken in order to examine the main research question and the propositions.

Core to this study is the exploratory nature of real-world cases and how IDT explains the relationships between examined behaviors and deceptive states. The characteristics of IDT and the real-world law enforcement interviews discussed above that are examined in this study are: dyadic communication, high-stakes, duration, and multi-channel speech communication. It is believed that cues extracted from audio recordings of these interviews would lead to identifiable behavioral patterns over time. Next is a description of the case study methodology. Initially, the use of a single case study methodology is clarified in this context followed by a description of the case itself. Next, is a description of the behavioral measures taken from the speech channel and a detailed explanation of the data processing model used to transform them into an analyzable format. Finally, a description of the analysis methods used is given, which are both statistical and exploratory in nature.

Case Study Methodology

When relevant behaviors cannot be manipulated or to do so would be unethical (e.g. felony, fraud, murder, assault) a case study is the ideal methodology (Yin, 2009). According to Yin (2009) a case study is, "an empirical inquiry that investigates a contemporary phenomenon in depth and within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident". The case study is an exploratory method that allows the retention of the holistic and meaningful characteristics of real-life events during the exploration of the complex and often ill-defined topic of deception. Exploratory studies attempt to answer how and why questions with the hope of better understanding a problem or situation that has not been clearly defined. In contrast, explanatory studies are more suitable for causal questioning. They build on exploratory studies and attempt to explain why a well-defined event occurs. In an exploratory, real-world case study the questions deal with operational associations that need to be examined over time, rather than statistical frequencies alone. A major strength of case studies is their ability to trace changes over time while not being limited to cross-sectional or static assessments (Yin 2009). Because of the wicked problem of establishing ground truth in RWHS deception and the unethical feasibility of laboratory experiments, case studies based on field data seem to be the experimental design with the greatest chance to further the understanding of deception detection.

In regards to experimental design, case studies are commonly regarded as quasi-experimental where behaviors cannot be manipulated but the logic of experimental design can still be applied (Cook & Campbell, 1979).

Many examples of successful case studies can be found in top pier journals across every discipline. From the field of MIS, Ravishankar, Pan, and Leidner, (2011) published a case study in *Information Systems Research* (*ISR*) on the implementation of a knowledge management system (KMS). This example points outs the importance of using a case study to fill research gaps. Ko and Dennis (2011) also performed a KMS case study, this time within a single pharmaceutical company exploring the impact of time and experience. The Ko and Dennis (2011) case study is a good example of a study where N=1 (i.e. the pharmaceutical company) but the focus of the study and the unit of analysis was a subunit within N. A final example of a top tier journal using case studies is by Malhotra, Majchrzak, Carman, and Lott, (2001) who published in *MIS Quarterly*. Malhotra et al.'s, (2001) case study involved the behavior of team members utilizing virtual computer-mediated collaboration technology to develop a radically new rocket engine. This is a good example of using a case study when conditions are unique and difficult to replicate. Similar examples can be found in any discipline, the point being that case studies, if implemented properly, can produce solid contributions to the body of knowledge in a field of study.

There are some prejudices against the case study method according to Yin (2009); the greatest is a concern for a lack of rigor. Too often case studies are attacked for being sloppy and not systematic in their procedures (Yin, 2009). To address this concern the utmost care must be given to meticulous documentation of processes and methods to a level equivalent or beyond those of experimental design. Another prejudice critical of case studies is an apprehension that they provide little basis for scientific generalization. Case studies, like experiments, are generalizable to theoretical propositions and not to

populations (Thomas, 2010). The goal of doing a case study and an experiment are the same, to expand and generalize theories (analytic generalization) and not to enumerate frequencies (statistical generalization). A third concern is that case studies take too long, result in massive, unreadable documents. This does not have to be the case and stems from confusion with data collection methods like ethnography and participant-observation. Ethnography is a study of cultures and requires vast periods of time in the field (Fetterman, 1989). Participant-observation research requires intimate familiarity with the group being studied (Jorgensen, 1989). Case studies are a form of inquiry and do not rely on these methods alone (Yin, 2009).

Table 8, Complementarily of Case Studies and Statistical Methods

	Case Studies	Statistical Methods
Strengths	Depth	Breadth
	High conceptual validity	Understanding how widespread a phenomenon is across a population
	Understanding of context and process	Measures of correlation for populations of cases
	Understanding of what causes a phenomenon, linking causes and outcomes	Establishment of probabilistic levels of confidence
	Fostering new hypotheses and new research questions	
Weaknesses	Selection bias may overstate or understate relationships	Conceptual stretching, by grouping together dissimilar cases to get larger samples
	Weak understanding of occurrence in population of phenomena under study	Weak understanding of context, process, and causal mechanisms
	Statistical significance often unknown or unclear	Correlation does not imply causation
		Weak mechanisms for fostering new hypotheses

Case studies and statistical methods can complement each other. Table 8 lists some strengths and weaknesses of both. One final justification for using case studies is that they can offer important evidence to complement experiments; an adjunct to experiments rather than an alternative to them. This is the approach taken for this study. Next, is a description of the case itself including steps taken to collect data.

Case Description - The James Perry Case

Please note: this case has been adjudicated and all identifiable information is publically available upon proper request. In Nov 2004, James Perry was sentenced in federal court in Madison, Wisconsin to 470 years in prison for creating child pornography, rape, sexual exploitation of children, child sexual assault and kidnapping, a crime spree that spanned over a five year period. It is the longest sentence for sex crimes in Wisconsin history and there is no option for parole. What follows is a detailed background of the subject and crimes committed, followed by a description of the data collection process.

Background

In 2004 James Perry committed his final assault which led to his capture. Perry, a husband and father of two young girls, entered a Madison, Wisconsin hotel with the intent of committing a sexual assault. This incident was only one of two times Perry was ever caught on film despite targeting very public locations. It was a key piece linking him to a long series of rapes and assaults over five years and four states. Upon entering

the hotel Perry proceeded to a back elevator where he saw a 13 year old girl entering an elevator alone. He assaulted the young girl at gun point when she tried to exit the elevator. The girl resisted, at one point saying, "Jesus, Jesus I love you." to which Perry told her to shut up; Perry later said that this statement impacted him and made him feel guilty for his actions. Perry exited the building through a back stairwell door with the girl where she saw friends unpacking their car down at the end of the lot. She screamed, broke free of his hold, and ran; Perry ran in the other direction and escaped. At the same time the FBI was investigating a child pornography ring of which Perry was involved. Their investigation led them to a cabin in upstate New York where they arrested Thomas Redeker. A video of Redeker and Perry engaged in sexual activity with young children was found. Redeker turned in Perry and gave the FBI his contact information. Within hours 17 FBI agents and local police surrounded Perry's house.

Only a few days after the assault and attempted abduction of the 13 year old girl the FBI arrested Perry for his involvement in the internet child pornography ring. When they searched Perry's house they found videotapes and 121 CD-ROMs containing hours of child pornography and dozens of pictures almost all of pre-teen girls. Perry stalked these children in malls, county fairs, public pools, parking lots, and toy aisles in department stores. Included among the evidence were videos of Perry sexually assaulting two 8-year-old girls.

Det. Maureen Wall of the Madison PD became aware that the serial rapist she had been hunting was in FBI custody. The FBI was not aware of the rape and assault charges at that time. Perry did not want to talk to Madison PD because he was trying to negotiate with the FBI on a proffer plea deal to try to get only six years in prison for manufacturing

child pornography. A proffer agreement is a written agreement between the prosecutor and an individual allowing the individual to provide information about a crime or possible crime to the government, with the assurance that his words will not be used against him in the government's case in the event of a subsequent trial. Det. Wall informed the FBI about the plethora of crimes he committed as "The Mall Rapist", a name the news stations called him at the time; all plea bargaining on federal charges stopped so Det. Wall could conduct the interview. The Madison PD had a list of 45 victims but believed there were hundreds more.

Meanwhile the media began running stories portraying Perry as a monster. He later told Det. Wall those stories made him really mad and that he did not want his mother, wife, and daughters to believe that he was a cold, brutal, monster who beat and raped women and children. Det. Wall said in her statement that Perry wanted the chance to craft the story in his own words, from his own perspective, and give admissions to only a handful of crimes that did not appear to be violent – even though he actually did do those things.

Only after the interview and when Perry became aware of all the evidence against him a plea agreement was made to stop adding on charges (over and above the 125 he was now being charged with). Again, after the interview Perry wanted to expedite his sentencing and avoid a trial and a media circus that would paint him as a monster. There were no other considerations in the plea agreement. The time of initial FBI arrest and sentencing took place within a one year period.

Data Collection

Law enforcement videotaped three consecutive days of interviews totaling 14 hours and 27 minutes of video. Interviews were conducted by the same lead detective and her partner in the same room and under the same conditions with Mr. Perry and his attorney. Interaction was primarily between lead Det. Wall and Mr. Perry, only minor contributions (less than five minutes total) were made by the second detective and Mr. Perry's attorney; their voices were removed before analysis. A 200 page law enforcement transcript was generated by the lead detective immediately after the interviews. The law enforcement transcript contains all questions asked and the responses, often in quotations with additional pertinent notes by Det. Wall. Both the videotaped interviews and law enforcement transcripts were used in federal court.

Ground Truth

Ground truth was established by the lead detective based on credible evidence admissible in a federal court. The lead detective identified four types of statements: confirmed lies, suspected lies with evidence, suspected lies without evidence, and the truth. Known lies were those statements proven to be false by evidence admissible in court. When the sender made these statements law enforcement personnel knew for a fact he was lying. Suspected lies with evidence were those statements law enforcement personnel had disputing evidence on, however for various reasons that evidence was not or could not be admitted into federal court. Suspected lies without evidence were those statements law enforcement personnel believed, in their expert opinion, to be false but for

which they had little or no evidence. The final type of statements are truthful, were the law enforcement personnel knew were the truth or had no reason to believe they were false.

Prior to receiving any data the Oklahoma State University Institutional Review Board (IRB) was contacted and all IRB procedures were completed. Following IRB approval contact was made with the Madison, Wisconsin Chief of Police and the lead detective in this case. A written request for the data stating the purpose of the request and detailing procedures to keep the data secure was sent. Following approval by the Chief of Police the raw data was copied to CD and mailed along with the written law enforcement transcript coded with the lead detective's level of suspicion. The following section describes the dependent variables and how they were captured. Next is a description of the constructs and their measurements followed by a description of the independent variable and moderator variable, Suspicion.

Measurements

The model consists of 41 total measures across the 12 deception detection constructs. The linguistic-based cue constructs are: *Quantity*, *Specificity*, *Uncertainty*, *Clarity*, *Immediacy*, *Affect*, and *Cognitive Processing*. Paralinguistic-based cue constructs are: *Time*, *Intensity*, *Frequency*, *Fluency*, and *Duration*.

Linguistic-Based Constructs Development

The initial model contained linguistic-based constructs from Fuller, Biros, and Wilson (2009). Fuller et al.'s constructs were chosen because they generated almost 74% accuracy in deception detection, the data was RWHS field data taken in law enforcement environments with solid ground truth validation, and the units of measure were written statements. This matches the current data set with the exceptions that it is a transcript of a law enforcement interview and the unit of measure varies from topic level to question/response pairs; both of which are a focus of the study. In addition to the seven linguistic constructs by Fuller, listed above, an eighth construct of Severity was also considered by them to be important. However it is not a part of the current study because its measure would be constant across the current data set. The current data comes from a serial rapist, the punishment for which was life in prison. The lead detective in this case would assign the maximum severity score of five on the one to five scale used by Fuller. The Fuller constructs along with their measurements are described in Table 9, Linguistic-Based Constructs and Their Measurements.

Table 9, Linguistic-Based Constructs and Their Measurements

Construct	Construct Measurement	Brief Description
Quantity	# of Words, Verbs, & Sentences	Length of message
Specificity	Sensory ratio, Spatial ratio, Temporal ratio, Content Word Diversity, Bilogarithmic Type-Token-Ratio	Amount and type of details in the message
Uncertainty	Certainty Terms, Tentative Terms, Modal Verbs, Passive	Relevance, directness, and

Construct	Construct Measurement	Brief Description
	Voice, Generalizing Terms	certainty of message
Clarity	Redundancy, Sentence Length, Complexity Ratio, Average Word Length, Causation Terms.	Message clarity and comprehensibility
Immediacy	1st person pronouns, 2nd person pronouns, 3rd person pronouns	Attempts to disassociate oneself from the events described
Affect	Activation, Imagery, Pleasantness*	Emotions present in the message
Cognitive Processing	Exclusive Verbs, Motion Words, Cognitive Processing Terms.	Increased or decreased cognitive processing and cognitive information present in the message related to veracity

* Note, Fuller et al. (2009) used positive and negative measures for each *Affect* measure, this study combines the positive and negative into a single bi-polar measure for ease of processing.

Paralinguistic-Based Constructs Development

The initial model is based on vocal constructs examined by Meservy (2007). These constructs and their measures were selected for this study because they represent a thorough coverage of the audio channel and tools exist to measure each consisting of: *Fluency*, *Duration*, *Tempo*, *Intensity*, *Frequency*, and *Voice Quality* (Anolli & Ciceri, 1997; DePaulo et al. 2003; Rockwell, Buller et al. 1996; Scherer, Feldstein et al. 1985; Zuckerman et al., 1981). However, because the construct Voice Quality contains cues that are difficult to measure objectively without the aid of a human evaluation this construct was removed; a focus of this study is on identifying behavioral cues that can be objectively measured and potentially automated. The Meservy (2007) constructs are

described along with their measurements in Table 10, Vocal-Based Constructs and Their Measurements.

Table 10, Vocal-Based Constructs and Their Measurements

Construct	Construct Measurement	Brief Description
Fluency	1. Non-ah disturbances 2. Speech errors 3. *Interruptions** 4. Silent pauses 5. Filled pauses	1. Speech disturbances other than "um", "er", "ah", and other such words 2. General speech errors 3. *Overlaps of subject and interviewer that results in a change of turns* 4. Various pauses in conversation
Duration	1. Length of interaction 2. Response length 3. talking time 4. *Response latency**	1. Total time of dyadic interaction 2. Length of sender's response 3. Proportion of total time sender talks 4. *time between end of question and beginning of senders response*
Tempo	1. Rate of speaking 2. Rate change	1. Average number of words per minute 2. Rate of speaking in the epoch minus the average rate of speaking for all responses
Intensity	1. Amplitude 2. Amplitude variety	1. loudness of senders voice 2. variation of loudness of a sender's voice
Frequency	1. Pitch 2. Pitch change 3. Pitch variety	1. The average fundamental frequency of sender's voice 2. variation of pitch of a sender's voice 3. Frequency of changes of pitch of a sender's voice

*note, the measures Interruptions and Response Latency are not considered due to the difficulty in automating these measures.

Moderator

Proposition 14 from IDT (Buller & Burgoon, 1996, p. 231) states, "Suspicion (perceived or actual) increases senders' (a) strategic and (b) nonstrategic behavior". They stated mixed supporting evidence to which the present study provides an opportunity to test this proposition. Suspicion level is the independent variable (IV) in this study but it is also a moderator of the senders' behavior. The IV is ordinal and the dependent variables (DV) vary in type including categorical, continuous, and ratios. In general, the moderator effect is measured by correlating the independent variable (IV) with the dependent variable (DV) (See Table 9 & Table 10) for each state of moderator and then testing the difference. The dyadic circular communication pattern depicted in the IDT model (Figure 1 & Figure 3) is very similar to moderator model seen in Figure 4 below in that the communication is dyadic and circular, providing senders the opportunity to change their behavior based on perceived or actual suspicion from the receiver.

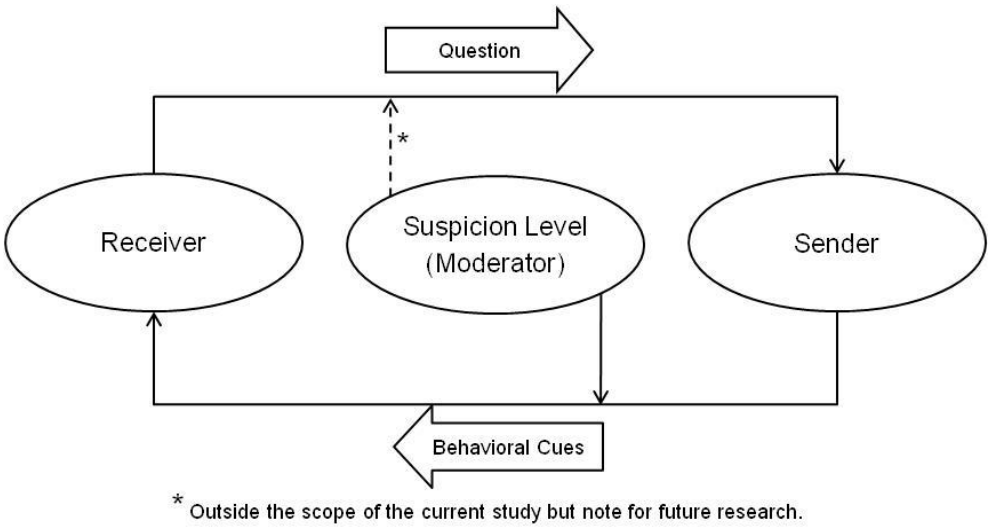


Figure 4, Moderator Model

The research question (*Are speech cues to deceptive behavior moderated over time by receiver suspicion during dyadic interactions in a real-world high-stakes setting?)* required an examination of 41 behavioral cues to deception within four levels of suspicion across three levels of granularity of epochs. So the moderator model was run 12 different times with repeated measures of the 41 cues taken over time. At each run in an attempt to identify which combination of suspicion and granularity best addresses the research question. The way to measure the moderator effect is to correlate IV with the DV separately for each level of suspicion and examine the impact on behavioral cues as well as compare between suspicion level impacts (Baron & Kenny 1986). The following sections will describe the data preparation as well as a description of the analysis methods utilized.

Data Preparation

The data preparation process followed the steps shown in Figure 5. First, the raw video stored on DVD was processed with Adobe Soundbooth to isolate the audio from the video portion; there was no loss of audio data during this step. The digital audio files were then passed through DC Live Forensic 7.5 to improve audibility in preparation for segmentation. Global filters were applied to remove audio signals outside the abilities of humans to hear as well as make. It should be noted that any filters or transformations to improve audibility were applied universally. It should also be noted that all recording took place in the same room with the same recording device and same environmental settings. Once global filters removed noise and audibility quality was improved, audio was segmented into question/response pairs.

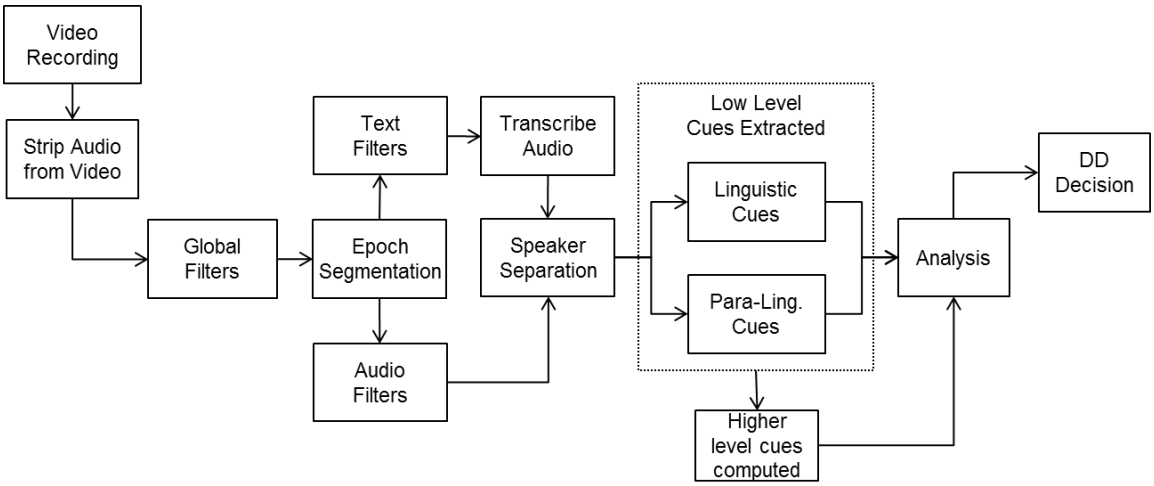


Figure 5, Data Processing Model

The audio data was then duplicated for split processing for the two categories of cues, linguistic and paralinguistic. In preparing the audio for transcription any audio or

acoustic filter can be applied that improves transcription accuracy (i.e. pitch, tone, cadence, etc. have no impact on linguistic cues). Identification of which filters to run was done by hand. However, unknown, multi-source, noisy speaker separation and transcription research is advancing and many commercial products are available that can perform a known, single-source, quiet transcription such as Dragon Naturally Speak $_{tm}$.

The goal of processing the data for paralinguistic cue measurement is the removal of noise without removing, degrading, or changing the speech signal. There are several techniques for removing and improving clarity of audio however, some can be very aggressive and rely on human physical and cognitive audio processing characteristics to "trick" the listener into hearing clearer voices (Campbell, 2008; Roweis, 2000; Schimmel, Atlas, & Nie, 2007). This study took a conservative approach to audio filter selection to retain as much of the voice signal as possible.

Transcription Automation Difficulties

The audio data segments then underwent transcription using Docsoft: AV. Unfortunately this was very unsuccessful. Docsoft: AV uses at its' core a Dragon Naturally Speaking voice recognition engine combined with specialized hardware for speech capture and transcription (e.g. news and TV broadcasts use similar systems for the closed captioning function). However Dragon's speech recognition, like almost all commercially available transcription systems, is not designed to transcribe multiple voices, it is speaker dependent (Transcribing-interview, n.d.). Unfortunately, interview style audio almost always contains multiple speakers (e.g. police interviews) and presents

several challenges to transcribers (MacLean, Meyer, & Estable, 2004). The most common solution and the one used in almost every courtroom in the world is called "voice writing" (Voice Writer, n.d.). During voice writing the transcriber repeats the words of the subject verbatim, typically into a stenomask to block their own voice from being heard and to block outside sounds from being recorded.

Because the audio for this research comes from a real-world law enforcement interview containing multiple speakers on a single channel microphone, automating transcriptions is nearly impossible with today's technology. Though several interview processes were considered that would reduce or remove these barriers (i.e. multiple microphones, instruct interviewers to not overlap sender speech), they were left for future research when new data can be gathered. For the present study, voice writing was performed to capture the linguistic cues. The transcript was then manually verified for 100% accuracy.

Linguistic cues were measured from the transcript using Structured Programming for Linguistic Cue Extraction (SPLICE) and Linguistic Inquiry and Word Count (LIWC) software. Waikato Environment for Knowledge Analysis (WEKA) (Witten & Frank, 2000) is used for classification based on the initial text processing steps. This transcript was compared to the law enforcement transcript were ground truth and deceptive statements were coded into the full transcript.

Following processing, linguistic and paralinguistic measures were recombined into one data set while maintaining chronological order. The following section discussed several analysis methods that were employed.

Analysis Methods

To better appreciate the difficulty of pursuing multi-method research one must realize that over 40 unique approaches to qualitative methods have been identified (Tesch, 1990) including ethnography, grounded theory, phenomenology, case study, narrative research, and historical research to name a few. The outcome of qualitative studies is what one learns or comes to understand about the phenomenon, organized by category or theme (Herndon & Kreps, 2001).

The purpose of exploratory research is to clarify the research questions that guide the entire research project. This suggests that it precedes some larger more formal (i.e. quantitative) research project or stream (Merriam, 1998). John W. Tukey (1980) strongly supported exploratory research stating, "Finding the question is often more important than finding the answer". Tukey developed one tool called Exploratory Data Analysis (EDA) for just that purpose.

In EDA, like all exploratory tools, the emphasis is on insight and flexibility, in contrast to hypothesizing a specific function, estimating factors, and testing for model adequacy (Tukey, 1977). Exploratory data analysis is itself an iterative process. Its first use is to examine the raw data. This may identify additional aspects of the data which in turn encourage further investigation (Kundzewicz & Robson, 2000). EDA is also a very visually-based analysis technique (Jewitt & Leeuwen, 2010) hence the use of graphical analysis.

From the above, an attempt to clarify the main research question was made by exploring narrower sub-questions and reporting insights obtained in a detailed descriptive exploratory case study of RWHS deception. In exploring these sub-questions a mixed method approach was utilized in order to maximize the understanding of the phenomenon in a very specific environment (Axinn & Pearce, 2006). In the true nature of exploratory research the below method of evaluation (MOE) and measure of performance (MOP) were a starting point (Creswell, & Clark, 2007). Given the limitations of a single case study where the unit of analysis is a sub-unit and with the assumption that although the unit of analysis comes from the same source, there is some element of independence and empirical comparisons are made were possible (Yin, 2009). The goal of the outcome is to offer propositions to stimulate future research about RWHS deception (Saunders, Lewis, & Thornhill, 2007).

RQ: *Are speech cues to deceptive behavior moderated over time by receiver suspicion during dyadic interactions in a real-world high-stakes setting?*

What follows are the specific questions, their MOE, and MOP. MOEs are the suggested tools or techniques implemented to investigate the question (Table 11). Because this study is exploratory the MOEs may change in response to findings. MOPs are qualitative or quantitative measures of system capabilities or characteristics.

Table 11, Prospective Analysis Methods

PROPSITION	*P1: The difference in speech cues to deceptive behavior will be moderated **over time** by the level of suspicion.*
MOE	(1) Trend lines with nested base line comparisons (Monmonier, 1990; Shumway & Stoffer, 2011); (2) Change-Point Analysis (Yamanishi & Takeuchi, 2002).
MOP	(1) Compare slopes of regression lines of different suspicion levels.

	This is a mixed method approach utilizing the linear regression test method in an exploratory manner (i.e. without specifying a specific desired significance level). (2) Change-point analysis is a powerful tool for determining whether a change has taken place. It is capable of detecting subtle changes missed by other methods (Taylor, 2000a & 2000b). http://www.variation.com/cpa/tech/changepoint.html.
PROPSITION	*P1A: The **magnitude** of speech cues to deceptive behavior will be positively related to an increase in level of suspicion.*
MOE	(1) Cumulative stacked line charts Ward & Guo, (2011) ; (2) Change-Point Analysis (Yamanishi & Takeuchi, 2002); (3) 3-dimentional nonlinear smoothers (Conradie, deWet, & Jankowitz, 2009; Díaza, Domínguezb, Cuadradoa, & Fuertesb, (2008); (4) Heat maps (Wilknson & Friendly, 2009).
MOP	(1) Cumulative stacked line charts show the relationship of the parts to the whole over time. Visual analysis will be performed to look for changes in cue magnitude; (2) see MOP P1; (3) Inspection of surface area of 3D map with Z-score on vertical axis, suspicion on Y axis, and epoch on X axis (4) visual inspection of color change.
PROPSITION	*P1B: The **level of moderation** by suspicion will be strongest when epochs are at the topic level.*
MOE	(1) Boxplots (Seo & Shneiderman, 2005), display differences between populations without making any assumptions of the underlying statistical distribution (i.e. non-parametric). (2) Tukey's test.
MOP	Visual comparisons across minimum, lower quartile, median, upper quartile, and maximum; (2) $P < 0.10$
PROPSITION	*P2: Speech cues during deceptive behavior will form **identifiable patterns** that differ from speech cue patterns during truthful behavior.*
MOE	(1) Scatter Plot Matrices, disaggregated by suspicion (Zhang, 2008); (2) data mining, cluster analysis (3) repeated measures ANOVA.
MOP	(1) Visual inspection of construct x epoch x z-score; (2) Compare squared Euclidean distance of centroids; (3) t-test.
PROPSITION	*P2A: Patterns of speech cues during deceptive behavior will be most detectable when epochs are at the **topic level.***
MOE	Data mining, cluster analysis (Comas, Turmo & Surdeanu, 2008).
MOP	Compare squared Euclidean distance of centroids.
PROPSITION	*P3: As the granularity of epochs decreases (epochs gets smaller), **the number of speech cues** to deceptive behavior will increase to an apex then decrease.*
MOE	(1) Trend lines with nested base line comparisons (Shumway & Stoffer, 2011; Monmonier, 1990).
MOP	(1) Compare slopes of regression lines of different granularity levels. This is a mixed method approach utilizing the linear regression test method in an exploratory manner (i.e. without specifying a specific desired significance level).

PROPSITION	*P3A: Speech cues will have the **highest correlation** with deceptive behavior when epochs are at the topic level.*
MOE	Linear regression comparisons Jensen, Lowry, & Jenkins, (2011).
MOP	Pearson product-moment correlation coefficient > .5; R-squared.

Several MOEs are listed above including descriptive statistics, ANOVA, principal components analysis, and several different graphical methods. What follows is a brief description of each.

Descriptive Statistics & ANOVA

Initially the data was examined with descriptive statistics to gain an understanding of the range and spread of the data. According to Trochim, (2000),

> "Descriptive statistics are used to describe the basic features of the data in a study. They provide simple summaries about the sample and the measures. Together with simple graphics analysis, they form the basis of virtually every quantitative analysis of data."

Descriptive statistics provide a valuable means to make initial comparisons across the units of analysis. Almost all published research uses descriptive statistics to some degree, deception detection research is no different. One recent article on deception detection which references many of the same theories discussed in this study comes from the Journal of Forensic & Investigative Accounting (Cefaratti & Barkhi, 2013). In which the authors use descriptive statistics to describe their sample population and even to lend support for their hypotheses. There are examples of descriptive statistical method used in

almost all deception detection research. The initial analysis method in the current study was basic descriptive statistics in the form of box plots and a mean comparison chart. A box plot is a way of graphically depicting numerical data via a five number summary: minimum observation, lower quartile, median, upper quartile, and the largest observation. Box plots make no statistical distribution assumptions but are good for displaying differences between populations. Box plots can indicate the spread and skewness of data as well as identifying possible outliers. In deception detection research box plots have been used to show, very quickly, the lack of significant differences in treatments (Duran, Dale, Kello, Street, & Richardson, 2013).

In addition to the box plots for each measure comparing truthful and deceptive samples, a table will present the means for each measure and construct (refer to example Table 12). This study makes extensive use of mean tables and frequency distributions, again common in deception detection research (Olson, 2013). The use of ANOVA and the various forms of regression are almost expected, even in exploratory research that their use should not need justification (Dunbar, Jensen, Bessarabova, Burgoon, Bernard, Harrison, & Eckstein, 2012; Fuller, Biros, & Delen, 2011).

With 41 raw measures across hundreds of question/response pairs, the data set contains over 31,000 raw data points. When multiple epochs, 12 constructs and data normalization are considered the data set contains well over 100,000 data points.

Table 12, Cue & Conststruct Means Table Format

Cues	Mean (Raw)	Mean (Raw w/ Direction)			
		Truthful	↕	Deceitful	↕
# of Words			↓		↑
…	…	…		…	

…

Constructs	Mean (Z-Score)			
	Truthful	↕	Deceitful	↕
Quantity		↓		↑
…	…		…	

Because the DV's in this study vary in type and scale, a z-score was computed in order to make unitless comparisons. A standard score henceforth referred to as z-scores was computed for each cue at each epoch then averaged for each construct at each epoch with the following function: $z = \frac{x-\mu}{\sigma}$, where x is a raw score, μ is the mean, and σ is the standard deviation. The z-score is dimensionless with a mean of zero and a standard deviation of one (Marx & Larsen, 2006).

Principal Component Analysis (PCA)

One goal of this study is to confirm the constructs used by previous researchers when measuring deceptive behavior. To ensure that the constructs and cues within them are measuring what it is believed they should, principle components analysis (PCA) was conducted. PCA has been used in deception detection research on low-quality data (Raiman, Hung, & Englebienne, 2011); the current study's data came from a real-world law enforcement interview video tape, not the best quality. Another benefit of PCA is a reduction in dimensionality or factor reduction. Often one finds that large variances

associated with the first $k < m$ principal components, and then a precipitous drop-off (Shlens, 2005). It can be concluded that the most interesting dynamics occur only in the first k dimensions. Therefore the first step of PCA is to identify the components with the highest variance utilizing eigenvalues as a cutoff guide.

The number of components selected was based on their eigenvalue or the percent variation explained by the corresponding principal component. A minimum of 80% of the variance explained was selected as a cut point. Because these components are perpendicular to each other they are uncorrelated.

Graphical Analysis

To help understand and visualize patterns over time a graphical analysis was performed. Implementing data visualization as a means to analyze data is as much an art as it is a science (Fayyad, 2002; McBurney & White, 2009). Several methods were explored and those most informative were reported in this study. For example: bar charts to compare across suspicion levels, line charts or bump charts to visualize trends over time, scatter plots to look for clusters and outliers, and moving ranges to find patterns.

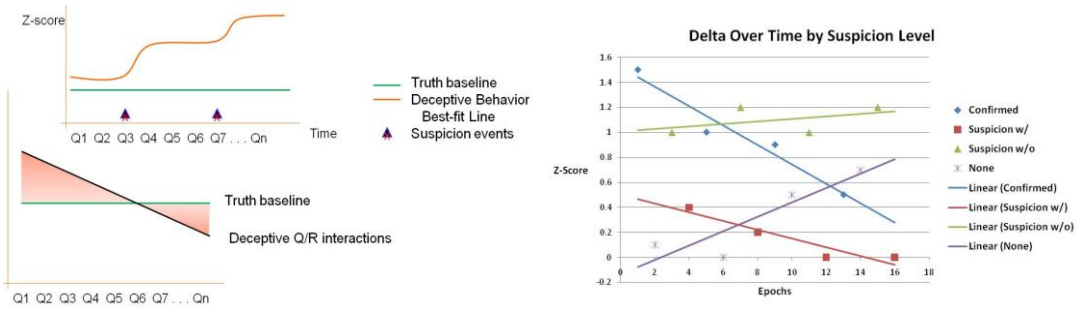


Figure 6, Example Graphical Analysis Charts

Above are only a few preliminary examples; final graphical analysis depended on what the data said and how best to display that information. A final graphical analysis method compared the epoch means of various granularities. Data was segmented into the following epoch sizes: grand means, by topic, by 1.6 hour block, and by question/response pairs. It was believed that the question/response pairs would show the most difference between truthful and deceptive measures. What follows are the results of exploring this case study data.

CHAPTER 4

RESULTS AND ANALYSIS

This chapter describes the results of the study and methodology established in the former chapters. On the first day, the interview lasted just over four hours and 10 minutes, during which 711 individual questions were asked covering 209 different topics.

Due to the methodology of this study and the volume of data analyzed the presentation of the results and analysis is a bit unconventional. First, descriptive statistics were generated to describe the basic features of the data including the distribution, central tendency, and dispersion. Analysis of variance was conducted to identify the constructs and measurements where there was a significant effect by the level of suspicion. This was followed by regressing the different individual measures of deceptive behavior as well as their constructs on to Suspicion for each level of granularity. Descriptive statistics, ANOVA, and regression results are presented in table-form in their entirety so they may be referenced. Next principal components analysis (PCA) was run to validate the constructs and to confirm the cues for each construct are measuring the same thing. Then, a graphical analysis was run on the data to better understand the relationships between levels of Suspicion and measures of deceptive behavior over time. Finally, the results of each proposition are presented. Again, this way of presentation is done because of the exploratory nature of the study and the volume of data analyzed.

Descriptive Statistics

With 70.7% of the measures showing increases during deceptive responses there is a general rise in behavior measures (Table 13). This could be explained by deceiver's tendency to over compensate because he is anxious to appear honest (Boltz, Dyer, & Miller, 2010).

Table 13, Descriptive Means of Measures

Cues	Mean (Raw)	Mean (Raw w/ Direction)			
		Truthful	↕	Deceitful	↕
# of Words	43.729	39.253	↓	66.429	↑
# of Verbs	3.338	3.000	↓	5.038	↑
# of Sentences	8.318	7.418	↓	12.829	↑
Sensory ratio	0.790	0.819	↑	0.656	↓
Spatial ratio	6.030	6.017	↓	6.111	↑
Temporal ratio	4.772	4.711	↓	5.150	↑
Content Word Diversity	0.803	0.817	↑	0.739	↓
Bilo. Type-Token-Ratio	79.378	80.742	↑	73.083	↓
Certainty Terms	3.022	3.037	↑	2.914	↓
Tentative Terms	3.111	3.072	↓	3.454	↑
Modal Verbs	10.495	10.248	↓	11.827	↑
Passive Voice	0.006	0.007	↑	0.005	↓
Generalizing Terms	2.331	2.273	↓	2.447	↑
Redundancy	18.926	18.667	↓	20.323	↑
Sentence Length	12.664	12.532	↓	13.495	↑
Complexity Ratio	2.5093	2.5086	↓	2.5108	↑
Average Word Length	3.822	3.827	↑	3.789	↓
Causation Terms	1.033	0.858	↓	1.990	↑
1st person pronouns	9.631	9.368	↓	11.294	↑
2nd person pronouns	0.740	0.714	↓	0.778	↑
3rd person pronouns	3.031	3.027	↓	3.113	↑
Activation	1.595	1.585	↓	1.660	↑
Imagery	1.405	1.399	↓	1.443	↑
Pleasantness	1.732	1.724	↓	1.783	↑

Cues	Mean (Raw)	Mean (Raw w/ Direction)			
		Truthful	↕	Deceitful	↕
Exclusive Verbs	3.137	2.943	↓	4.097	↑
Motion Words	2.106	2.042	↓	2.506	↑
Cognitive Proc. Terms	16.828	16.390	↓	19.250	↑
Non-ah disturbances	2.306	2.195	↓	3.051	↑
Speech errors	0.0097	0.0100	↑	0.0076	↓
Silent pauses	0.103	0.102	↓	0.100	↓
Filled pauses	2.010	2.058	↑	1.895	↓
Length of interaction	20.949	19.915	↓	26.334	↑
Response length	13.005	11.789	↓	19.279	↑
Talking time	13.001	11.784	↓	19.275	↑
Rate of speaking	4.940	4.915	↓	5.066	↑
Rate change	0.657	0.690	↑	0.514	↓
Amplitude	53.727	53.684	↓	53.886	↑
Amplitude variety	0.0142	0.0141	↓	0.0143	↑
Pitch	135.057	136.465	↑	125.577	↓
Pitch change	0.0530	0.0533	↑	0.0517	↓
Pitch variety	49.556	49.269	↓	51.050	↑

The above raw score mean table also gives a good initial understanding of the spread of the data. For example, *# of Words* averaged just over 43 with truthful statements, less at 39 and deceitful statements, and much more at 66 words on average. Looking at Construct means requires converting the individual measures to z-scores and averaging for each response. Again, the mean scores for all but three constructs (75%) increased during deceitful behavior (Table 14).

Table 14, Descriptive Means of Constructs

Constructs	Mean = 0 (Z-Score)			
	Truthful	↕	Deceitful	↕
Quantity	-0.096	↓	0.483	↑
Specificity	0.068	↑	-0.312	↓
Uncertainty	-0.018	↓	0.096	↑
Clarity	-0.034	↓	0.181	↑
Immediacy	-0.028	↓	0.151	↑
Affect	-0.021	↓	0.138	↑
Cognitive Proc.	-0.049	↓	0.266	↑
Fluency	-0.002	↓	0.013	↑
Duration	-0.077	↓	0.398	↑
Tempo	0.022	↑	-0.081	↓
Intensity	-0.013	↓	0.048	↑
Frequency	0.013	↑	-0.102	↓

Because all of the constructs are reflective (vs formative) it follows that changes in the individual cues reflect the changes in the latent constructs (Coltman, Devinney, Midgley, & Venaik, 2008; Petter, Straub, & Rai, 2007).

Before exploring each proposition, only those measures that showed potential for explanatory power were considered. To address the overarching research question, "*Are speech cues to deceptive behavior moderated over time by receiver suspicion during dyadic interactions in a real-world high-stakes setting?*" an ANOVA was run on the effect of suspicion on the measures to determine which cues are showing deceptive behavior.

ANOVA

An initial step to reporting ANOVA results should be to define what is "extreme". In other words, what is the cutoff value of α level of significance given the nature of the

study. Most linguistic and psycholinguistic as well as MIS journals enforce the conventional α of 0.05 (Baayen 2008, p 68). Because of the exploratory nature of this study Type II errors (failing to reject when the null hypothesis is in fact false) are more acceptable than Type I (rejecting the null hypothesis when in fact it is true). In practical terms, believing a treatment has an effect when in fact there is none (Type II error) is less damaging than dismissing a treatment that in fact has an effect (Type I error) (Murphy, Myors, & Wolach, 2009). Furthermore, given the uncontrolled environment from which the data was collected, a more relaxed **α of 0.10** is adopted. Where α levels of greater significance are noticed, they are reported. What follows in Table 15 are the ANOVA statistics on the individual measures z-score data.

Table 15, ANOVA of Suspicion on Individual Measures by Granularity

CONSTRUCT	MEASURES	Q/R Pairs		Topic		1.6Hr Blocks	
		$F_{3, 707}$	Sig.	F	Sig.	F	Sig.
Quantity	WordCount	9.64	0.000	1.22	0.192	2.88	0.035
	NumSentences	14.01	0.000	2.14	0.000	13.02	0.000
	NumVerbs	8.33	0.000	1.22	0.198	4.27	0.005
Specificity	SensoryRatio	0.37	0.778	0.69	0.892	1.17	0.320
	SpatialRatio	0.02	0.995	1.04	0.410	1.75	0.155
	TemporalRatio	0.41	0.744	0.99	0.490	0.21	0.889
	ContentWordDiversity	10.09	0.000	1.38	0.087	3.32	0.020
	BiloTTR	9.15	0.000	1.31	0.129	1.90	0.128
Uncertainty	CertaintyTerms	0.24	0.869	0.92	0.587	0.40	0.754
	TentativeTerms	0.51	0.674	0.73	0.849	1.05	0.371
	ModalVerbs	1.24	0.293	1.08	0.360	0.25	0.863
	PassiveVoice	0.33	0.801	0.50	0.988	0.59	0.624
	GeneralizingTerms	1.10	0.348	1.31	0.127	2.11	0.098
Clarity	Redundancy	1.05	0.370	1.15	0.262	0.48	0.695
	SentenceLength	0.68	0.563	0.62	0.947	5.22	0.001
	ComplexityRatio	0.06	0.983	1.72	0.010	1.53	0.204
	AvgWordLength	0.29	0.835	1.97	0.002	2.66	0.048

CONSTRUCT	MEASURES	Q/R Pairs		Topic		1.6Hr Blocks	
		$F_{3, 707}$	Sig.	F	Sig.	F	Sig.
	CausationTerms	6.33	0.000	1.68	0.013	0.88	0.452
Immediacy	1stppronoun	3.28	0.021	1.37	0.090	1.43	0.233
	2ndppronoun	0.95	0.414	1.51	0.040	1.01	0.387
	3rdppronoun	0.06	0.982	2.41	0.000	2.38	0.069
Affect	ActivationScore	1.48	0.219	1.20	0.214	1.60	0.187
	ImageryScore	0.40	0.755	1.32	0.122	2.01	0.111
	PleasantnessScore	0.73	0.536	1.16	0.253	0.72	0.540
Cognitive Processing	ExclusiveVerbs	1.95	0.121	1.13	0.286	0.28	0.839
	MotionWords	0.50	0.679	1.67	0.014	3.01	0.029
	CogProcTerms	1.76	0.154	1.16	0.256	1.01	0.387
Fluency	NonAhDisturbances	2.11	0.098	1.30	0.132	0.36	0.782
	SpeechErrors	0.35	0.788	0.90	0.619	0.09	0.967
	SilentPauses	1.27	0.283	2.37	0.000	20.51	0.000
	FilledPauses	0.43	0.734	0.98	0.499	1.12	0.338
Time-Duration	LengthOfInteraction	1.03	0.378	0.44	0.997	2.16	0.092
	ResponseLength	8.97	0.000	1.13	0.289	4.35	0.005
	TalkingTime	8.99	0.000	1.13	0.287	4.35	0.005
Time-Tempo	RateOfSpeaking	0.92	0.433	1.21	0.205	11.67	0.000
	RateChange	2.89	0.035	1.00	0.469	1.50	0.213
Intensity	AmpMeandB	0.75	0.520	2.40	0.000	5.77	0.001
	AmpVarietyPascals	0.60	0.616	2.27	0.000	2.97	0.031
Frequency	PitchHz	3.41	0.017	1.11	0.312	0.64	0.587
	PitchChange	0.74	0.527	1.64	0.018	0.78	0.504
	PitchVariety	1.53	0.206	1.26	0.165	4.37	0.005
# of Significant Measures		12		13		17	
% of Significant Measures		29.3%		31.7%		41.5%	

There was a significant effect of Suspicion on 12 measures of behavior namely WordCount, NumSentances, and NumVerbs, ContentWordDiversity, BiloTTR, CausationTerms, 1stppronoun, NonAhDisturbances, ResponseLength, TalkingTime, RateChange, and PitchHz ($2.89 < F_{3, 707} > 14.01$, $p < .035$). The measure of *Cognitive Processing*, ExclusiveVerbs, was close to being significant ($F_{3, 707} = 1.98$, $p = 0.121$). All 13 measures were included in further analysis to see if patterns could be identified.

In order to run the ANOVA on the constructs the data required manipulation so the aggregate of the different measures could be computed. All measures were given a z-score, the minimum score in a range was found and added back onto each individual measure. This generated a positive scale across individual measures allowing for a meaningful average for each construct for each level of granularity.

Table 16, ANOVA of Suspicion on Constructs by Granularity

CONSTRUCTS	By QR Pairs		By Topic		By 1.6 Hr Blocks	
	$F_{3, 707}$	Sig.	$F_{3, 707}$	Sig.	$F_{3, 707}$	Sig.
Quantity	11.567	.000	1.527	.037	5.981	.000
Specificity	5.932	.001	1.440	.062	3.918	.009
Uncertainty	.839	.473	.621	.945	.031	.993
Clarity	1.967	.118	1.208	.207	2.557	.054
Immediacy	1.015	.386	1.858	.004	2.370	.069
Affect	.915	.433	1.222	.194	1.473	.221
Cognitive Processing	3.138	.025	1.478	.049	.736	.531
Fluency	.705	.549	1.494	.045	7.470	.000
Time Duration	7.456	.000	1.131	.289	4.809	.003
Time Tempo	.922	.430	1.690	.013	25.530	.000
Intensity	.659	.578	2.372	.000	4.376	.005
Frequency	.684	.562	1.474	.050	.947	.417
# of Sig. Measures	4		8		8	
% of Sig. Measures	33.3%		50%		50%	

As seen in Table 16, there was a significant effect of Suspicion on four constructs by QR Pairs, eight by topic, and eight by 1.6 Hr Blocks ($1.46 < F_{3, 707} > 25.530$, $p < .055$). The construct *Clarity* by QR Pairs, was close to being significant ($F_{3, 707} = 1.98$, $p = 0.118$). 10 of the 12 Constructs are included in further analysis to see if patterns can be identified; the two that were omitted are *Uncertainty* and *Affect* due to their lack of significance.

There was a significant effect of Suspicion on behavioral measures when combined into a single mode Linguistics for both QR Pairs and topic level of granularity ($1.710 < F_{3, 707} > 4.653$, $p < .011$). However, the Paralinguistic mode was effected by Suspicion at all three levels of granularity ($3.277 < F_{3, 707} > 4.673$, $p < .021$).

Table 17, ANOVA of Suspicion on Modes by Granularity

Modes	QR Pairs		Topic		1.6 Hr Blocks	
	F	Sig.	F	Sig.	F	Sig.
Linguistic	4.653	.003	1.710	.011	.107	.956
Paralinguistic	3.277	.021	1.690	.013	4.673	.003

It appears that ANOVA provides strong support for the primary research question. Not only do ~32% of the individual cues measured show a significant effect of Suspicion but 10 of 12 constructs showed some degree of effect of Suspicion.

Regression

Regressing the behavioral measures on to the levels of Suspicion was also run in order to show how the dependent variable changes with respect to the independent variable. What follows is a regression of the individual measures and the constructs onto Suspicion to better understand the relationship between these variables. The intent of both ANOVA and regression are to identify those measures and constructs that warrant further analysis.

Table 18, Regression on Mode by Granularity

Dependent Variable	Predictors	R	R^2	P-Value
Linguistics	QRSUSP	0.115	0.013	0.002
	TOPICSUSP	0.063	0.004	0.092
	1.6HRSUSP	0.005	0.000	0.893
Paralinguistics	QRSUSP	0.089	0.008	0.018
	TOPICSUSP	0.085	0.007	0.023
	1.6HRSUSP	0.092	0.008	0.015

Table 18 shows that both linguistic and paralinguistic has a significant effect across almost all levels of granularity. However, the percentage of variance explained was extremely low. This would suggest that with additional data sets or when examined in a controlled environment these modes could be fruitful.

Table 19 also shows a significant effect across constructs by level of granularity with 50% being significant. Again the percentage of variance explained is extremely small. It can be argued that a small but reliable relationship in a real-world deception detection setting where no controls are in place is still valuable.

Table 19, Regression on Constructs by Granularity

Dependent Variable	IV	R	R^2	P-Value	IV	R	R^2	P-Value	IV	R	R^2	P-Value
Quantity	QRSUSP	0.200	0.040	0.000	TOPICSUSP	0.150	0.022	0.000	1.6HRSUSP	0.100	0.010	0.008
Specificity		0.139	0.019	0.000		0.114	0.013	0.002		0.093	0.009	0.013
Uncertainty		0.030	0.001	0.419		0.044	0.002	0.237		0.002	0.000	0.959
Clarity		0.065	0.004	0.085		0.005	0.000	0.901		0.086	0.007	0.022
Immediacy		0.065	0.004	0.081		0.095	0.009	0.011		0.053	0.003	0.160
Affect		0.057	0.003	0.131		0.015	0.000	0.692		0.046	0.002	0.221
Cognitive Processing		0.105	0.011	0.005		0.078	0.006	0.037		0.049	0.002	0.194
Fluency		0.006	0.000	0.870		0.012	0.000	0.741		0.140	0.020	0.000
Time Duration		0.162	0.026	0.000		0.101	0.010	0.007		0.076	0.006	0.044
Time Tempo		0.045	0.002	0.234		0.003	0.000	0.937		0.300	0.090	0.000
Intensity		0.015	0.000	0.682		0.023	0.001	0.537		0.130	0.017	0.001
Frequency		0.035	0.001	0.346		0.005	0.000	0.887		0.053	0.003	0.156

When regressing individual measures by level of granularity a pattern of increasing percentage of variance explained becomes evident as the dependent variable goes from *mode* to *construct* to *measure*.

Table 20, Regression on Measures by Granularity

Dependent Variable	IV	R	R^2	P-Value	IV	R	R^2	P-Value	IV	R	R^2	P-Value
WC		0.181	0.033	0.000		0.122	0.015	0.001		0.043	0.002	0.252
NumSentences		0.223	0.050	0.000		0.186	0.034	0.000		0.192	0.037	0.000
NumVerbs		0.170	0.029	0.000		0.121	0.015	0.001		0.051	0.003	0.173
SensoryRatio		0.033	0.001	0.381		0.015	0.000	0.695		0.050	0.003	0.181
SpatialRatio		0.002	0.000	0.963		0.032	0.001	0.393		0.085	0.007	0.024
TemporalRatio		0.014	0.000	0.702		0.009	0.000	0.817		0.025	0.001	0.505
ContentWordDiversity		0.182	0.033	0.000		0.134	0.018	0.000		0.052	0.003	0.165
BiloTTR		0.169	0.029	0.000		0.129	0.017	0.001		0.034	0.001	0.370
CertaintyTerms		0.011	0.000	0.778		0.030	0.001	0.426		0.020	0.000	0.589
TentativeTerms		0.017	0.000	0.654		0.004	0.000	0.925		0.027	0.001	0.469
ModalVerbs		0.072	0.005	0.054		0.057	0.003	0.132		0.002	0.000	0.952
PassiveVoice		0.022	0.000	0.554		0.018	0.000	0.636		0.017	0.000	0.648
GeneralizingTerms		0.013	0.000	0.737		0.028	0.001	0.451		0.033	0.001	0.387
Redundancy		0.042	0.002	0.262		0.004	0.000	0.915		0.001	0.000	0.968
SentenceLength		0.024	0.001	0.522		0.008	0.000	0.822		0.147	0.022	0.000
ComplexityRatio		0.001	0.000	0.988		0.033	0.001	0.385		0.018	0.000	0.624
AvgWordLength		0.031	0.001	0.409	T	0.059	0.003	0.116	1	0.056	0.003	0.139
CausationTerms		0.150	0.023	0.000	O	0.083	0.007	0.028	.	0.028	0.001	0.458
1stppronoun	Q	0.099	0.010	0.009	P	0.075	0.006	0.045	6	0.069	0.005	0.065
2ndppronoun	R	0.006	0.000	0.883	I	0.089	0.008	0.018	H	0.035	0.001	0.358
3rdppronoun	S	0.002	0.000	0.959	C	0.010	0.000	0.787	R	0.051	0.003	0.176
ActivationScore	U	0.073	0.005	0.052	S	0.009	0.000	0.801	S	0.069	0.005	0.067
ImageryScore	S	0.039	0.002	0.302	U	0.020	0.000	0.588	U	0.029	0.001	0.447
PleasantnessScore	P	0.046	0.002	0.219	S	0.012	0.000	0.758	S	0.031	0.001	0.414
ExclusiveVerbs		0.084	0.007	0.024	P	0.052	0.003	0.167	P	0.034	0.001	0.368
MotionWords		0.041	0.002	0.271		0.019	0.000	0.605		0.041	0.002	0.277
CogProcTerms		0.069	0.005	0.064		0.074	0.005	0.049		0.016	0.000	0.670
NonAhDisturbances		0.056	0.003	0.137		0.062	0.004	0.099		0.036	0.001	0.344
SpeechErrors		0.033	0.001	0.374		0.029	0.001	0.441		0.007	0.000	0.862
SilentPauses		0.018	0.000	0.639		0.021	0.000	0.568		0.213	0.045	0.000
FilledPauses		0.018	0.000	0.639		0.014	0.000	0.704		0.035	0.001	0.346
LengthOfInteraction		0.060	0.004	0.109		0.034	0.001	0.361		0.037	0.001	0.322
ResponseLength		0.178	0.032	0.000		0.112	0.013	0.003		0.079	0.006	0.036
TalkingTime		0.178	0.032	0.000		0.113	0.013	0.003		0.078	0.006	0.037
RateOfSpeaking		0.053	0.003	0.160		0.028	0.001	0.453		0.209	0.044	0.000
RateChange		0.095	0.009	0.011		0.031	0.001	0.409		0.077	0.006	0.041
AmpMeandB		0.024	0.001	0.515		0.021	0.000	0.582		0.149	0.022	0.000
AmpVarietyPascals		0.006	0.000	0.875		0.025	0.001	0.505		0.108	0.012	0.004
PitchHz		0.101	0.010	0.007		0.043	0.002	0.250		0.043	0.002	0.255
PitchChange		0.054	0.003	0.153		0.026	0.001	0.483		0.035	0.001	0.358
PitchVariety		0.079	0.006	0.034		0.081	0.007	0.031		0.122	0.015	0.001

Both ANOVA and regression examinations provide evidence that further analysis is warranted on the following: *modes* (linguistic & paralinguistic), *constructs* (Quantity, Specificity, Clarity, Immediacy, Cognitive Processing, Time-Duration, and Time-Tempo), and significant *Measures* in Table 20 above.

Next, PCA was run to test the degree to which the measures within the constructs have a single dimension; in other words, are the cues measuring the same construct and only that construct.

Principal Components Analysis

A PCA was run not only on the measures but on the constructs in the hopes to better understand the measures of each and to see if the constructs used by previous researchers (Fuller et al., 2009; Meservy 2007) held up in a RWHS case. PCA provides a method to reduce a complex data set to one of lower dimensionality to potentially reveal any hidden, simplified structures.

The number of components to extract was determined first by looking at eigenvalues greater than one in a Scree Plot (Figure 7).

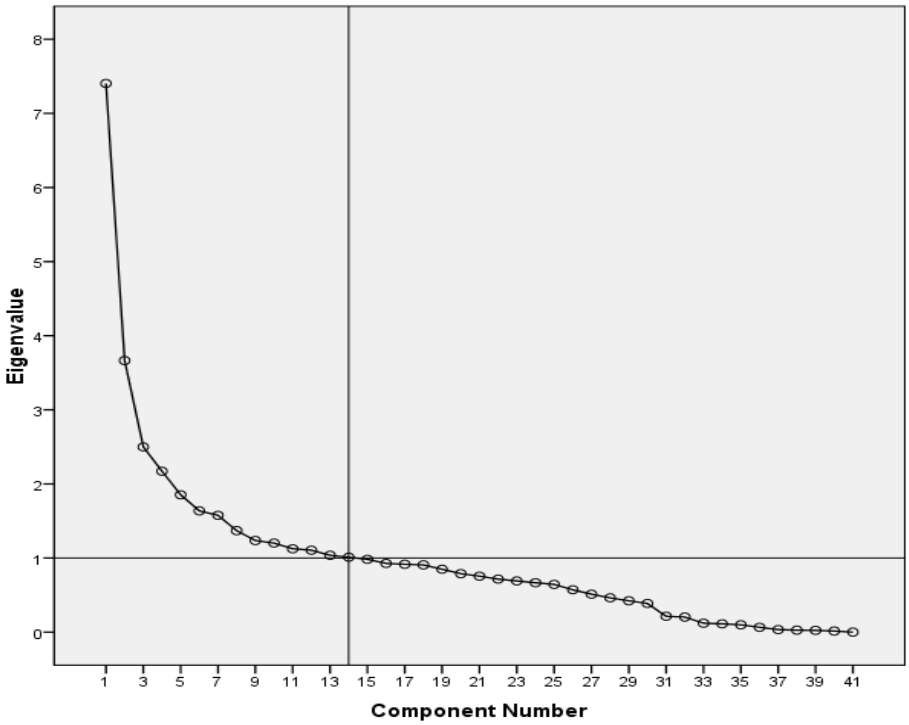


Figure 7, Scree Plot

The Scree Plot shows 14 components with eigenvalues greater than one. It is noted there appears to be a drop after seven components which may be a better number of components to extract. 14 components were fixed in the following PCA and special attention given to the first seven components if their cumulative variance explained is noteworthy.

A KMO Bartlett's Test of Sphericity was run to test whether a PCA is worthwhile in the first place. In order to proceed, Bartlett's Test should be significant and the PCA Kaiser-Meyer-Olkin Measure of Sampling Adequacy should be > 0.5.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.727
Bartlett's Test of Sphericity	Approx. Chi-Square	26520.117
	Df	820
	Sig.	.000

Figure 8, KMO & Bartlett's Test

As seen in Figure 8, the KMO test was significant and the Sampling Adequacy was 0.727, well above the common 0.5 accepted standard.

Communalities were also run which represent the percent of variance that is being accounted for by the components analysis for each factor. There is no standard percentage that is too low but for the current study 0.25 would be considered low. In Table 21, none of the factor communalities are below 0.25, in fact most are very high.

Table 21, PCA Communalities

Communalities					
	Extraction		Extraction		Extraction
WC	.952	SentenceLength	.520	SpeechErrors	.821
NumSentences	.801	ComplexityRatio	.902	SilentPauses	.566
NumVerbs	.918	AvgWordLength	.909	FilledPauses	.523
SensoryRatio	.701	CausationTerms	.675	LengthOfInteraction	.333
SpatialRatio	.581	1stppronoun	.632	ResponseLength	.957
TemporalRatio	.654	2ndppronoun	.614	TalkingTime	.958
ContentWordDiversity	.857	3rdppronoun	.338	RateOfSpeaking	.703
BiloTTR	.807	ActivationScore	.874	RateChange	.707
CertaintyTerms	.775	ImageryScore	.803	AmpMeandB	.928
TentativeTerms	.613	PleasantnessScore	.855	AmpVarietyPascals	.857
ModalVerbs	.605	ExclusiveVerbs	.652	PitchHz	.643
PassiveVoice	.669	MotionWords	.590	PitchChange	.783
GeneralizingTerms	.497	CogProcTerms	.857	PitchVariety	.340
Redundancy	.609	NonAhDisturbances	.506		

Because the measures are believed to be correlated a Direct Oblimin rotation was used. The Direct Oblimin is an approach to producing an oblique factor rotation, which means the factors can be correlated with each other.

If the factor solution that is the most appropriate in an orthogonal uncorrelated factor solution then the Direct Oblimin factor rotation procedure will yield a more or less orthogonal factor solution. Normally, there will be some correlation between factors in any real world examination, especially the current study, and Direct Oblimin will estimate those correlations. In other words, the best of both orthogonal and oblique rotation procedures are recognized.

Table 22, PCA Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	7.403	18.056	18.056	7.403	18.056	18.056	6.748
2	3.664	8.937	26.993	3.664	8.937	26.993	3.698
3	2.499	6.095	33.089	2.499	6.095	33.089	2.217
4	2.171	5.295	38.384	2.171	5.295	38.384	3.234
5	1.853	4.520	42.903	1.853	4.520	42.903	1.644
6	1.636	3.991	46.895	1.636	3.991	46.895	1.442
7	1.576	3.844	50.738	1.576	3.844	50.738	2.932
8	1.369	3.340	54.078	1.369	3.340	54.078	1.916
9	1.236	3.014	57.092	1.236	3.014	57.092	1.696
10	1.201	2.930	60.023	1.201	2.930	60.023	1.258
11	1.125	2.743	62.766	1.125	2.743	62.766	1.568
12	1.106	2.696	65.462	1.106	2.696	65.462	1.247
13	1.037	2.530	67.992	1.037	2.530	67.992	1.188
14	1.011	2.465	70.457	1.011	2.465	70.457	1.265
…							
41	.000	.000	100.000				

Extraction Method: Principal Component Analysis.

With 14 components 70.457% of the variance is explained (Table 22). Given the nature of the study and the large number of measures taken (41) this is a respectable reduction of factors and percent of variance explained.

What follows is the Pattern Matrix (Table 23) with absolute values < 0.3 subdued to help identify the nature of the components. Additionally PCA was run on the Constructs and their Pattern Matrix is on the right side of the table.

Table 23, PCA Pattern Matrix

MEASURES	Pattern Matrix														CONSTRUCT	Pattern Matrix				
	Component															Component				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14		1	2	3	4	5
WC	**.974**	.018	.012	.037	-.031	-.048	.041	.011	.035	.019	.023	.012	.008	.027						
NumSentences	**.898**	.044	.080	-.081	-.011	.124	.029	-.052	.000	.034	-.042	-.014	.021	.073	Quantity	**.933**	.028	.042	.035	-.043
NumVerbs	**.959**	.013	-.009	.022	.019	-.034	.054	.017	.023	.015	.044	.020	.013	.027						
SensoryRatio	-.015	-.026	.046	.042	.139	-.210	.172	-.026	-.049	.084	.003	.172	**.717**	.250						
SpatialRatio	.092	.259	-.045	.065	**-.615**	.059	.279	-.066	.014	.060	-.094	-.115	-.079	-.106						
TemporalRatio	-.006	.063	.046	.015	-.109	-.049	-.221	-.197	**.734**	-.082	-.107	.139	.008	.143	Specificity	**-.738**	.009	.026	.032	-.087
ContentWordDiversity	**-.764**	-.007	.076	-.084	-.090	.111	.134	-.024	.020	.289	-.072	.001	-.008	.080						
BiloTTR	**-.730**	.024	.047	-.076	-.120	.106	.138	.013	.040	**.329**	-.078	.046	.003	.069						
CertaintyTerms	.000	**-.318**	.137	-.083	.092	.127	.168	.033	**.712**	.105	.049	-.080	-.058	.023						
TentativeTerms	-.035	.034	.090	-.041	-.042	.107	-.039	**.747**	-.031	-.120	-.039	-.060	.074	-.049						
ModalVerbs	.089	**.303**	-.096	.085	**.643**	.030	-.062	-.017	.021	.083	-.147	-.158	-.025	-.087	Undertainty	-.012	-.015	**.745**	-.116	-.088
PassiveVoice	.028	-.048	.003	-.063	-.111	**.360**	-.188	.005	-.025	-.089	.052	-.147	**.651**	-.227						
GeneralizingTerms	-.036	.048	.133	.049	-.009	-.009	-.171	.290	-.098	.042	-.026	**-.482**	-.016	**.316**						
Redundancy	.129	.235	-.125	.040	**-.513**	-.062	**-.339**	.098	.034	.072	.151	.010	-.031	-.088						
SentenceLength	**.302**	.066	-.086	.247	-.019	**-.374**	-.142	.154	.091	-.038	.143	.050	-.006	-.085						
ComplexityRatio	.068	.202	**.892**	.091	-.014	-.029	-.038	.073	.104	.062	-.015	-.057	.045	-.040	Clarity	.135	.087	**.438**	**.399**	.202
AvgWordLength	.035	.092	**.932**	.073	-.022	-.027	.069	.031	.078	.053	-.053	-.151	.039	-.022						
CausationTerms	.002	-.035	.028	-.001	.127	-.010	.018	.010	.000	.032	**.808**	.102	-.014	.054						
1stppronoun	.047	**.358**	-.236	.024	**.536**	.101	.023	-.046	.029	.047	.169	-.101	-.067	-.237						
2ndppronoun	.065	.119	-.104	.056	-.037	.016	-.007	-.061	.115	-.072	.036	-.029	.023	**.772**	Immediacy	.109	-.056	-.204	**.704**	-.295
3rdppronoun	.084	-.079	.128	-.038	.045	**-.493**	-.037	-.021	-.243	.060	.009	.037	-.019	.025						
ActivationScore	.002	**.902**	.045	.024	.082	.077	-.027	.078	-.045	-.028	.041	.064	-.008	.037						
ImageryScore	-.015	**.818**	.187	.010	-.121	-.027	-.107	-.080	-.087	.114	.070	-.012	.026	.048	Affect	-.083	.086	.103	**.760**	.287
PleasantnessScore	.010	**.907**	.086	-.014	.049	.028	.008	.071	-.018	-.097	.020	.055	-.022	.076						
ExclusiveVerbs	.015	.023	-.051	.024	-.021	-.101	.017	**.775**	-.053	.036	-.043	.201	-.074	-.003						
MotionWords	.032	.130	-.073	.055	-.250	.014	.032	-.101	-.032	.038	**.654**	-.156	.067	.011	Cognitive Processing	.041	-.026	**.783**	.014	-.024
CogProcTerms	.072	-.118	.104	-.096	.187	.111	.075	**.498**	**.618**	.101	.212	-.063	-.062	-.068						
NonAhDisturbances	-.016	-.018	**.449**	-.077	-.031	.022	-.111	-.081	-.069	-.239	.103	**.312**	-.151	-.220						
SpeechErrors	.048	.009	-.043	.005	-.022	.064	.093	.069	-.004	**-.910**	-.072	-.076	.002	.095	Fluency	.053	-.025	-.102	.017	**.903**
SilentPauses	.097	.025	.051	.259	.114	**.648**	.106	.040	-.114	.016	-.010	.156	.001	.058						
FilledPauses	-.015	.072	-.055	.055	-.075	.035	.019	.190	.018	.065	-.005	**.695**	.009	.044						
LengthOfInteraction	**.462**	-.193	-.057	.030	-.076	.176	-.079	-.007	-.068	.172	-.077	-.027	-.024	-.038						
ResponseLength	**.984**	.025	.057	-.056	-.029	-.041	.024	.006	.020	.024	.010	.018	-.004	.027	Time-Duration	**.935**	.000	.030	-.029	-.047
TalkingTime	**.984**	.025	.057	-.056	-.029	-.041	.024	.006	.020	.024	.010	.018	-.005	.027						
RateOfSpeaking	.050	.006	.059	-.064	.074	-.077	**-.807**	.000	.013	.033	-.028	-.056	.041	-.022	Time-Tempo	-.006	**-.659**	.076	.205	-.068
RateChange	.009	-.039	.041	-.197	-.028	.033	**.796**	-.022	-.054	-.020	.014	.025	.049	-.031						
AmpMeandB	-.015	-.046	.056	**.954**	.028	.185	-.023	-.017	-.067	-.004	.036	.012	-.037	.110	Intensity	.079	**.763**	-.039	.113	.053
AmpVarietyPascals	-.004	-.087	.074	**.893**	.021	.228	-.103	-.039	-.106	.003	.057	.032	-.072	.134						
PitchHz	-.066	-.191	.060	**-.444**	.035	.251	.003	-.098	-.188	-.003	.056	-.015	-.259	**.309**						
PitchChange	.030	-.112	-.002	**-.749**	.040	.250	-.028	-.017	-.108	-.002	.026	.046	-.121	.189	Frequency	.081	**-.670**	-.125	-.147	.205
PitchVariety	.191	.151	.011	-.162	.086	.164	-.263	.018	-.045	.037	-.042	.292	.092	.068						

Extraction Method: Principal Component Analysis. | Rotation Method: Oblimin with Kaiser Normalization

For the Construct PCA the KMO was significant at 0.000 and Sampling

Adequacy of 0.638. The Communalities ranged from 0.452 - 0.887 and the total variance

explained by five components was 64.881%. Finally the largest Component Correlation Matric value was 0.172 showing little correlation between components.

There are several very strong constructs identified in the Pattern Matrix which are discussed, namely components 1-4, 7 and 10. The remaining constructs explain a fair amount of variance but do not seem to fall into a logical pattern. Component one of the Measures PCA encompasses all the *Quantity* and *Time-Duration* measures and could be described collectively as Amount with a negative correlation with Content Word Diversity and Bilo Type Token Ratio. This would indicate that as the amount of speech generated increased the diversity and variety of linguistics decreased at the same time. Next, Component two seems to center on the *Affect* construct with minor impact from other measures so its measures of impact and emotion seems to correlate well and of equal strength. The third Component contains two strong measures from the *Clarity* Construct, Complexity Ration and Average Word Length and NonAhDisturbances from the *Fluency* Construct. This component seems to show a pattern of increased complexity paired with an increase in fluidity of speech. This could be described as Confidence in the subjects' speech behavior. Component four is composed of four paralinguistic measures of amplitude and frequency and could best be described as Voice Propagation. This is a very logical connection and would seem to suggest that the constructs *Intensity* and *Frequency* should be combined in future studies. The final two components of interest are #7 and #10. Component seven is primarily made up of the *Time-Tempo* construct measures as is best described as such. Finally, Component 10 has several contributing measures but none as strong as SpeechErrors (negatively). In future

research perhaps the Construct *Fluency* could be measured by counting speech errors alone.

If an information technology tool was built on the smallest set of measures the constructs Quantity, Affect, Intensity, and Time-Duration seem to score well on the PCA; these cues would also be easy to measure with existing technology.

Propositions

Next the results and analysis of each proposition will be discussed. References to the above descriptive statistics, ANOVA, and regression tables will be made along with other pertinent methods like PCA and graphical analysis as evidences for their support.

P1: The difference in speech cues to deceptive behavior will be moderated over time by the level of suspicion.

Supported: Considering the mean scores from the descriptive statistics Table 13 and categorizing them by suspicion level, by time (quartiles) results in Figure 9 below. Behavioral scores from beginning to end changed and they changed more drastically for deception than for truth. The choice of quartiles was because the individual units of time between question and response pairs differed.

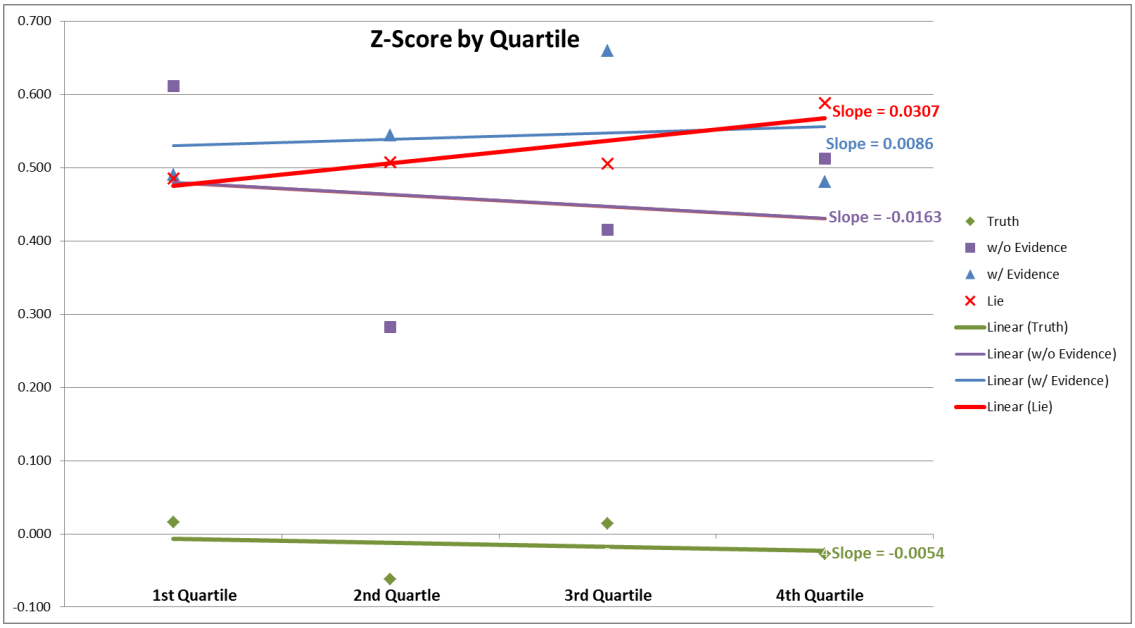


Figure 9, Z-Score by Suspicion by Quartile

The above chart shows the actual z-scores but is difficult to interpret or see the differences. Therefore, the scores were adjusted to have a common starting point to make their slope difference more apparent.

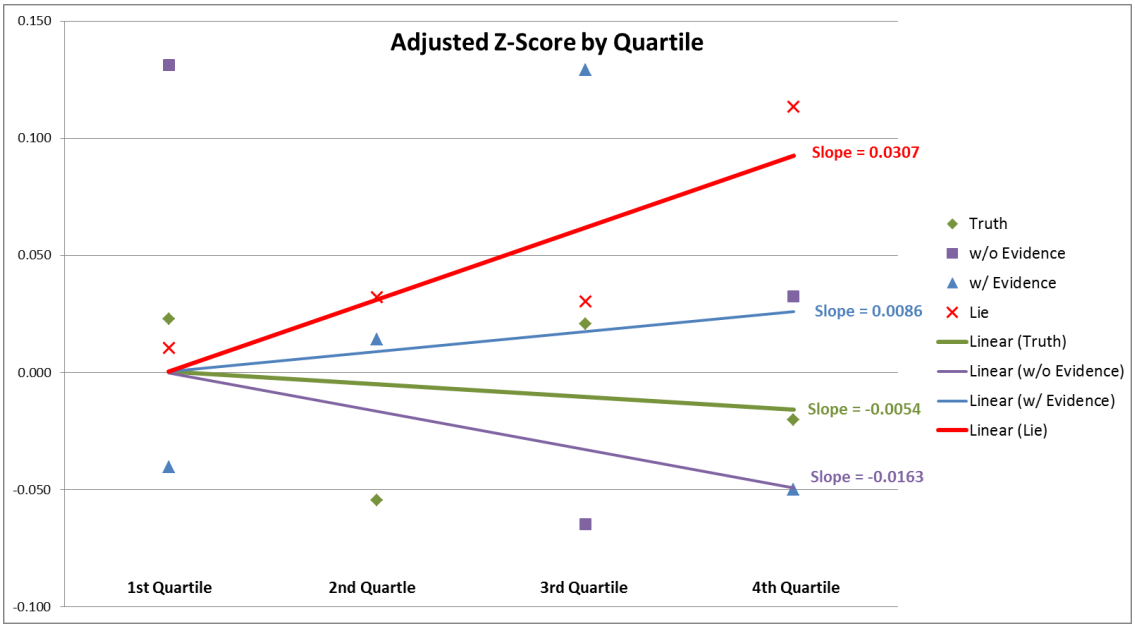


Figure 10, Adjusted Z-Score by Quartile

The z-scores for each level of suspicion were adjusted so they all start at zero but maintain their slope. This clearly shows that scores for truthful statements tended to decrease overtime while deceptive statements increased. It should be noted that these charts looked at all measures, not just the significant ones in the ANOVA Table 15. If only the significant measures from ANOVA Table 15 are used the difference over time becomes even more noticeable (Figure 11).

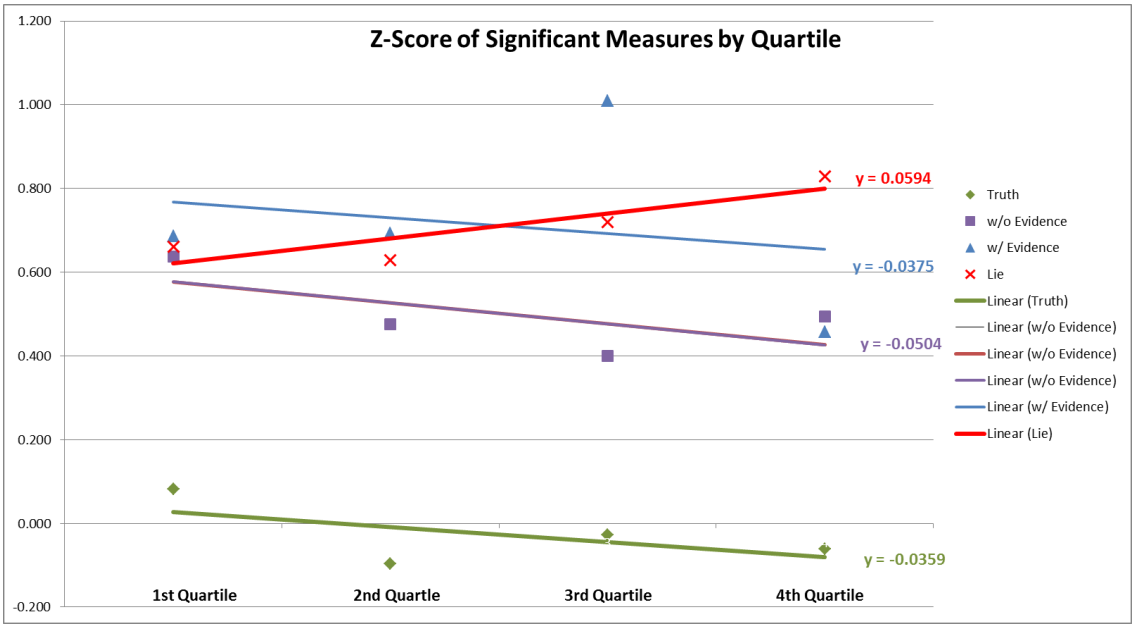


Figure 11, Z-Score of Significant Measures by Quartile

Though still small, the slopes change over time even more so; this is to be expected, but makes the differences easier to see.

P1A: The magnitude of speech cues to deceptive behavior will be positively related to an increase in level of suspicion.

Supported: The magnitude of deceptive behavioral cues increased as suspicion increased. One concern is the drop from w/ Evidence to Lie which should be looked at in future research.

Figure 12 – Figure 15 examine the significant constructs of *Quantity*, *Specificity*, *Cognitive Processing*, and *Time-Duration* looking at average cue scores across the four levels of suspicion. All scores in these charts are z-scores to enable direct comparisons.

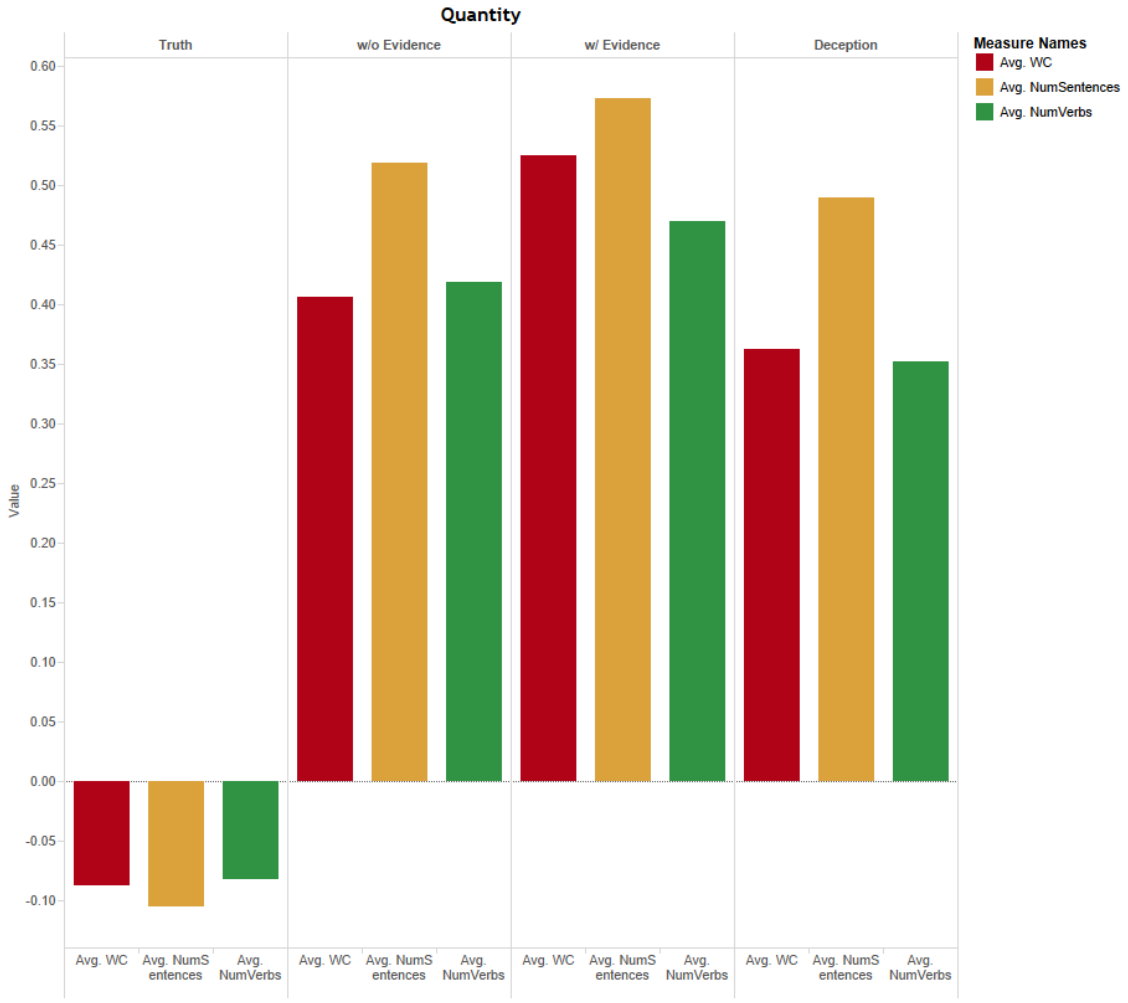


Figure 12, Quantity

The average magnitude for cues in the Quantity construct clearly increase as Suspicion increases from Truth to Deception. However, the differences within the degrees of evidence are not clearly increasing. This may not be of concern with the exception of the w/o Evidence deception scores. One explanation for this maybe that the w/o Evidence level of suspicion had a very small sample size.

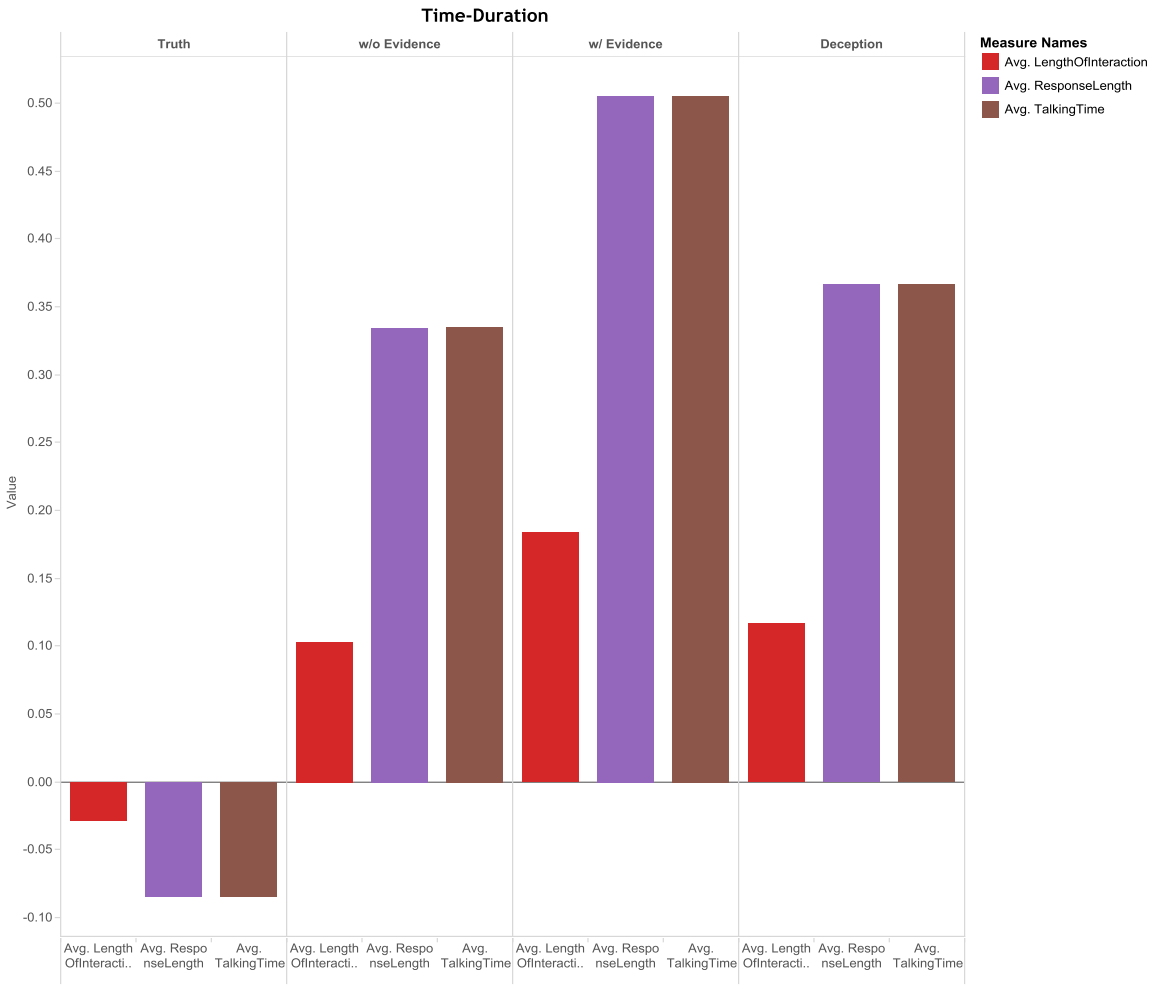


Figure 13, Time-Duration

The scores for Quantity and Time-Duration construct measures mirror each other. This is to be expected, the more the subject says the longer it takes. Again, the magnitude of each cue increases as Suspicion moves from Truth to Deception.

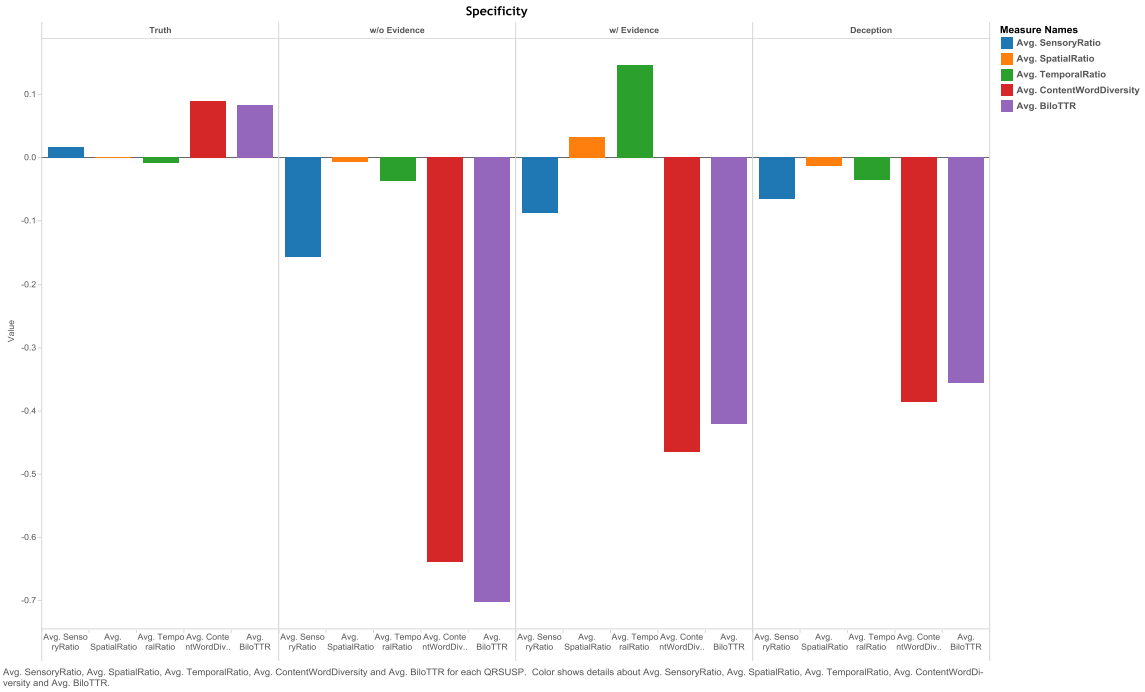
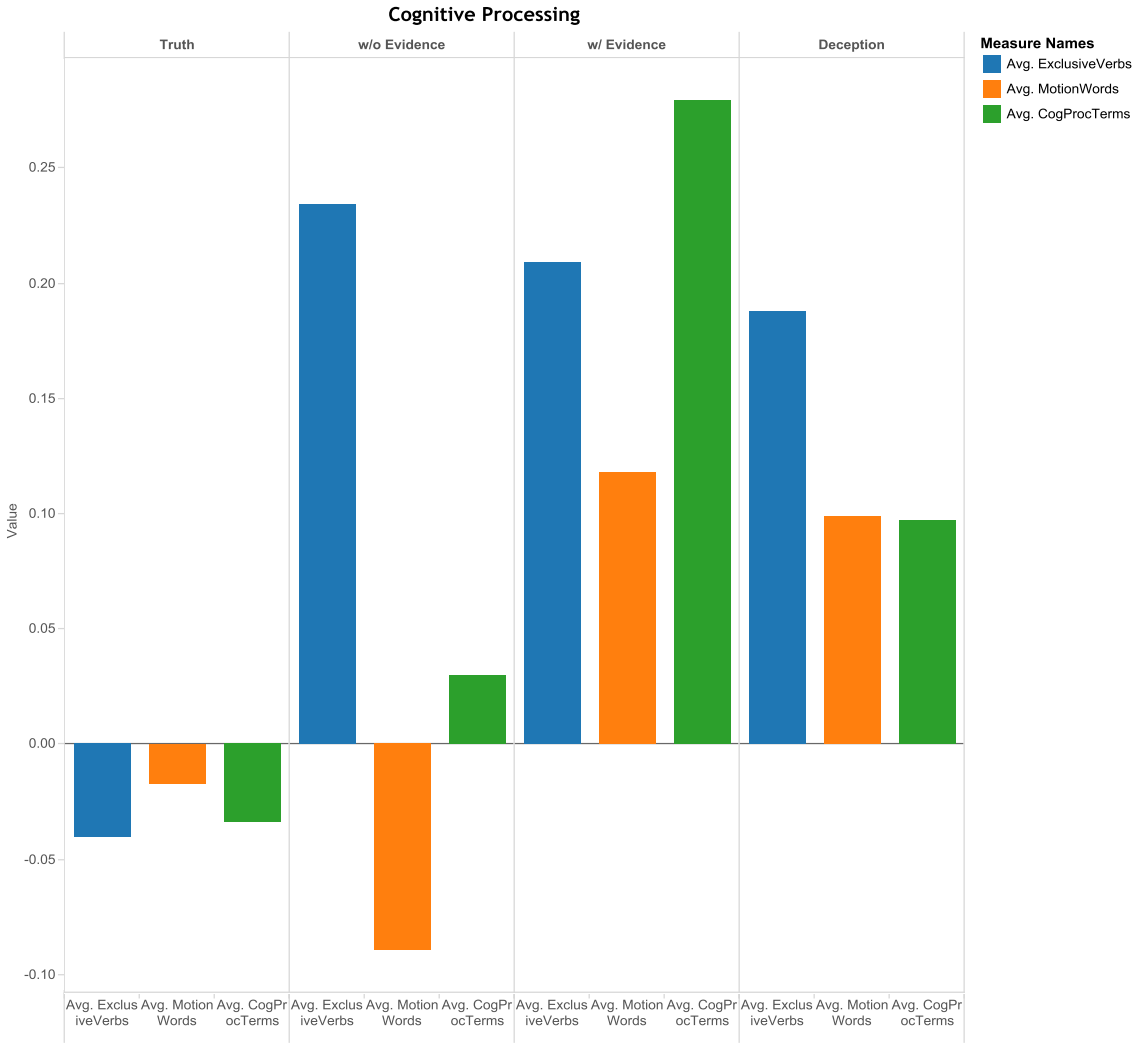


Figure 14, Specificity

Only two measures of Specificity, Content Diversity and BiloTTR significantly contributed to this construct with a decrease in average z-score magnitude as Suspicion moved from Truth to Deception.



Avg. ExclusiveVerbs, Avg. MotionWords and Avg. CogProcTerms for each QRSUSP. Color shows details about Avg. ExclusiveVerbs, Avg. MotionWords and Avg. CogProcTerms.

Figure 15, Cognitive Processing

The Cognitive Processing construct measures are not as consistent across levels of suspicion as the other significant constructs. Again, the w/o Evidence Suspicion level data may be skewed due to small sample size. The following charts consider the average z-scores for individual measures.

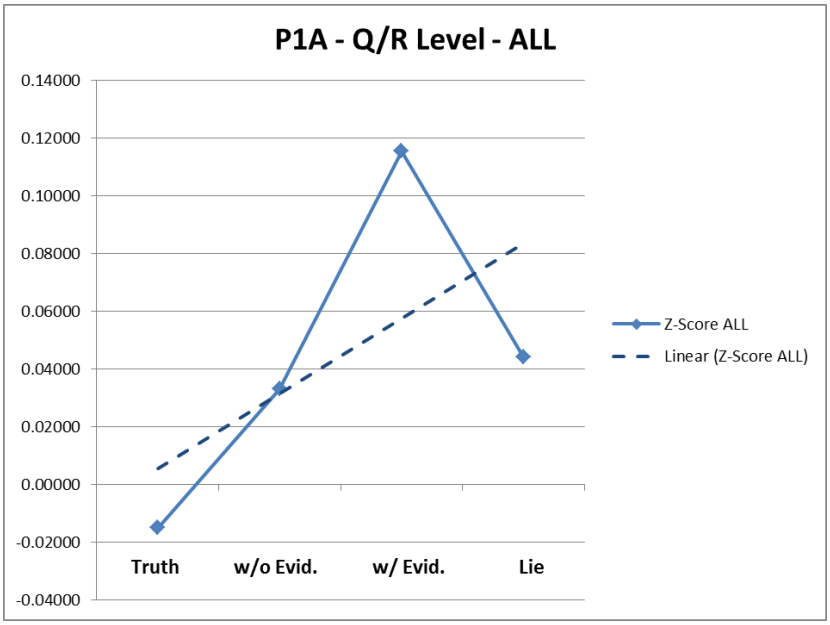


Figure 16, P1A - Q/R Level - All

Figure 16 shows all measures with a significant increase in z-score by level of suspicion. To better evaluate the magnitude Figure 17 shows only z-scores for significant measures from Table 20, Regression on Measures by Granularity, which is at Q/R Level.

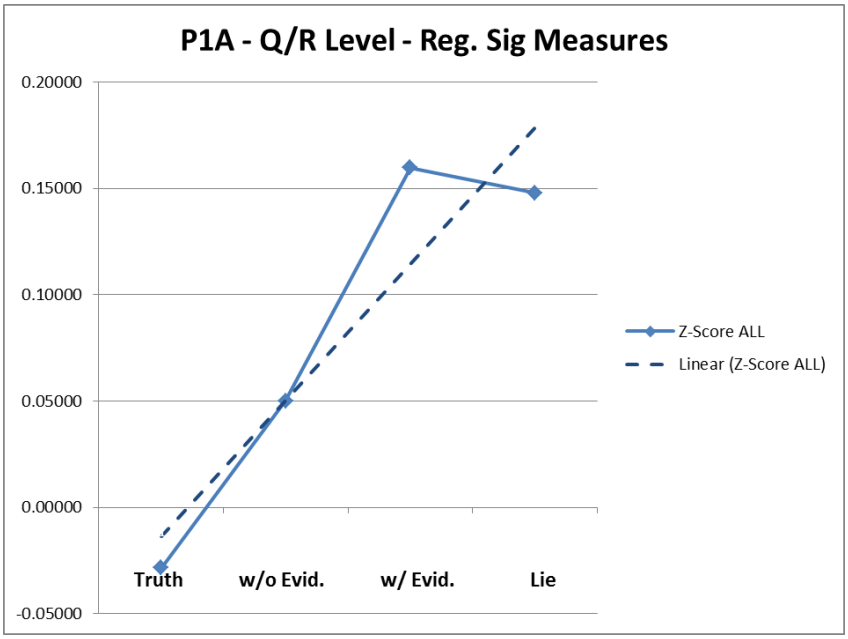


Figure 17, P1A - Q/R Level - Regression Significant Measures

P1B: The level of moderation by suspicion will be most detectable when epochs are at the topic level.

Opposite Supported: The opposite of this proposition was measured. When considering all behavioral cues as a single measure the highest variance explained occurred when looking at a granularity of 1.6 hrs and the lowest during topic level. To test if suspicion predicts behavior across different levels of epoch look at Table 20, Regression on Measures by Granularity, and count the number of behavioral cues which are significant at each level. The results are inconclusive at 16 significant measures by Q/R pairs, and 13 by both topic level and 1.6 Hr blocks.

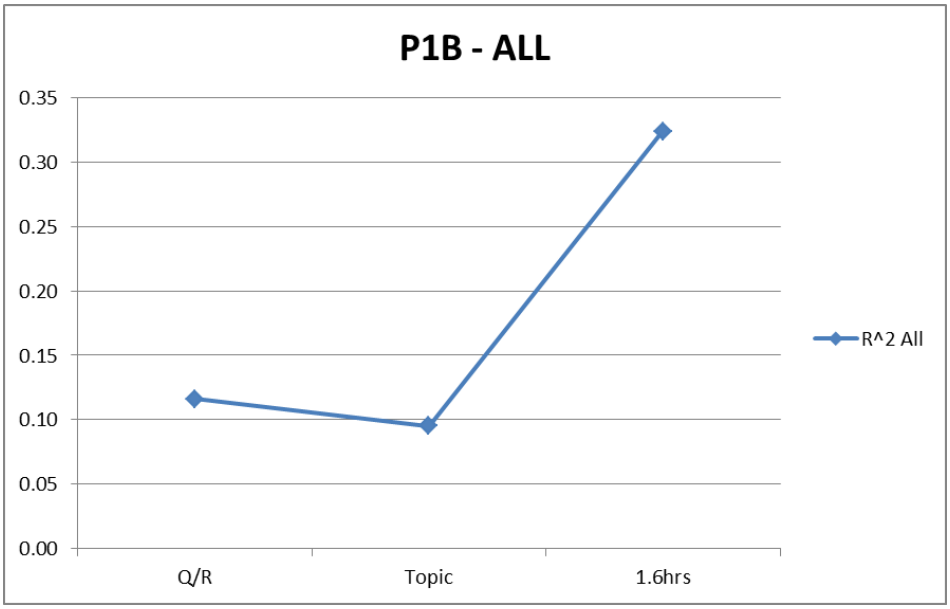


Figure 18, P1B - All

The following figure then considered modality to see if the strength of the relationship between IV and DV was strongest at the topic level, again it was not.

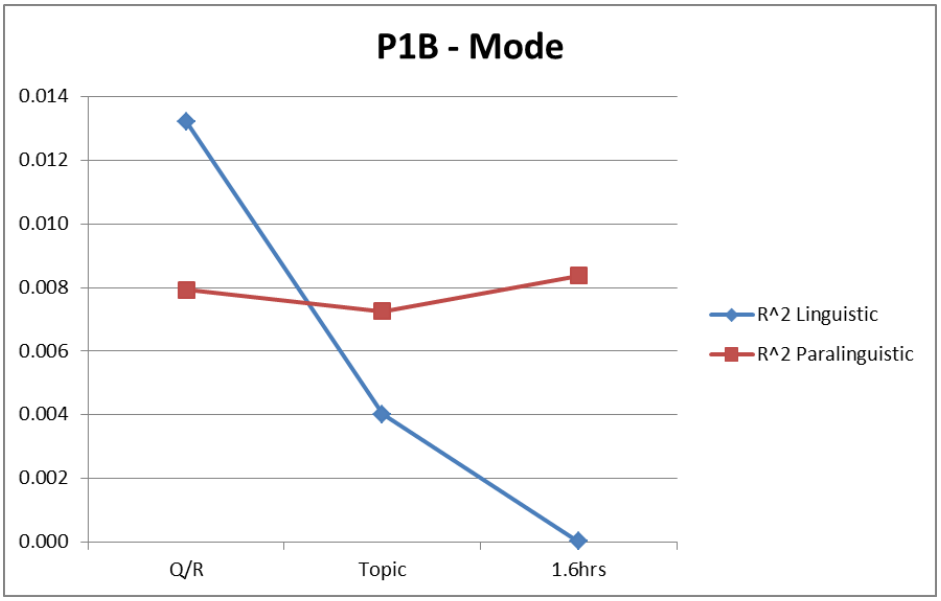


Figure 19, P1B - Mode

Finally, each significant construct was considered and the strength of the IV DV relationship for each was compared. Again, topic level granularity is not the strongest.

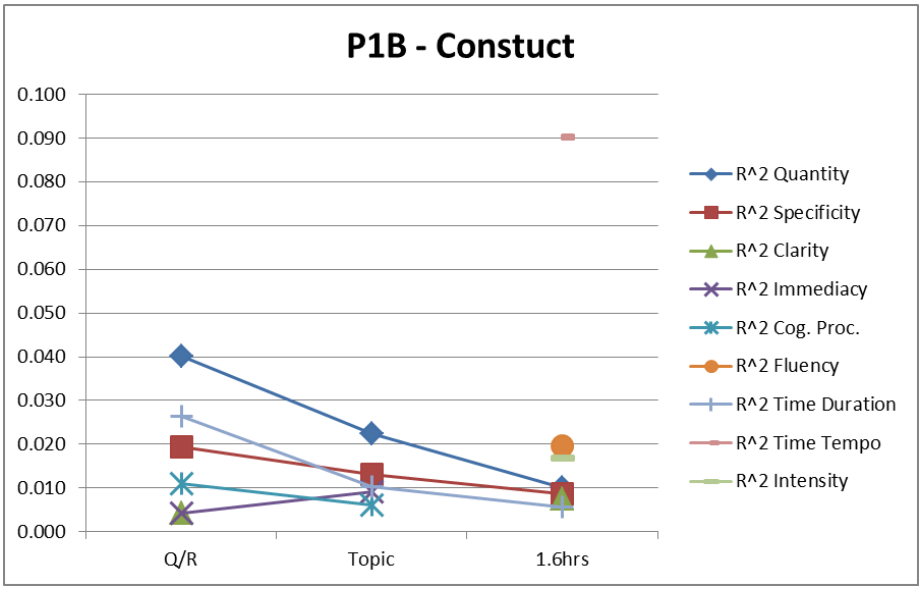


Figure 20, P1B - Construct

Though the behavior of the cues across different epochs is easy to observe, without additional data sets this proposition cannot be proven.

P2: Speech cues during deceptive behavior will form identifiable patterns that differ from speech cue patterns during truthful behavior.

From the descriptive statistics a general pattern appears; over 70% of the individual measures increased during deception (Table 13). The same general patterns hold when looking at Constructs, the mean scores for all but three constructs (75%) increased during deceitful behavior (Table 14).

Cluster analysis results were not conclusive. The matrix scatter plot of the significant constructs show dispersion in all but *Quantity* vs Time Duration. It is logical for these two constructs to be linearly correlated; as the quantity of words, verb, and sentences increases so does the amount of time to say those words.

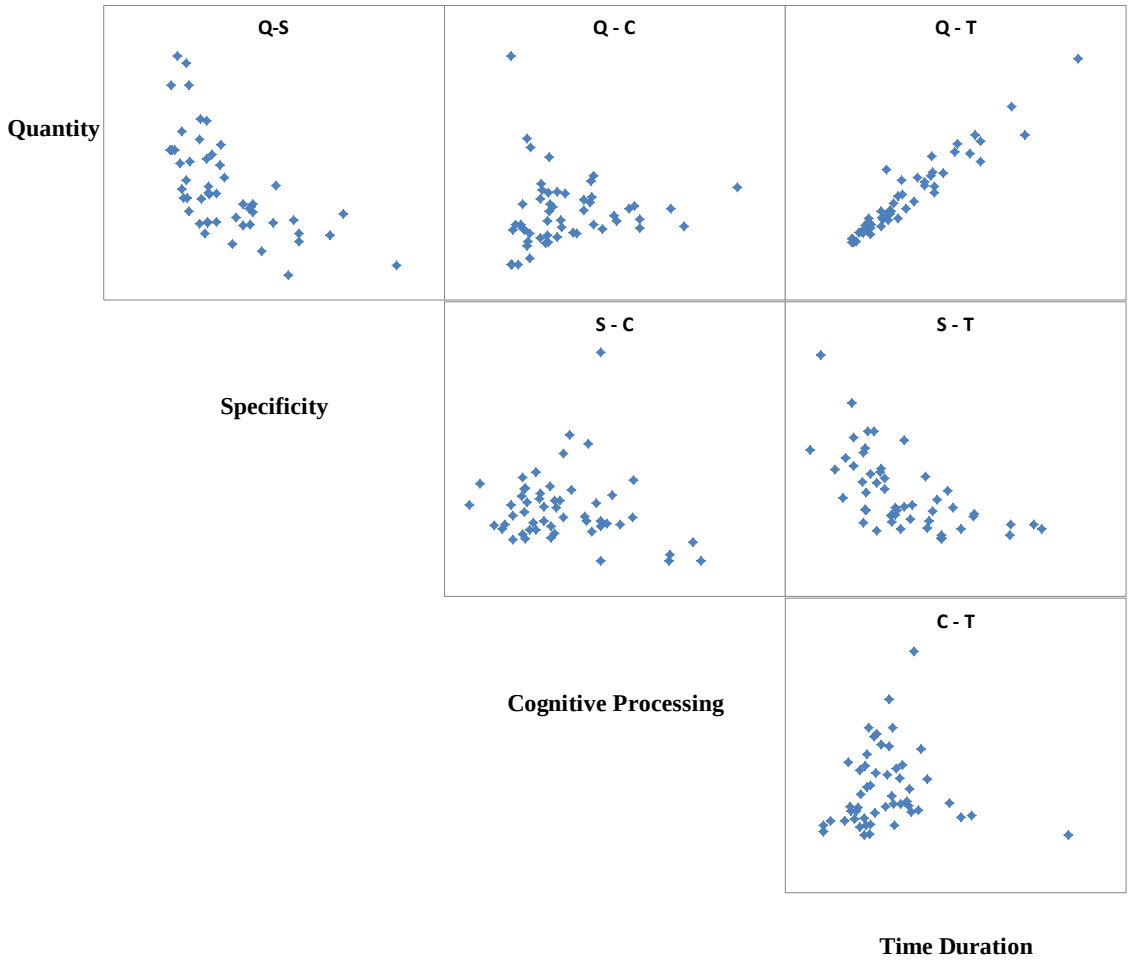


Figure 21, Cluster Matrix

Graphical analysis revealed promising results. Figure 22 through Figure 25 show a comparison of truths (blue) to lies (red) over time for three constructs and both modes (the remaining construct graphs can be found in Appendix B).

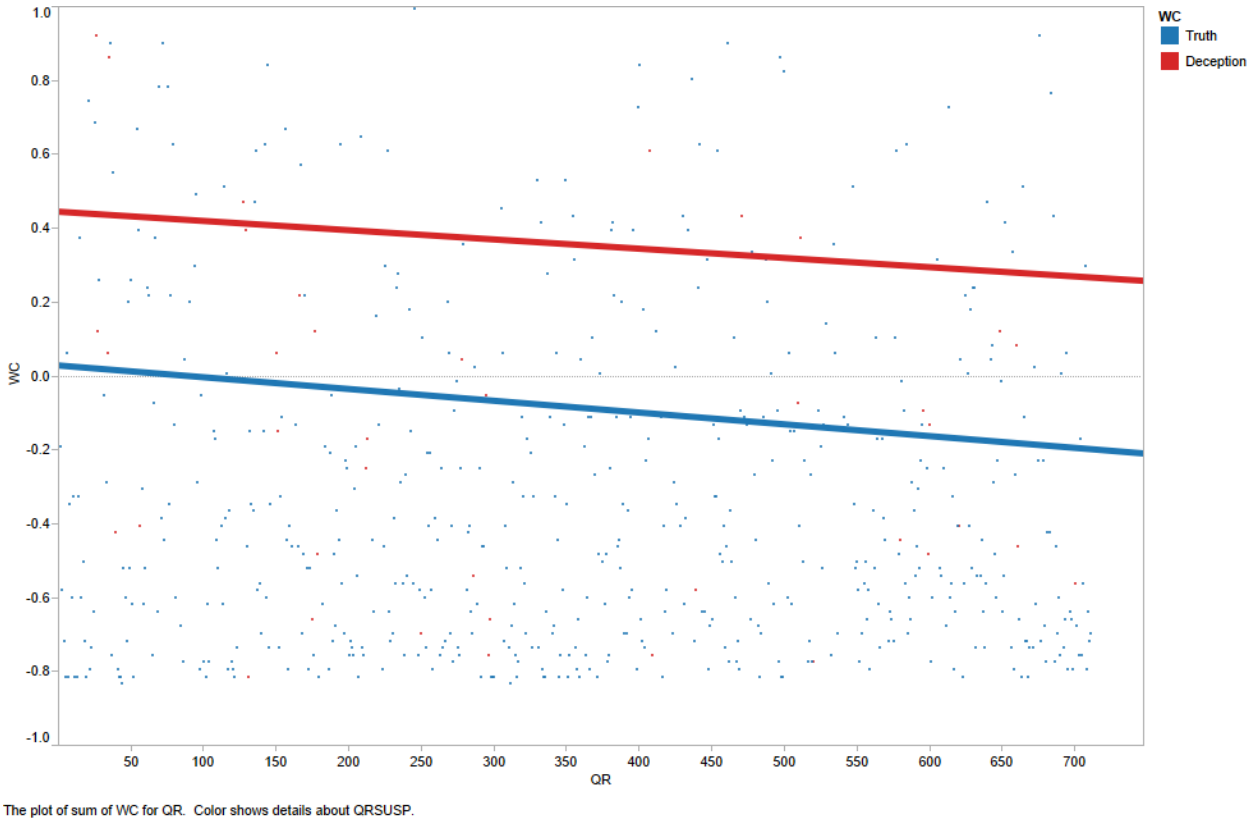


Figure 22, Word Count

Figure 22 shows how Word Count decreases almost uniformly regardless of level of suspicion. One explanation for this pattern could be fatigue (Ramdharry, Thornhill, Mein, Reilly, & Marsden, 2012). After four hours the subject could just be tired of talking. However, there is a stark difference in the number of words spoken when comparing truthful vs deceptive patterns which stay relatively constant, a pattern in and of itself.

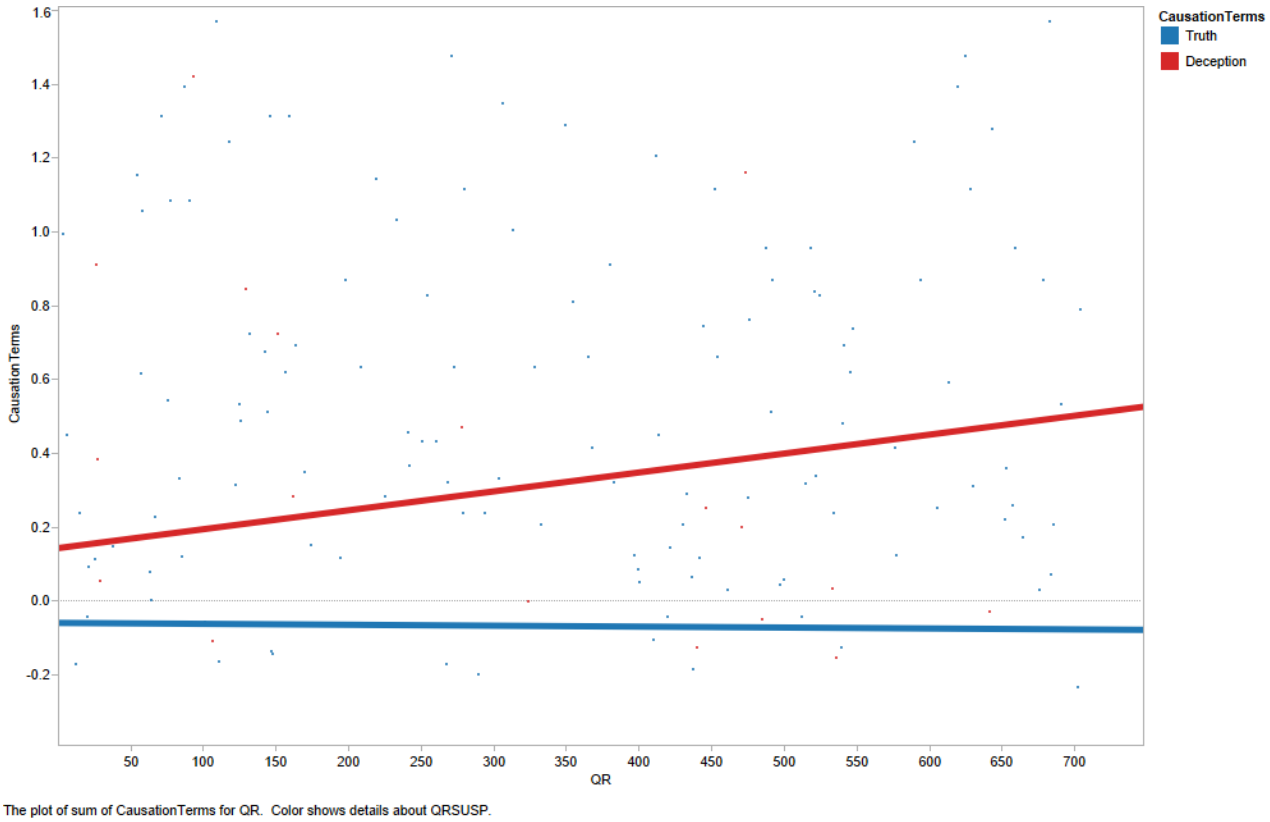


Figure 23, Causation Terms

In comparison to Word Count, the Causation Terms construct shows a relative constant truthful behavior while behavior measures during deception increases overtime. One explanation for this pattern could be overcompensation continuously during attempts to deceive (ten Brinke, 2012; Vrij, Mann, Leal, & Fisher, 2010).

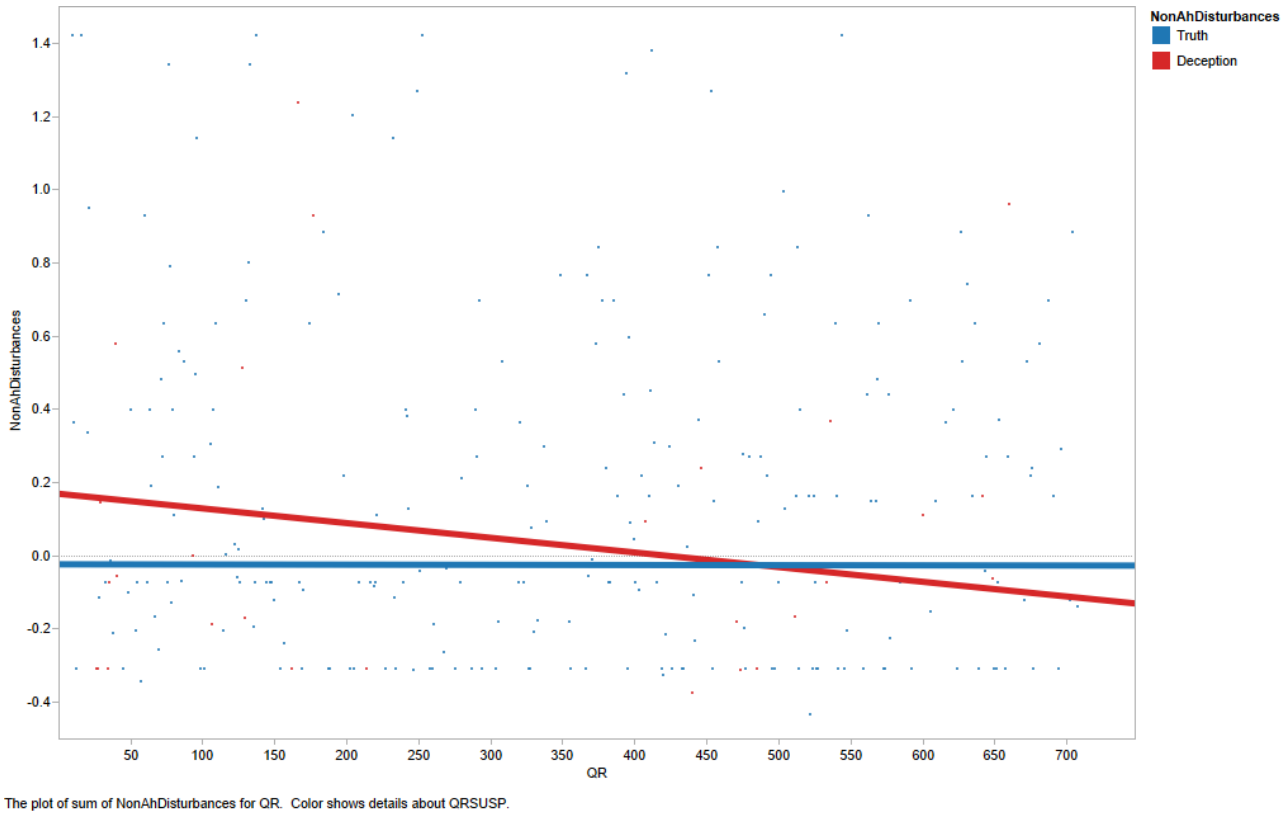


Figure 24, NonAhDisturbances

Similar to the Causation Terms constructs, the NonAhDisturbances construct shows consistent scores during truthful behavior. However deceptive behavior resulted in a slight decrease in the number of NonAhDisturbances. During a long interview the subject could become more comfortable in his deception resulting in a more fluid speech (Hu, Chen, & Fu, 2012). However this change is very slight and without additional subjects to test this could be just a chance occurrence.

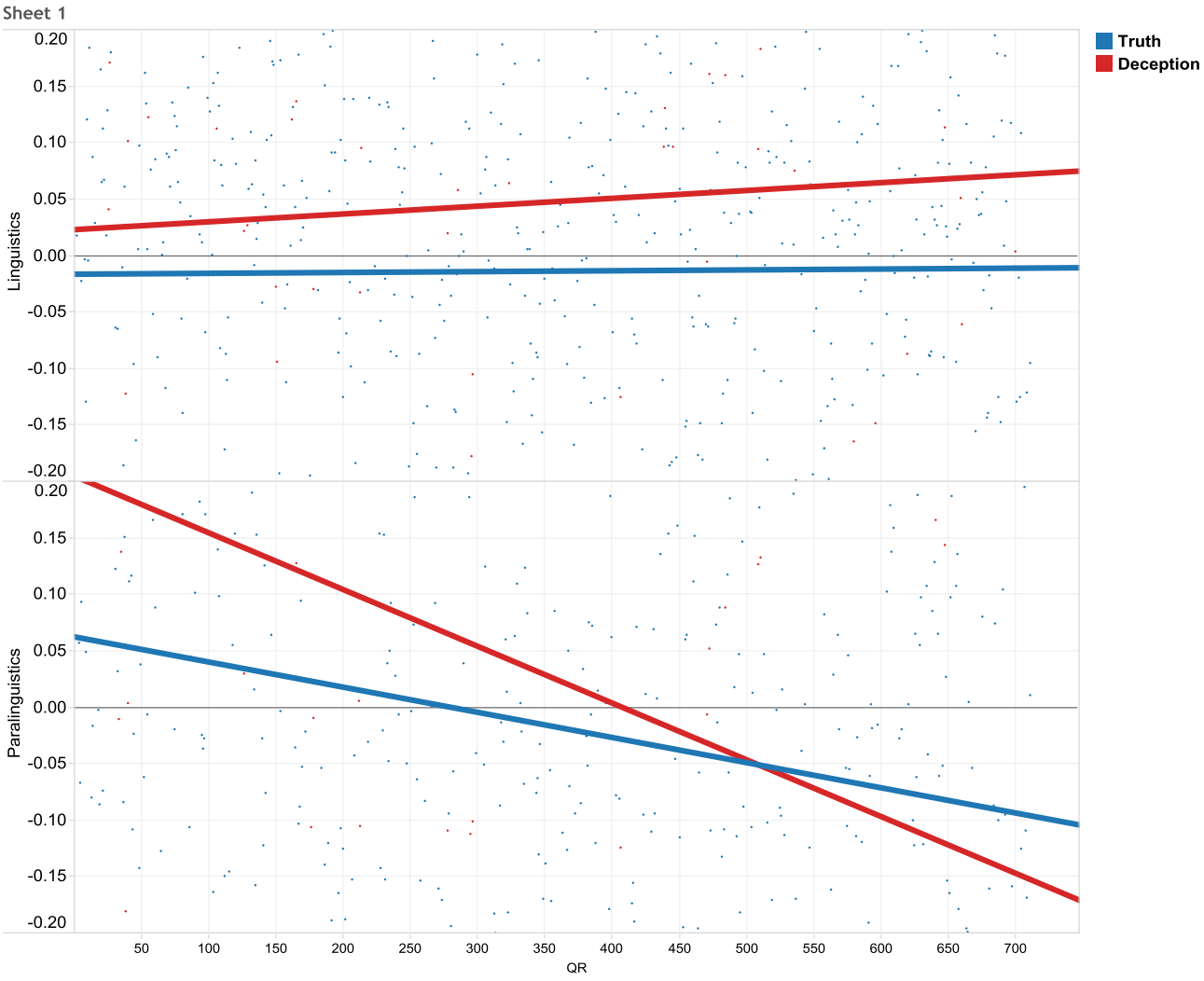


Figure 25, Linguistic & Paralinguistic

The final trend line graph compares overall behavior scores for linguistic and paralinguistic modes throughout the interview (Figure 25). Overall linguistic behavior scores, though slightly elevated during deception were relatively constant. There appears to be no support to looking at all linguistic cues when they are combined together. For the paralinguistic mode, scores trend downward (i.e. possible fatigue) but deceptive scores decrease more quickly.

P2A: Patterns of speech cues during deceptive behavior will be most detectable when epochs are at the topic level.

Not Supported: Though identifiable patterns of speech cues during deceptive behavior were substantiated, no support could be found that would suggest these patterns were more detectable at the topic level. The basic statistics did not show support for proposition P2A therefore no further analysis was warranted. Summarizing the ANOVA results (Table 24) from below, the number of significant differences between means by epoch showed no consistent trends to base a pattern on.

Table 24, ANOVA Summary

ANOVA Summary	Total	Q/R Pairs	Topic	1.6Hr Blocks
# of Significant Measures	41	12	13	17
# of Significant Constructs	12	4	8	8
# of Significant Modes	2	2	2	1

Similar to ANOVA, considering the Regression Summary (Table 25) and counting the number of significant (linear) relationship between a dependent and independent variables gives no indication of an increase in significance at the epoch topic level.

Table 25, Regression Summary

Regression Summary	Total	Q/R Pairs	Topic	1.6Hr Blocks
# of Significant Measures	41	16	13	13
# of Significant Constructs	12	6	5	7
# of Significant Modes	2	2	2	1

P3: As the granularity of epochs decreases (epochs gets smaller), the number of speech cues to deceptive behavior will increase to an apex then decrease.

Opposite Supported: The opposite of this proposition was measured; as the granularity increased (epochs got larger) the number of significant speech cues increased but not to an apex then decreased.

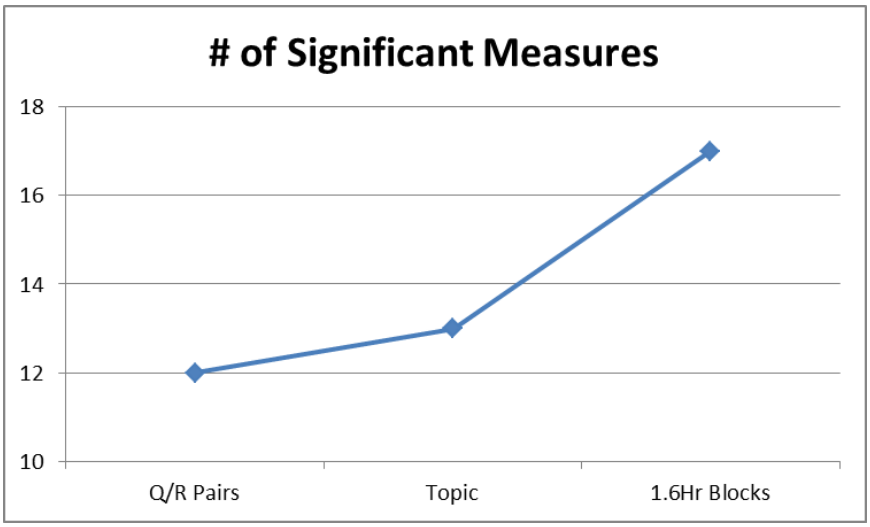


Figure 26, # of Significant Measures

Due to a lack of support when looking at individual measures, consideration of the number of significant constructs was warranted. Again, the number of significant constructs decreased but this time plateaued.

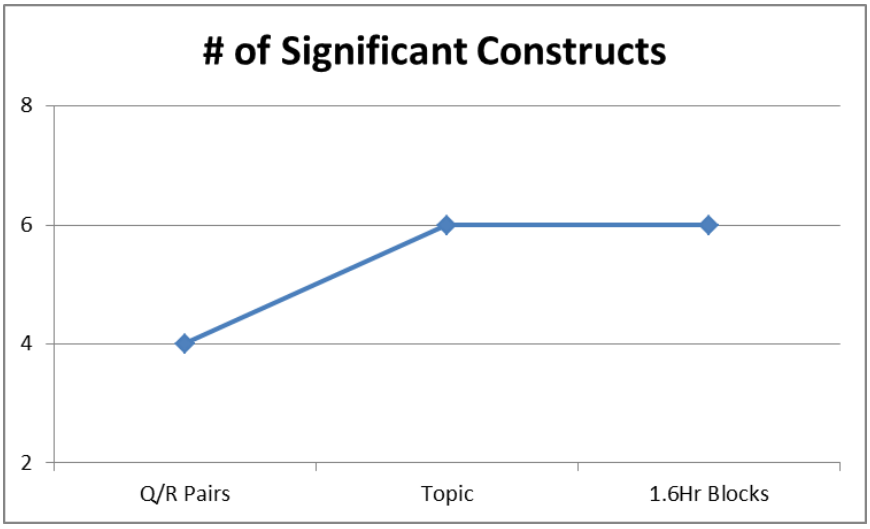


Figure 27, # of Significant Constructs

Note: Upon closer examination proposition P3A *"Speech cues will have the highest correlation with deceptive behavior when epochs are at the topic level"* was similar enough to P1B that the two could be combined.

Summary of Results

The following table is a summary of the propositions and whether they are supported or not by the analysis above. As a point of reference the research question is repeated here:

RQ: *Are speech cues to deceptive behavior moderated over time by receiver suspicion during dyadic interactions in a real-world high-stakes setting?*

Table 26, Proposition Summary

Proposition	Summary
P1	Supported
P1A	Supported
P1B	Opposite of Proposition Supported
P2	Partial Support
P2A	Not Supported
P3	Opposite of Proposition Supported

Based on the above it is reasonable to state that the research question was supported and that speech cues to deceptive behavior are impacted by receiver's suspicion during dyadic interactions in real-world high-stakes settings. This dissertation also look at whether measurements and constructs, developed by previous researchers, could hold up under a RWHS case. The ANOVA of the 41 behavioral cues measured 29.3% as significant at the Q/R pair epoch level, 31.7% at the topic level, and 41.5% at the 1.6 hr block level. Given the poor quality of the audio data, this is strong support for utilizing these measures in future deception detection research. Regression also showed a strong relationship between the levels of suspicion and the individual measures with 39% at the Q/R pair level and 31.7% at both the topic and 1.6 hr levels of granularity as significant. ANOVA and regression showed similar support for the constructs and their relationship to suspicion levels. However, PCA results were more telling, showing five very strong constructs and two which could be reduced to fewer measure. The linguistic construct of *Quantity* and the paralinguistic construct of *Time-Duration* had all their measures load on the same component. These could be combined into a single construct of *Voice-Quantity*. The linguistic construct of *Affect* showed strong loadings for all of its measures as did the paralinguistic constructs *Intensity* and *Time-Tempo*. The construct

Clarity loaded high on only two measures and *Fluency* in only one; both of these could be reduced in dimension.

As with any case study, more questions arose than were answered. The final chapter will discuss contributions limitations, and areas for future research.

CHAPTER 5

DISCUSSION & CONCLUSION

This dissertation had three tenets, derived from perceived gaps in existing deception detection research. Specifically, that existing theories on deception crave validation outside of the lab in a RWHS setting, where typical dyadic interactions are long and more complex than those studied in a controlled laboratory experiment. There are good reasons research in a RWHS setting are rare. First, it is impossible to ethically replicate a RWHS environment and second, it is very difficult to capture the contextual complexities like duration and interactivity of free-flowing dyadic communications.

To focus the study a leading deception detection theory, IDT, was chosen because it most closely models real-world dyadic communication. From this the author set out to explore if speech cues to deceptive behavior were moderated over time by receiver's suspicion during dyadic interactions in a RWHS setting. Fortunately, a wonderful data set presented itself in the form of an adjudicated criminal case study. Because it was an exploratory case study, propositions were tested, not hypotheses. As Yin (2009) pointed out, exploratory studies attempt to answer how and why questions with the hope of better understanding a problem and assist in formulating quantifiable hypotheses to examine later. Though the methods employed here were exploratory in their implementation they did reasonably well in describing the deceptive behavior of the subject.

Proposition P1, *the difference in speech cues to deceptive behavior will be moderated over time by the level of suspicion*, was supported. The Figure 11, Z-Score of Significant Measures by Quartile, showed a strong increase in deceptive behavior over time vs truthful behavior. P1A, *the magnitude of speech cues to deceptive behavior will be positively related to an increase in level of suspicion*, also showed strong differences in magnitude between truthful and deceptive states (see Figure 12 - Figure 15). However, the difference between the three levels of suspicion which include deception showed mixed levels of magnitude. This could be an artifact of the single case study and the small sample size of level 2 of suspicion, those without evidence.

Proposition P1B, *the level of moderation by suspicion will be most detectable when epochs are at the topic level* and P3, *As the granularity of epochs decreases (epochs gets smaller), the number of speech cues to deceptive behavior will increase to an apex then decrease* were supported but in the opposite direction than proposed. Proposition P1B and P3 related to the idea that during a conversation the different types of lies that revolve around a topic (i.e. omission, distortion, and half-truths) are more numerous than blatant lies about a specific question (Gillespie, Hybnerova, Esmark, & Noble, 2014; Kihlstrom, 1995). So the notion that deception would be most detectable and in greater number at the topic level vs any other level was tested. Surprisingly the opposite happened. As granularity *increased* (got larger), the number of significant cues increased, but not to an apex at the topic level. The reason for this is difficult to explain. However, exploring the relationship between cues to deceptive behavior and the different levels of granularity of a conversation will make an interesting future study.

Proposition P2, looked for patterns that differentiate deceptive from truthful behavior. Very clear patterns were found when graphical analysis was done, but this is not enough evidence to identify patterns in new data sets. The changes in behavior over time may aid other practitioners in the development of hypothesis but for this study the author could only concede a partial support.

A difficulty with this dissertation came from the exploratory methodology itself. The tools used to explore and explain what was happening in the data were numerous. It was difficult for the author to choose which tools to use and with each new analysis, new questions arose. This made the focus of the dissertation difficult to hold down. A dissertation should generate four or five future research questions to study. During this dissertation at least a dozen new research questions came up, each one a distraction from the current study.

The length of the interaction was also a good opportunity to examine IDT and how a lengthy dyadic communication can be dissected into reasonable units of analysis. Several measures and constructs, utilized and validated in existing research, were explored and validated in this study. However, many of the measures and their constructs were not significant predictors of deceptive behavior or explained only a fraction of the variance. The reason for their poor predictive power could be explained because the study was a single case and the fact that all measurements were taken from an uncontrolled environment. However, this fact does add weight to those measures and constructs that were significant predictors of deceptive behavior.

Contributions

In regards to IDT, one contribution of this study is a better understanding of the impact suspicion has in a RWHS setting. IDT was validated to the extent that suspicion play a role in sender's behavior and it affected cue intensity. It is apparent that not only does suspicion play a central role in IDT but that its impact on deceptive speech behaviors is measurable in a RWHS environment. This point is important to unlocking future studies involving IDT, suspicion, and RWHS cases.

Another contribution, this time for practitioners, is the creation of procedures for extracting behavioral cues from speech in real-world environments. This study was focused on speech cues that could be easily automated and processed. The procedures followed here were executed in series with human assistance. However, each step was chosen from existing information technology tools which, with a sizable amount of coding, could be combined into a single RWHS deceptive behavior analyzer requiring very little human input.

A contribution that was not realized completely was the identification of temporal speech patterns for deceptive behavior. Similar to a finger print it was hoped that a pattern would emerge that identified deceptive behavior. Though patterns which followed the constructs developed by others were realized, the temporal aspect was only visible in the graphical analysis, not statistically proven. One complication in doing a time series analysis looking for temporal patterns is the non-uniform nature of real-world dyadic interviews (Bar-Joseph, 2004). Though not impossible to overcome, the difficulty is much greater and was left for future study (Kahveci & Singh, 2001).

The examination of granularity of epochs is another contribution to both IDT and deception detection research as a whole. Some constructs and measures performed better at higher levels of granularity (e.g. Time-Tempo) while some were significant across all levels (e.g. Quantity). The behavior of these measures and constructs at different epochs could be the subject of a controlled study to isolate the cause and effect of granularity with more internal validity than an exploratory study.

Limitations

Several limitations are common to any case study. In the current study an emphasis was made to limiting research only to a RWHS environment, this raises a number of questions. Was this a typical high-stakes interview? Mr. Perry was more than a suspect, he knew the FBI had evidence against him, but he did not how much evidence Det. Wall had against him. Before the interview he wanted to plea down to six years for trafficking in child pornography; after the interview he received life in prison, 470 years to be exact. One could argue that having been caught, even on one criminal charge, he did not think he had much to lose by his deception.

The nature and environment of this real-world case is another limitation and potential area for further study. Longer, dyadic communication indicative of law enforcement interviews combined with a lack of fine granularity of episodes suggests the need for further research in interview-style communications. The difficulty is two-fold; longer duration interviews will be more difficult to gather in a controlled manner simply because volunteers are not going to sit for hours without proper compensation. Secondly,

the free-flowing nature of longer communications makes controlling the study more complex.

Was the quality of audio a factor? Audio was limited to a monophonic, relatively low quality recording in a typical law enforcement interview room, not a controlled sound studio. To put it bluntly, the audio quality was very poor so any noticeable results are noteworthy. Could higher fidelity and stereo make a difference? One advantage of stereo recordings where sender and receiver are on different tracks is the ability to utilize additional tools to improve noise reduction and speaker separation. One such tool is dynamic spectral subtraction (DSS), which can be used to extract speech from a single noisy channel (He, Xu, & Zhu, 2008). DSS subtracts an unwanted signal from a voice recording even if the unwanted signal is louder than the voice. In the case where the unwanted signal is on a separate channel but bleeds over to the voice channel DSS can almost completely separate the unwanted signal with minimal loss of fidelity.

Another limitation of this study is its scope and application of IDT. IDT has three distinct phases, this study focused only on the central interaction phase. The incorporation of the pre-interaction phase may impact the interactions in a RWHS setting more than anticipated.

With any N=1 study, specific subject characteristics not being observed or manipulated may have significant impact on results. In traditional studies which include randomization and much larger sample sizes, unobserved variables can be accounted for and confounding factors minimized. For example, in the present case the subject was evaluated by psychologist and deemed not to be a psychopath or suffering from some

other mental condition. Even though an individual is "normal" in the eyes of a psychologist, who is to say the mind of a habitual criminal and their expressed behaviors are similar to non-criminals. For instance, Mr. Perry did not perceive rape as violent of an act as punching someone or threatening their life. The subject was not mentally unstable or psychotic but his value system was so far from society's acceptable norms that to most people they appear unstable. This is a limitation of the current study because the subject was a confirmed criminal with a long history of crimes. The results may not be applicable to non-criminals who face a unique RWHS deception situation (e.g. spouse caught having an affair). A related limitation, studied many times, is the impact of practice on deception. Successful criminals notoriously hone their deceptive skills in a RWHS setting over and over while non-criminals are typically limited to infrequent RWHS deception. With so many severe limitations of real-world data sets, any observable behavior changes which distinguish truth from deception are noteworthy. The effort necessary to gather additional RWHS cases to examine should be pursued to further the knowledge of deception detection in RWHS situations.

While processing the data for analysis several issues and questions arose. First, encoding of distinct topics was difficult. The speech was conversational, so topics were often fluid and changed back and forth or overlapped. If the level of granularity is less concrete than words and sentences which have clear delineations, researchers should decide ahead of time how the epochs will be handled. In the case of topic level granularity, one could apply a linear dissection, considering any change in topic a new instance. The more complex approach would be to regroup the conversation, putting common topics together for analysis. This would be almost impossible to automate not

to mention destroying the integrity of the data timeline. Implementation of an interface program where the investigator could indicate with a press of a button at each new question or topic could mitigate this issue. On a related note, it was found that more than one lie can come from a single question/response pair. If the response was verbose, two or more distinct lies could be made. This made encoding a challenge; does the researcher count each deception separately or as one? In this study if a single response containing multiple deceptions was exceedingly long (which was rare) the deceptions were not split into smaller segments.

Another issue that came up during data encoding was the question; should laughs, crying, and sobbing be kept and if so how are vocal emotions categorized? These sounds are obviously significant emotionally but from a paralinguistic point of view their wide variety may be problematic. Laughing, crying, and sobbing are often much louder than normal speech, which can also create outliers. Though outside the scope of this study, emotional outbursts may hold another key to measuring deception not considered in linguistic or paralinguistic behaviors.

Future Research

The exploratory nature of the study, the volume of data, and the numerous methods of analysis used generated many possibilities for future research. One aspect of IDT which should be examined in greater detail is the view that deception involves strategic and non-strategic behaviors. This study's initial view into a RWHS deceptive case did not look for strategic motives. However, such an examination could produce

new insightful knowledge about deception, specifically in the case of longer more realistic dyadic interactions. This study kept IDT at the forefront when choosing the research question and subsequent propositions. However, as mentioned in the literature review there are several theories on deceptive behavior, all of which could benefit if looked at through a RWHS case study.

The original data from this case is in video format which could be used as another communication channel in which to examine other theories on deception detection. Furthermore, the original data covers three separate sessions, one each day in succession. This volume of data should be examined in its entirety. The levels of granularity could then include a comparison of sessions to see if the subject's behavior changes from day to day. Another level of granularity that needs to be examined is the individual word level. Individual words were considered in this study but several characteristics were not; for example, individual word count comparisons, key words (i.e. highly emotional words, unique words, etc.). One limitation of the current data set which should be considered in a separate study is the mono vs stereo audio. Having all audio on a single track limits the tools available for processing audio signals. With the cooperation of law enforcement agencies, excellent audio data can be gathered if they are willing to put a separate microphone on the investigator and place the suspects microphone as close to them as possible. In addition, if the investigators are willing to pause slightly between subjects and try to stay on topic as much as possible the crossover of Q/R pairs and topic could be minimized or avoided completely.

Another area of future research could lie in the definition of ground truth used in this study. Ground truth was well defined; however, a closer examination of the type of

evidence investigators have and their relationship to deceptive behavior could prove very interesting. Several studies have looked at using evidence to assist with detecting deception but much could be gained if analysis could be run on RWHS cases (Granhag, Strömwall, Willén, & Hartwig, 2013; Hartwig, Granhag, Strömwall, & Kronkvist, 2006; Hartwig, Granhag, Strömwall, & Vrij, 2005).

The final potential future research areas which will be mentioned are the possibility of deception detection automation and the development of a collection of RWHS deception case studies. If a database of RWHS cases in which ground truth is established could be collected, it would be invaluable to the field of deception research. With the advancement of voice recognition programs which can capture linguistic and paralinguistic measures comes a need for processes and procedures to automate deception detection. All the pieces are there, what has not been done is assembling them into one complex deception detection system. The efforts to process the video or audio files in this dissertation could be the basis on which such a database could be populated with behavioral measures and data sets. This would be a huge undertaking, most likely requiring many interdisciplinary researchers working together. Before such collaboration can begin a large dataset of RWHS cases with a common well defined and validated definition of ground truth must be collected. This deception case data warehouse, like this dissertation, will be difficult and painful to collect but well worth the effort in the end.

REFERENCES

Aamodt, M. G., & Custer, H. (2006). Who can best catch a liar?: A meta-analysis of individual differences in detecting deception. *The Forensic Examiner, 15*(1), 6-11.

Aguirre, G. K. (2010). Experimental Design and Data Analysis for fMRI. In S. H. Faro & F. B. Mohamed (Eds.), *BOLD fMRI* (pp. 55-69): Springer New York.

Aguirre, G. K., & D'Esposito, M. (1999). Experimental design for brain fMRI. In C. T. W. Moonen & P. A. Bandettini (Eds.), Functional MRI (pp. 369–380). New York: Springer.

Akehurst, L., Bull, R., Vrij, A., & Koehnken, G. (2004). The Effects of Training Professional Groups and Lay Persons to use Criteria-Based Content Analysis to Detect Deception. Applied Cognitive Psychology, 18(7), 877-891.

Anolli, L., & Ciceri, R. (1997). The voice of deception: Vocal strategies of naive and able liars. *Journal of Nonverbal Behavior, 21*(4), 259-284.

Axinn, W. G., & Pearce, L. D. (2006). *Mixed method data collection strategies*: Cambridge University Press.

Baayen, R. H. (2008). *Analyzing linguistic data* (Vol. 505): Cambridge University Press Cambridge, UK.

Bar-Joseph, Z. (2004). Analyzing time series gene expression data. *Bioinformatics, 20*(16), 2493-2503.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173.

Bazin, A. (2010). Speaker recognition finds its voice. *Biometric Technology Today, 2010*(9), 5-7.

Ben-Shakhar, G., & Elaad, E. (2003). The validity of psychophysiological detection of information with the Guilty Knowledge Test: A meta-analytic review. *Journal of Applied Psychology, 88*(1), 131-151.

Biros, D. P., Hass, M. C., Wiers, K., Twitchell, D., Adkins, M., Burgoon, J. K., Nunamaker, J. F. (2005). Task performance under deceptive conditions: Using military scenarios in deception detection research, Paper presented at the System Sciences, 2005. HICSS '05. Proceedings of the 38th Annual Hawaii International Conference on System Sciences.

Boltz, M. G., Dyer, R. L., & Miller, A. R. (2010). Jo are you lying to me? Temporal cues for deception. *Journal of Language and Social Psychology*, *29*(4), 458-466.

Bond Jr., C. F., & DePaulo, B. M. (2006). Accuracy of Deception Judgments. *Personality & Social Psychology Review (Lawrence Erlbaum Associates), 10*(3), 214-234.

Bond Jr., C. F., & Fahey, W. E. (1987). False suspicion and the misperception of deceit. British Journal of Social Psychology, 26, 41-46.

Bond Jr., C. F., Kahler, K. N., & Paolicelli, L. M. (1985). The miscommunication of deception: An adaptive perspective. *Journal of Experimental Social Psychology*, 21(4), 331-345.

Brandt, D. R., Miller, G. R., & Hocking, J. E. (1980). Effects of self-monitoring and familiarity on deception. Communication Quarterly, 22, 3-10.

Bruneau, T. J. (1980). Chronemics and the Verbal-Nonverbal Interface. In M. R. Key (Ed.), The Relationship of Verbal and Nonverbal Communication. Walter de Gruyter.

Bull, R., Baron, H., Gudjonsson, G., Hampson, S., Rippon, G., & Vrij, A. (2004). A review of the current scientific status and fields of application of polygraphic deception detection. *London: British Psychological Society*.

Buller, D. B., & Aune, R. K. (1987). Nonverbal cues to deception among intimates, friends, and strangers. *Journal of Nonverbal Behavior, 11*(4), 269-290.

Buller, D. B., & Burgoon, J. K. (1994). Deception: Strategic and nonstrategic communication. In J. A. Daly & J. M. Wiemann (Eds.), *Strategic interpersonal communication.* (pp. 191-223): Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

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

Buller, D. B., Comstock, J., Aune, R. K., & Strzyzewski, K. D. (1989). The Effect Of Probing On Deceivers And Truthtellers. *Journal of Nonverbal Behavior, 13*(3), 155-170.

Buller, D. B., & Hunsaker, F. (1995). Interpersonal deception: XIII. Suspicion and the truth-bias of conversational participants. In J. Aitken (Ed.), Intrapersonal communication processes reader (pp. 239-251). Westland, MI: Hayden-McNeil.

Buller, D. B., Strzyzewski, K. D., & Comstock, J. (1991). Interpersonal deception: I. deceivers' reactions to receivers' suspicions and probing. Communication Monographs, 58, 1-24.

Buller, D. B., Strzyzewski, K. D., & Hunsaker, F. G. (1991). Interpersonal deception: 11. The inferiority of conversational participants as deception detectors. Communication Monographs, 58, 25-40.

Burgoon, J. K., Buller, D. B., Dillman, L., & Walther, J. B. (1995). Interpersonal deception IV. Effects of suspicion on perceived communication and nonverbal behavior dynamics. Human Communication Research, 22(2), 163-196.

Burgoon, J. K., Buller, D. B., & Woodall, W. G. (1996). *Nonverbal communication: The unspoken dialogue*: McGraw-Hill New York, NY.

Burgoon, J. K., Nunamaker, J. F., George, J. F., & Biros, D. P. (2007). *Detecting Deception in the Military Infosphere: Improving and Integrating Human Detection Capabilities with Automated Tools*: DTIC Document.

Burgoon, J. K., & Qin, T. (2006). The Dynamic Nature of Deceptive Verbal Communication. Journal of Language and Social Psychology, 25(1), 76-96.

Camden, C., Motley, M. M., & Wilson, A. (1984). White lies in interpersonal communication: A taxonomy and preliminary investigation of social motivations. The Western Journal of Speech Communication, 48, 309-325.

Campbell, R. (2008). The processing of audio-visual speech: empirical and neural bases. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*(1493), 1001-1010.

Cefaratti, M., & Barkhi, R. (2013). The Effects of Communication Media and Client Familiarity on Auditors' Confidence in Deception Detection. *Journal of Forensic & Investigative Accounting, 5*(2).

Cestaro, Victor L. (1996). A Comparison of Accuracy Rates Between Detection of Deception Examinations Using the Polygraph and the Computer Voice Stress Analyzer in a Mock Crime Scenario. Ft. Belvoir: Defense Technical Information Center. http://handle.dtic.mil/100.2/ADA318985.

Clark, H. H., & Clark, E. V. (1977). Psychology and language: an introduction to psycholinguistics. New York: Harcourt Brace Jovanovich.

Cody, M. J., Marston, P. J., & Foster, M. (1984, May). Paralinguistic and verbal leakage of deception as a function of attempted control and timing of questions. Paper presented at the annual meeting of the International Communication Association, San Francisco.

Cody, M. J. & O'Hair, H. D. (1983). Nonverbal communication and deception cues to gender and communicator dominance. Communication Monographs, 50, 175-192.

Coltman, T., Devinney, T. M., Midgley, D. F., & Venaik, S. (2008). Formative versus reflective measurement models: Two applications of formative measurement. *Journal of Business Research, 61*(12), 1250-1262.

Comas, P., Turmo, J., & Surdeanu, M. (2008). Robust Question Answering for Speech Transcripts Using Minimal Syntactic. *Analysis Advances in Multilingual and Multimodal Information Retrieval*. Vol. 5152, pp. 424-432.

Conradie, W., de Wet, T., & Jankowitz, M. (2009). Performance of nonlinear smoothers in signal recovery. *Applied Stochastic Models in Business and Industry*, 25(4), 425-444.

Cook, T. D., & Campbell, D. T. (1979). Quasi-experimentation design & analysis issues for field settings. Chicago: Rand McNally.

Daft, R., L., & Lengel, R., H. (1986). Organization Information Requirements, Media Richness and Structural Design. Management Science, 32(5), 554-571.

Damphousse, K.R., Pointon, L., Upchurch, D., and Moore, R.K. "Assessing the validity of Voice Stress Analysis tools in a jail setting," *Report submitted to the US Department of Justice: http://www. ncjrs. gov/pdffiles1/nij/grants/219031. pdf*) 2007.

Davatzikos, C., Ruparel, K., Fan, Y., Shen, D., Acharyya, M., Loughead, J., Gur, R., and Langleben, D.D. "Classifying spatial patterns of brain activity with machine learning methods: application to lie detection," *Neuroimage* (28:3) 2005, pp 663-668.

Davis, M., Markus, K., Walters, S., Vorus, N., & Connors, B. (2005). Behavioral Cues to Deception vs. Topic Incriminating Potential in Criminal Confessions. *Law and Human Behavior, 29*(6), 683-704. doi: 10.1007/s10979-005-7370-z

Dennis, A. R., Fuller, R. M., & Valacich, J. S. (2008). Media, Tasks, and Communication Processes: A Theory of Media Synchronicity. Management Information Systems Quarterly, 32(3), 575-600.

Dennis, A. R., & 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), 256-274.

Dennis, A. R., & Valacich, J. S. (1999). Rethinking media richness: towards a theory of media synchronicity. Paper presented at the System Sciences, 1999. HICSS '99. Proceedings of the 32nd Annual Hawaii International Conference on System Sciences.

Department of Homeland Security, Quadrennial Homeland Security Review Report: A Strategic Framework for a Secure Homeland (Feb 2010). Retrieved from http://www.dhs.gov/xlibrary/assets/qhsr_report.pdf

DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin, 111*(2), 203.

DePaulo, B. M, & Kashy, D. A. (1998). Everyday lies in close and casual relationships. Journal of Personality and Social Psychology, 74, 63-79.

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), 979-995.

DePaulo, B. M., Kirkendol, S. E., Tang, J., & O'Brien, T. P. (1988).The motivational impairment effect in the communication of deception: Replications and extensions. Journal of Nonverbal Behavior, 12, 177-202.

DePaulo, B. M., Lanier, P. S., & Davis, T. (1983). Detecting the deceit of the motivated liar. Journal of Personality and Social Psychology, 45, 1096-1103.

DePaulo, B. M., Lassiter, G. D., & Stone, J. I. (1982). Attitudinal determinants of success at detecting deception and truth. Personality and Social Psychology Bulletin, 8, 273-279.

DePaulo, B.M., LeMay, C.S., & Epstein, J.A. (1991). Effects of importance of success and expectations for success on effectiveness at deceiving. Personality and Social Psychology Bulletin, 17, 14-24.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. Psychological Bulletin, 129(1), 74-118.

DePaulo, B. M., Stone, J. I., & Lassiter, G. D. (1985). Deceiving and detecting deceit. In B. R. Schlenker (Ed.), *The self and social life* (pp. 323-370). New York: McGraw-Hill.

deTurck, M. A., & Miller, G. R. (1985). Deception and arousal: Isolating the behavioral correlates of deception. *Human Communication Research*.

deTurck, M. A., & Miller, G. R. (1990). Training Observers to Detect Deception: Effects of Self-Monitoring and Rehearsal. *Human Communication Research, 16*(4), 603-620.

Díaz, I., Domínguez, M., Cuadrado, A. A., & Fuertes, J. J. (2008). A new approach to exploratory analysis of system dynamics using SOM. Applications to industrial processes. *Expert Systems with Applications*, 34(4), 2953-2965.

Dunbar, N. E., Jensen, M. L., Bessarabova, E., Burgoon, J. K., Bernard, D. R., Harrison, K. J., . . . Eckstein, J. M. (2012). Empowered by Persuasive Deception:: The Effects of Power and Deception on Dominance, Credibility, and Decision Making. *Communication Research*.

Duran, N. D., Dale, R., Kello, C. T., Street, C. N., & Richardson, D. C. (2013). Exploring the movement dynamics of deception. *Frontiers in psychology, 4*.

D'Mello, S. K., & Graesser, A. (2010). Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features. User Modeling and User-Adapted Interaction, 20(2), 147-187.

Ekman, P. (1985). Telling lies: Clues to deceit in the marketplace, politics, and marriage. New York: WW Norton & Company.

Ekman, P. (1992). Facial Expressions of Emotion: New Findings, New Questions. Psychological Science. Wiley-Blackwell, 3(1), 34-38.

Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. Psychiatry: Journal for the Study of Interpersonal Processes, 32, 88-105.

Ekman, P., & Friesen, W. V. (1974). Detecting deception from the body or face. Journal of Personality and Social Psychology, 29(3), 288-298.

Elkins, A. C., Burgoon, J., & Nunamaker, J. (2012). Vocal Analysis Software for Security Screening: Validity and Deception Detection Potential. *Homeland Security Affairs, 8*.

Fangfang, Z. (2008, 20-22 Dec. 2008). *The Application of Visualization Technology on E-Commerce Data Mining*. Paper presented at the Intelligent Information Technology Application, 2008. IITA '08. Second International Symposium on.

Fay, P. J., & Middleton, W. C. (1941). The ability to judge truth-telling, or lying, from the voice as transmitted over a public address system. *Journal of General Psychology*.

Fayyad, U. M., Wierse, A., & Grinstein, G. G. (2002). *Information visualization in data mining and knowledge discovery*: Morgan Kaufmann.

Feeley, T. H., & deTurck, M. A. (1998). The Behavioral Correlates of Sanctioned and Unsanctioned Deceptive Communication. Journal of Nonverbal Behavior, 22(3), 189-204.

Feeley, T.H., & Young, M.J. (1998). Humans as lie detectors: Some more second thoughts. Communication Quarterly, 46 (2), 109-126.

Fetterman, D. M. (1998). Ethnography: Step by step. Newbury Park, CA: Sage.

Ford, E. B. (2006). Lie detection: Historical, neuropsychiatric and legal dimensions. International Journal of Law and Psychiatry, 29(3), 159-177.

Forrest, J.A., & Feldman, R.S. (2000). Detecting deception and judge's involvement: Lower task involvement leads to better lie detection. Personality and Social Psychology Bulletin, 26 (1), 118-125.

Frank, M.G., & Ekman, P. (1997). The ability to detect deceit generalizes across different types of high-stake lies. Journal of Personality and Social Psychology, 72 (6), 1429-1439.

Frank, M. G., & Feeley, T. H. (2003). To catch a liar: Challenges for research in lie detection training. *Journal of Applied Communication Research, 31*(1), 58-75.

Fuller, C.M. (2008). "High-stakes, real-world deception: An examination of the process of deception and deception detection using linguistic-based cues," *PhD Dissertation, Oklahoma State University*), Retrieved 13 Jun 2010, from Dissertations & Theses: A&I database.

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.

Fuller, C. M., Biros, D. P., & Delen, D. (2011). An investigation of data and text mining methods for real world deception detection. *Expert Systems with Applications, 38*(7), 8392-8398.

Fuller, C. M., Biros, D. P., & Wilson, R. L. (2009). Decision support for determining veracity via linguistic-based cues. *Decision Support Systems, 46*(3), 695-703.

Fuller, C. M., Marett, K., & Twitchell, D. P. (2012). An examination of deception in virtual teams: Effects of deception on task performance, mutuality, and trust. Professional Communication, IEEE Transactions on, 55(1), 20-35.

Gamer, M., Rill, H.-G., Vossel, G., & Godert, H. W. (2006). Psychophysiological and vocal measures in the detection of guilty knowledge. International Journal of Psychophysiology, 60(1), 76-87.

Ganis, G., Rosenfeld, J. P., Meixner, J., Kievit, R. A., & Schendan, H. E. (2011). Lying in the scanner: Covert countermeasures disrupt deception detection by functional magnetic resonance imaging. *Neuroimage, 55*(1), 312-319.

Ganis, G., Kosslyn, S., Stose, S., Thompson, W., & Yurgelun-Todd, D. (2003). Neural correlates of different types of deception: an fMRI investigation. Cerebral Cortex (New York, N.Y.: 1991), 13(8), 830-836.

Garson, G. D., 2011. *StatNotes: Topics in Multivariate Analysis*. Online at http://faculty.chass.ncsu.edu/garson/PA765/logistic.htm, accessed 18 Apr 11.

Geizer, R. S., Rarick, D. L., & Soldow, G. F. (1977). Deception judgment accuracy: A study of person perception. Personality and Psychology Bulletin, 3, 446-449.

Gillespie, E. A., Hybnerova, K., Esmark, C., & Noble, S. M. (2014). A tangled web: views of deception from the customer's perspective. *Business Ethics: A European Review*.

Goffman, E. (1974). Frame analysis: An essay on the organization of experience. New York: Harper & Row.

Granhag, P. A., Strömwall, L. A., Willén, R. M., & Hartwig, M. (2013). Eliciting cues to deception by tactical disclosure of evidence: The first test of the Evidence Framing Matrix. *Legal and Criminological Psychology*, *18*(2), 341-355.

Green D. M., Swets, J. A. (1966). Signal Detection Theory and Psychophysics: New York: Wiley.

Grice, H. P. (1989) Studies in the way of words, Harvard University Press, Cambridge, Mass.

Haddad, D. & Ratley, R. (2002). Investigation and evaluation of voice stress analysis technology. Technical report, National Criminal Justice Reference Service.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). Multivariate data analysis (5 ed.). Upper Saddle River, New Jersey: Prentice-Hall, Inc.

Hample, D. (1980). Purposes and effects of lying. The Southern Speech Communication Journal, 46, 33-47.

Hancock, J. T. (2007). Digital deception. *Oxford handbook of internet psychology*, 289-301.

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

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Kronkvist, O. (2006). Strategic use of evidence during police interviews: when training to detect deception works. *Law and Human Behavior*, *30*(5), 603.

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Vrij, A. (2005). Detecting deception via strategic disclosure of evidence. *Law and Human Behavior, 29*(4), 469.

He, L., Xu, B., & Zhu, Y. (2008). Method for spectral subtraction in speech enhancement: Google Patents.

Herndon, S., & Kreps, G. (2001). Qualitative research: Applications in organizational life. Cresskill: NJ: Hampton Press.

Hocking, J. E., Bauchner, J., Kaminski, E. P., & Miller, G. R. (1979). Detecting deceptive communication from verbal, visual, and paralinguistic cues. Human Communication Research, 6, 33-46.

Hollien, J. H. (2006). Voice stress analyzer instrumentation evaluation. Technical report, Counterintelligence Field Activity.

Hollien, H., & Harnsberger, J. D. (2006). Voice stress analyzer instrumentation evaluation (No. Final Report, CIFA Contract FA 4814-04-0011): IASCP, University of Florida.

Hopkins, C. S., Benincasa, D. S., Ratley, R. J., & Grieco, J. J. (2005). Evaluation of Voice Stress Analysis Technology. Paper presented at the System Sciences, 2005. HICSS '05. Proceedings of the 38th Annual Hawaii International Conference on.

Howard, D., & Kirchhübel, C. (2011). Acoustic Correlates of Deceptive Speech–An Exploratory Study. *Engineering Psychology and Cognitive Ergonomics*, 28-37.

Hu, X., Chen, H., & Fu, G. (2012). A repeated lie becomes a truth? The effect of intentional control and training on deception. *Frontiers in psychology, 3*.

Hubbell, A.P., Mitchell, M.M., & Gee, J.C. (2001). The relative effects of timing of suspicion and outcome involvement on biased message processing. Communication Monographs, 68 (2), 115-132.

Huettel, S. A., Song, A. W., & McCarthy, G. (2004). *Functional magnetic resonance imaging* (Vol. 1): Sinauer Associates Sunderland, MA.

Iacono, W. G. (1991). Can we determine the accuracy of polygraph tests. *Advances in psychophysiology*, *4*, 201-207.

Iacono, W.G. "Forensic lie detection: Procedures without scientific basis," *Journal of forensic psychology practice* (1:1) 2001, pp 75-86.

Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2001). *Criminal interrogation and confessions* (4th ed.). Gaithersburg, Maryland: Aspen Publishers.

Institute for Economics & Peace (2012). Global Terrorism Index, Capturing the Impact of Terrorism for the Last Decade. Retrieved from http://www.visionofhumanity.org/sites/default/files/2012_Global_Terrorism_Index_Report.pdf

Janniro, M., & Cestaro, V. (1996). Effectiveness of Detection of Deception Examinations Using the Computer Voice Stress Analyzer. Ft. McClellan, AL: Department of Defense Polygraph Institute. http://handle.dtic.mil/100.2/ADA318986

Jensen, M. L., Burgoon, J. K., & Jay F. Nunamaker, J. (2010). Judging the Credibility of Information Gathered from Face-to-Face Interactions. *J. Data and Information Quality*, *2*(1), 1-20.

Jensen, M. L., Lowry, P. B., & Jenkins, J. L. (2011). Effects of Automated and Participative Decision Support in Computer-Aided Credibility Assessment. *Journal of Management Information Systems*, 28(1), 201-234.

Jewitt, C., & Leeuwen, T. V. (2010). A Handbook of Visual Analysis. *Applied Linguistics*.

Johnson, R., Barnhardt, J., & Zhu, J. (2004). The contribution of executive processes to deceptive responding. *Neuropsychologia*, 42(7), 878-901.

Jorgensen, D. L. (1989). Participant observation: a methodology for human studies. Newbury Park, CA.: Sage.

Kahveci, T., & Singh, A. (2001). *Variable length queries for time series data. Paper presented at the Data Engineering*, 2001. Proceedings. 17th International Conference on.

Kalbfleisch, P. J. (1992). Deceit, distrust and the social milieu: Application of deception research in a troubled world. Journal of Applied Communication Research, 20(3), 308-334.

Kassin, S. M., & Fong, C. T. (1999). "I'm Innocent!": Effects of Training on Judgments of Truth and Deception in the Interrogation Room. Law and Human Behavior, 23(5), 499-516.

Kassin, S., Leo, R., Meissner, C., Richman, K., Colwell, L., Leach, A.-M., and Fon, D. (2007). "Police Interviewing and Interrogation: A Self-Report Survey of Police Practices and Beliefs," *Law and Human Behavior* (31:4), pp 381-400.

Kihlstrom, J. F. (1995). *From the subject's point of view: The experiment as conversation and collaboration between investigator and subject.* Paper presented at the Keynote address at the 7th annual convention of the American Psychological Society, New York, NY.

Ko, D. G., & Dennis, A. R. (2011). Profiting from Knowledge Management: The Impact of Time and Experience. *Information Systems Research*, 22(1), 134-152.

Kock, N. (2004). The Psychobiological Model: Towards a New Theory of Computer-Mediated Communication Based on Darwinian Evolution. Organization Science, 15(3), 327-348.

Koper, R. J., & Sahlman, J. M. (1991). *The Behavioral Correlates of Real-World Deceptive Communication*. Paper presented at the Annual Meeting of the International Communication Association, Chicago, IL

Kraut, R. E. (1978). Verbal and nonverbal cues in the perception of lying. Journal of Personality and Social Psychology, 36, 380-391.

Kraut, R. E. (1980). Humans as lie detectors: Some second thoughts. Journal of Communication, 30, 209-216.

Kulaney, E., (1982). Changes in language behavior as a function of veracity. *Human Communication Research*, 9(1):75–82.

Kundzewicz, Z. W., & Robson, A. (2000). Detecting trend and other changes in hydrological data. *World Climate Programme–Water, World Climate Programme Data and Monitoring, WCDMP-45, WMO/TD.*

Lane, J. D., & DePaulo, B.M. (1999). Completing Coyne's cycle: Dysphoric's ability to detect deception. Journal of Research in Personality, 33, 311-329.

Langleben, D. D. (2008). Detection of deception with fMRI: Are we there yet? *Legal and Criminological Psychology*, *13*(1), 1-9.

Leo, R. A., (2008). *Police interrogation and American justic*e. Harvard University Press.

Levine, T.R., Feeley, T.H., McCornack, S.A., Hughes, M., and Harms, C.M. "Testing the effects of nonverbal behavior training on accuracy in deception detection with the inclusion of a bogus training control group," Western Journal of Communication (69:3) 2005, p 203.

Levine, T.R., Park, H.S., & McCornack, S.A. (1999). Accuracy in detecting truths and lies: Documenting the "veracity effect." Communication Monographs, 66 (2), 125-144.

Lippard, P. V. (1988). "Ask me no questions, I'll tell you no lies": Situational exigencies for interpersonal deception. Western Journal of Speech Communication, 52, 91-103.

Littlepage, G. E., & Pineault, M. A. (1985). Detection of deception of planned and spontaneous communications. *The Journal of Social Psychology*, *125*(2), 195-201.

Littlepage, G. E. & Pineault, M. A. (1979). Detection of deceptive factual statements from the body and the face. Personality and Psychology Bulletin, 5, 325-328.

MacLean, L. M., Meyer, M., & Estable, A. (2004). Improving accuracy of transcripts in qualitative research. *Qualitative Health Research, 14*(1), 113-123.

Malhotra, A., Majchrzak, A., Carman, R., & Lott, V. (2001). Radical innovation without collocation: A case study at Boeing-Rocketdyne. *MIS Quarterly*, 229-249.

Mann, S., Vrij, A., & Bull, R. (2002). Suspects, lies, and videotape: An analysis of authentic high-stake liars. *Law and Human Behavior, 26*(3), 365-376.

Mann, S. Vrij A., & Bull, R. (2004). Detecting true lies: Police officers' ability to detect deceit. Journal of Applied Psychology, 89(1), 137-149.

Masip, J., Garrido, E., & Herrero, C. (2009). Heuristic Versus Systematic Processing of Information in Detecting Deception: Questioning The Truth Bias. *Psychological Reports, 105*(1), 11-36.

Mattson, M. (1994, November). Reactive strategies to suspicious information in intimate relationships. Paper presented to the annual meeting of the Speech Communication Association, New Orleans.

McBurney, D., & White, T. (2009). *Research methods*: Cengage Learning.

McCornack, S. A. (1992). Information Manipulation Theory. Communication Monographs, 59, 1-16.

McCornack, S.A. (1997). The generation of deceptive messages: Laying the groundwork for a viable theory of interpersonal deception. In J.O. Greene (Ed.), Message

production: Advances of communication theory (pp. 91-126). Mahwah, NJ: Lawrence Erlbaum.

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

McCornack, S. A., Levine, T. R., Solowczuk, K., Torres, H. I., & Campbell, D. M. (1992). When the alteration of information is viewed as deception: An empirical test of information manipulation theory. Communication Monographs, 59, 17-29.

McCornack, S. A., & Parks, M. R. (1986). Deception detection and relational development: The other side of trust. In M. L. McLaughlin (Ed.), Communication yearbook 9 (pp. 377-389). Beverly Hills, CA: Sage.

Mehrabian, A. (1971). Nonverbal betrayal of feeling. *Journal of Experimental Research in Personality*.

Meissner, C. A., & Kassin, S. M. (2002). “He's guilty!”: Investigator Bias in Judgments of Truth and Deception. Law and Human Behavior, 26(5), 469-480.

Merriam, S. B. (1998). *Qualitative Research and Case Study Applications in Education. Revised and Expanded from" Case Study Research in Education."*: Jossey-Bass Publishers.

Meservy, T.O. (2007). "Augmenting human intellect: Automatic recognition of nonverbal behavior with application in deception detection," *Ph.D. dissertation, The University of Arizona*), Retrieved March 10, 2009, from Dissertations & Theses: A&I database.

Miller, G. R., & Stiff, J. B. (1993). Deceptive communication. Newbury Park, CA: Sage.

Monmonier, M. (1990). Strategies for the visualization of geographic time-series data. *Cartographica: The International Journal for Geographic Information and Geovisualization, 27(1),* 30-45.(Comas, Turmo, & Surdeanu, 2008)

Murphy, K. R., Myors, B., & Wolach, A. H. (2009). *Statistical power analysis: A simple and general model for traditional and modern hypothesis tests*: Routledge.

Ogawa, S., Lee, T.M., Nayak, A.S., and Glynn, P. (1990). "Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields". *Magnetic Resonance in Medicine* **14**: 68–78

Olson, E. A. (2013). "You don't expect me to believe that, do you?" Expectations influence recall and belief of alibi information. *Journal of Applied Social Psychology, 43*(6), 1238-1247. doi: 10.1111/jasp.12086

O'Hair, D., Cody, M. J., & McLaughlin, M. L. (1981). Prepared lies, spontaneous lies, Machiavellianism, and nonverbal communication. Human Communication Research, 7, 325-339.

O'Sullivan, M., Ekman, P., & Friesen, W. V. (1988). The effect of comparisons on detecting deceit. Journal of Nonverbal Behavior, 12(3), 203-215.

Ott, M., Choi, Y., Cardie, C., & Hancock, J. T. (2011). Finding deceptive opinion spam by any stretch of the imagination. Paper presented at the Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1.

Paralinguistic, (2011) http://dictionary.reference.com/browse/paralinguistic. Retrieved on 11 Aug 2011.

Petter, S., Straub, D., & Rai, A. (2007). Specifying formative constructs in information systems research. *Mis Quarterly*, 623-656.

Petty, R.E., & Cacioppo, J.R. (1986). Communication and persuasion: Central and peripheral routes to attitude change. New York: Springer-Verlag.

Phillips, M. C., Meek, S. W., & Vendemia, J. (2011). Understanding the underlying structure of deceptive behaviors. *Personality and Individual Differences*, *50*(6), 783-789.

Podlesny, J. A., & Raskin, D. C. (1977). Physiological measures and the detection of deception. *Psychological Bulletin*, *84*(4), 782-799.

Pollina, D. A., Dollins, A. B., Senter, S. M., Krapohl, D. J., & Ryan, A. H. (2004). Comparison of Polygraph Data Obtained From Individuals Involved in Mock Crimes and Actual Criminal Investigations. *Journal of Applied Psychology*, *89*(6), 1099.

Porter, S., & Brinke, L. (2010). The truth about lies: What works in detecting high stakes deception? *Legal and Criminological Psychology*, *15*(1), 57-75.

Pudota, N., Dattolo, A., Baruzzo, A., Ferrara, F., & Tasso, C. (2010). Automatic keyphrase extraction and ontology mining for content-based tag recommendation. *International Journal of Intelligent Systems*, 25(12), 1158-1186.

Raiman, N., Hung, H., & Englebienne, G. (2011). *Move, and i will tell you who you are: detecting deceptive roles in low-quality data.* Paper presented at the Proceedings of the 13th international conference on multimodal interfaces.

Ramdharry, G. M., Thornhill, A., Mein, G., Reilly, M. M., & Marsden, J. F. (2012). Exploring the experience of fatigue in people with Charcot–Marie–Tooth disease. *Neuromuscular Disorders*, *22*, S208-S213.

Ravishankar, M., Pan, S. L., & Leidner, D. E. (2011). Examining the Strategic Alignment and Implementation Success of a KMS: A Subculture-Based Multilevel Analysis. *Information Systems Research*, *22*(1), 39-59.

Rice, B. (1978). The new truth machines. Psychology Today, 12(1).

Riggio, R. E. & Friedman, H. S. (1983). Individual differences and cues to deception. Journal of Personality and Social Psychology, 45, 899-915.

Riggio, R. E., Tucker, J., & Throckmorton, B. (1987). Social Skills and Deception Ability. Personality and Social Psychology Bulletin, 13(4), 568-577.

Ringle, C.M., Wende, S., &Will, S. (2005). SmartPLS 2.0 (M3) Beta, Hamburg, http://www.smartpls.de.

Rockwell, P., Buller, D. B., & Burgoon, J. K. (1997). The Voice of Deceit: Refining and Expanding Vocal Cues to Deception. *Communication Research Reports*, *14*(4), 451-459.

Saunders, M., Lewis, P., & Thornhill, A. (2007). Research Methods for Business Students. Harlow, England: Pearson Education Limited. 624 pp: ISBN 0-273-70148-7.

Schacter, D. L., & Buckner, R. L. (1998). Priming and the brain. *Neuron*, 20(2), 185-195.

Schimmel, S. M., Atlas, L. E., & Nie, K. (2007). *Feasibility of single channel speaker separation based on modulation frequency analysis.* Paper presented at the Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on.

Seo, J., & Shneiderman, B. (2005). A rank-by-feature framework for interactive exploration of multidimensional data. *Information Visualization*, 4(2), 96.

Shlens, J. (2005). A tutorial on principal component analysis. *Systems Neurobiology Laboratory, University of California at San Diego.*

Shumway, R. H., & Stoffer, D. S. (2011). *Time Series Regression and Exploratory Data Analysis Time Series Analysis and Its Applications* (pp. 47-82): Springer New York.

Sporer, S. L., & Schwandt, B. (2006). Paraverbal indicators of deception: a meta-analytic synthesis. *Applied Cognitive Psychology*, *20*(4), 421-446.

Stanislaw, H., Todorov, N., "Calculation of signal detection theory measures," Behavior Research Methods, Instruments, & Computers, vol. 31, pp. 137-149, 1999.

Stiff, J. G., Kim, H. J., & Ramesh, C. (1992). Truth biases and aroused suspicion in relational deception. Communication Research, 19, 326-345.

Stiff, J.B., Miller, G.R., Sleight, C., Mongeau, P., Garlick, R., & Rogan, R. (1989). Explanations for visual cue primacy in judgments of honesty and deceit. Journal of Personality and Social Psychology, 56, 555-564.

Sullivan, T.P. "Recording Federal Custodial Interviews," American Criminal Law Review (45) 2008, pp 1297-1439.

Sullivan, T.P. "The Time Has Come for Law Enforcement Recordings of Custodial Interviews, Start to Finish," Golden Gate University Law Review (37:1) 2010, p 7.

Taylor, Wayne (2000a), Change-Point Analyzer 2.0 shareware program, Taylor Enterprises, Libertyville, Illinois. Web: http://www.variation.com/cpa

Taylor, Wayne (2000b), "A Pattern Test for Distinguishing Between Autoregressive and Mean-Shift Data," submitted to *Journal of Quality Technologies*.

ten Brinke, L. M. (2012). *Darwin the detective: behavioural consequences of high-stakes emotional deception.* University of British Columbia.

tenBrinke, L., & Porter, S. (2012). Cry me a river: Identifying the behavioral consequences of extremely high-stakes interpersonal deception. *Law and Human Behavior, 36*(6), 469.

Tesch, R. (1990). *Qualitative research: Analysis types and software tools* (Vol. 337): Routledge.

Thackray, R. I., & Orne, M. T. (1968). Effects of stimulus employed and the level of subject awareness on the detection of deception. Journal of Applied Psychology, 52, 234-239.

Thomas, G. (2010). Doing case study: Abduction not induction, phronesis not theory. *Qualitative Inquiry, 16*(7), 575-582.

Toma, C. L., Hancock, J. T., & Ellison, N. B. (2008). Separating fact from fiction: An examination of deceptive self-presentation in online dating profiles. *Personality and Social Psychology Bulletin, 34*(8), 1023.

Tower, D. L., Jensen, M. L., Dunbar, N. E., & Elkins, A. C. (2013). Don't Lie to Me: The Impact of Deception on Vocalic and Linguistic Synchrony. Paper presented at the Intelligence and Security Informatics Conference (EISIC), 2013 European.

Transcribing-interview, n.d. In Nuance. Retrieved 22 March 2014, from http://www.nuance.com/for-individuals/by-industry/education-solutions/transcribing-interview/index.htm

Trochim, W. (2000). *The Research Methods Knowledge Base* (2nd Edition ed.). Cincinnati, OH.: Atomic Dog Publishing.

Tukey, J. W. (1980). We need both exploratory and confirmatory. *American Statistician*, 23-25.

Tukey, J. W. (1977). Exploratory data analysis. *Reading*, MA.

Turner, R. E., Edgley, C., & Olmstead, G. (1975). Information control in conversations: Honesty is not always the best policy. Kansas Journal of Sociology, 11, 69-89.

Utz, S. (2005). Types of deception and underlying motivation what people think. *Social Science Computer Review*, *23*(1), 49-56.

Voice Writer, n.d. In Wikipedia. Retrieved 22 March 2014, from http://en.wikipedia.org/wiki/Voice_writer.

Vrij, A. (2000). Detecting lies and deceit: The psychology of lying and the implications for professional practice. Chichester; New York: John Wiley.

Vrij, A. (2005). Criteria-Based Content Analysis: A Qualitative Review of the First 37 Studies. Psychology, Public Policy, and Law, 11(1), 3-41.

Vrij, A., Edward, K., & Bull, R. (2001). Police officers' ability to detect deceit: The benefit of indirect deception detection measures. Legal & Criminological Psychology, 6(2), 185.

Vrij, A., Edward, K., Roberts, K. P., & Bull, R. (2000). "Detecting deceit via analysis of verbal and nonverbal behavior." Journal of Nonverbal Behavior 24(4): 239-263.

Vrij, A., Fisher, R., Mann, S., & Leal, S. (2006). Detecting deception by manipulating cognitive load. *Trends in Cognitive Sciences, 10*(4), 141-142.

Vrij, A., Mann, S., Leal, S., & Fisher, R. (2010). 'Look into my eyes': can an instruction to maintain eye contact facilitate lie detection? *Psychology, Crime & Law, 16*(4), 327-348.

Vrij, A., & Mann, S. (2001a). Who killed my relative? police officers' ability to detect real-life high-stake lies. Psychology, Crime & Law, 7(1), 119 - 132.

Vrij, A., & Mann, S. (2001b). Telling and detecting lies in a high-stake situation: The case of a convicted murderer. *Applied Cognitive Psychology, 15*(2), 187-203.

Vrij, A., Mann, S., & Fisher, R. (2006). An Empirical Test of the Behaviour Analysis Interview. *Law and Human Behavior, 30*(3), 329-345.

Ward, M. O., & Guo, Z. (2011). Visual Exploration of Time-Series Data with Shape Space Projections. *Computer Graphics Forum*, 30(3), 701-710.

Whitty, M. T., & Carville, S. E. (2008). Would I lie to you? Self-serving lies and other-oriented lies told across different media. *Computers in Human Behavior, 24*(3), 1021-1031.

Wilkinson, L., & Friendly, M. (2009). The history of the cluster heat map. *The American Statistician*, 63(2), 179-184.

Williams, J. R. (1998). Guidelines for the use of multimedia in instruction. Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Witten, I. H., & Frank, E. (2000). Data mining: Practical machine learning tools and techniques with java. San Francisco: Morgan Kaufman.

Wong, L. (2014). Essential study skills: Cengage Learning.

Yamanishi, K., & Takeuchi, J.-i. (2002). *A unifying framework for detecting outliers and change points from non-stationary time series data*. Paper presented at the Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining.

Yin, R. K. (2009). Case study research : design and methods. Los Angeles, Calif.: Sage Publications.

Zhou, L. (2005). An empirical investigation of deception behavior in instant messaging. *Professional Communication, IEEE Transactions on, 48*(2), 147-160.

Zhou, L., Burgoon, J. K., Nunamaker, J. F., & Twitchell, D. (2004a). Automating Linguistics-Based Cues for Detecting Deception in Text-based Asynchronous Computer- Mediated Communication. Group Decision and Negotiation, 13(1), 81-106.

Zhou, L., Burgoon, J., & Twitchell, D. (2003). A Longitudinal Analysis of Language Behavior of Deception in E-mail. In Intelligence and Security Informatics (pp. 959-959).

Zhou, L., Burgoon, J. K., Twitchell, D. P., Tiantian, Q., & Nunamaker J. F. (2004b). A Comparison of Classification Methods for Predicting Deception in Computer-Mediated Communication. Journal of Management Information Systems, 20(4), 139-165.

Zhou, L., Twitchell, D. P., Tiantian, Q., Burgoon, J. K., & Nunamaker, J. F. (2003). An exploratory study into deception detection in text-based computer-mediated communication. Paper presented at the System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on.

Zuckerman, M., & Driver, R. E. (1985). Telling lies: Verbal and nonverbal correlates of deception. In A. W. Siegman and S. Feldstein (Eds.) Multichannel integrations of nonverbal behavior (pp. 129-147). Hillsdale, NJ: Erlbaum.

Zuckerman, M., & Driver, R. E. (1987). Telling Lies: Verbal and nonverbal correlates of deception. *Multichannel Integration of Nonverbal Behavior*. Hillsdale, NJ: Erlbaum.

Zuckerman, M., Driver, R., & Guadagno, N. S. (1985). Effects of segmentation patterns on the perception of deception. *Journal of Nonverbal Behavior, 9*(3), 160-168.

Zuckerman, M., DePaulo, B. M., Rosenthal, R., & Leonard, B. (1981). Verbal and Nonverbal Communication of Deception. In Advances in Experimental Social Psychology (Vol. Volume 14, pp. 1-59): Academic Press.

APPENDIX A

Construct Boxplot Graphs of Measures at Q/R Level

Construct bar graphs with errors.

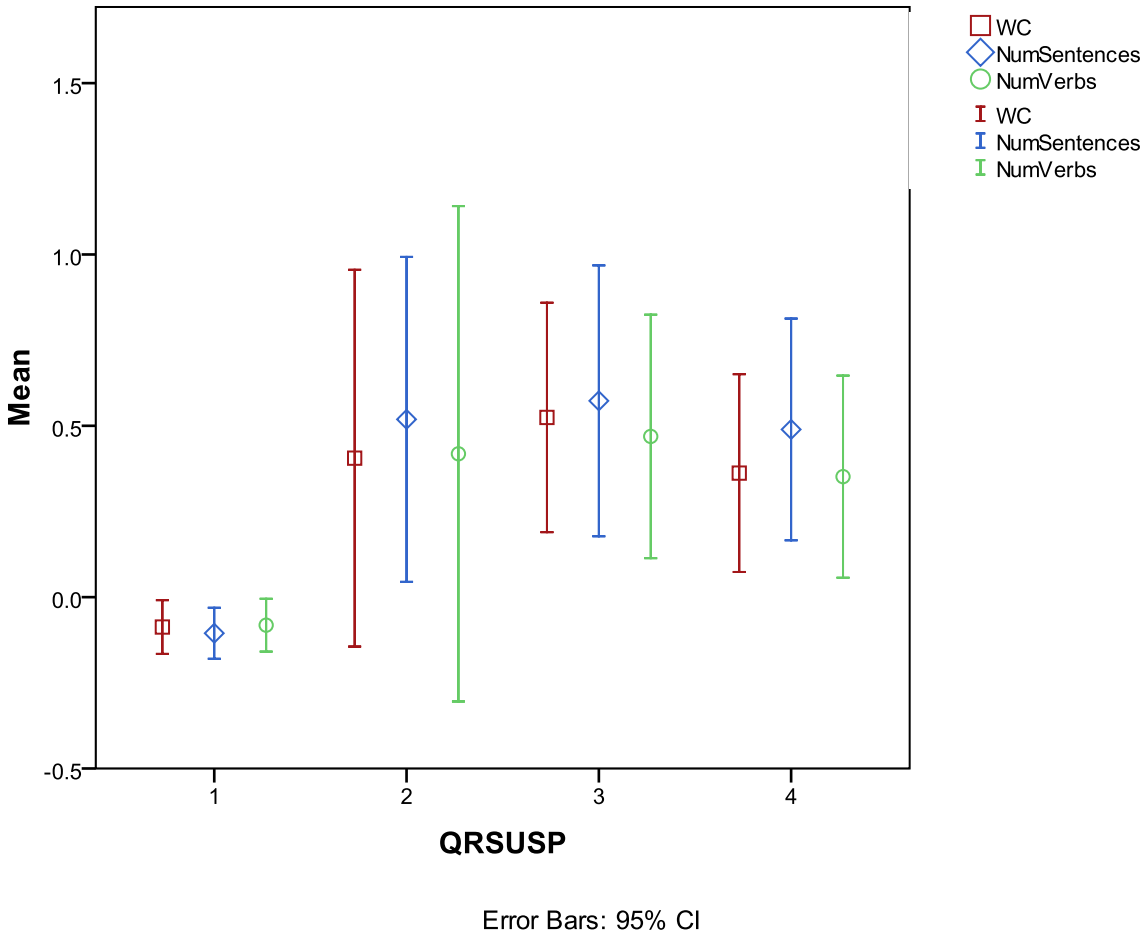


Figure 28, Quantity Construct

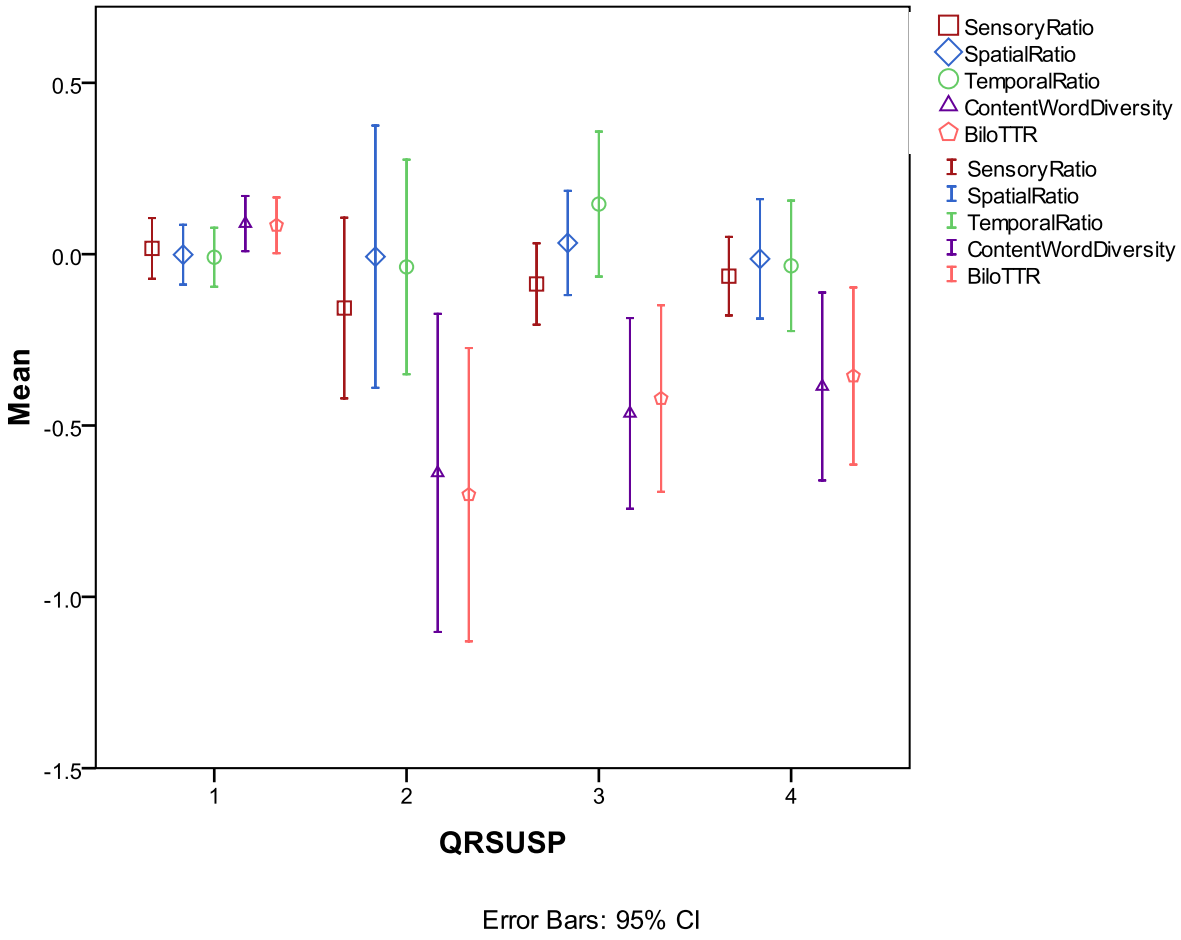


Figure 29, Specificity Construct

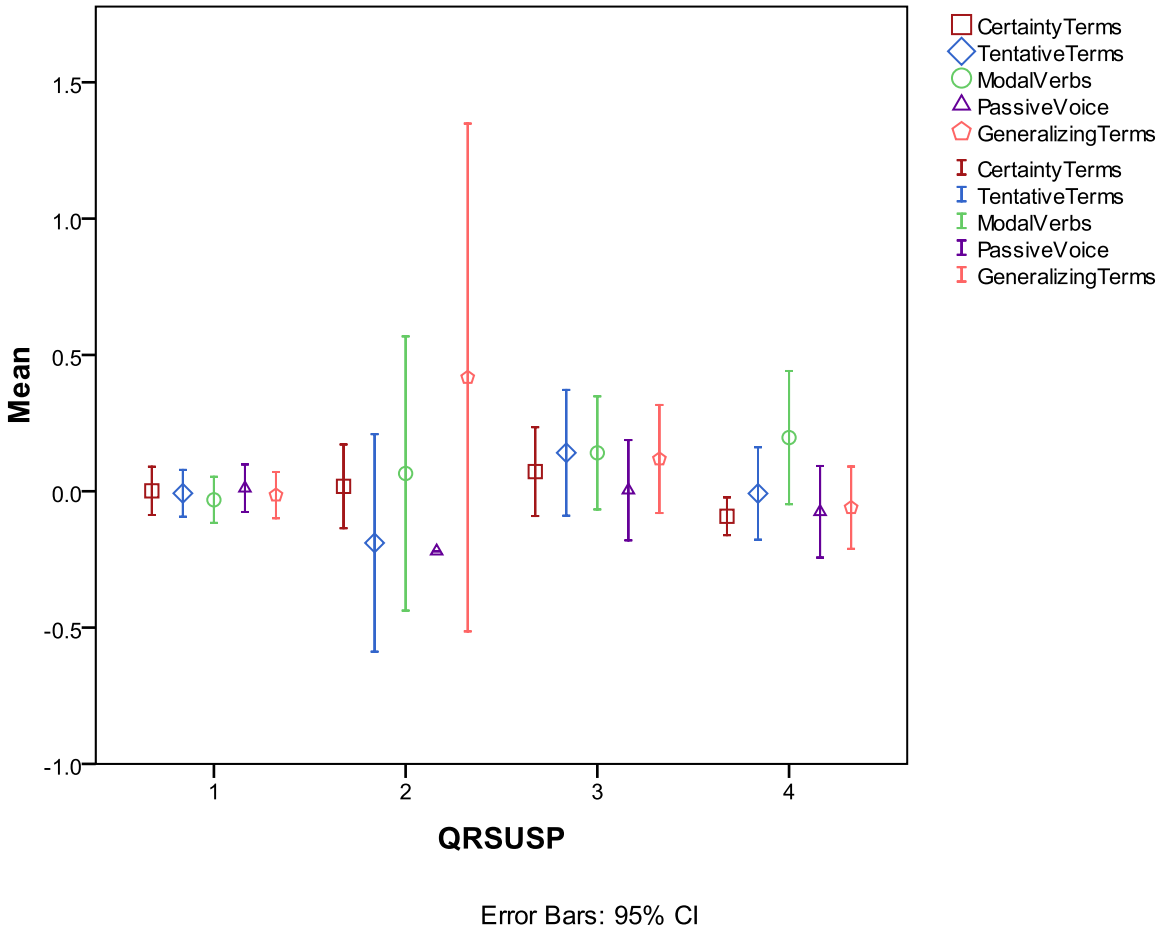


Figure 30, Uncertainty Construct

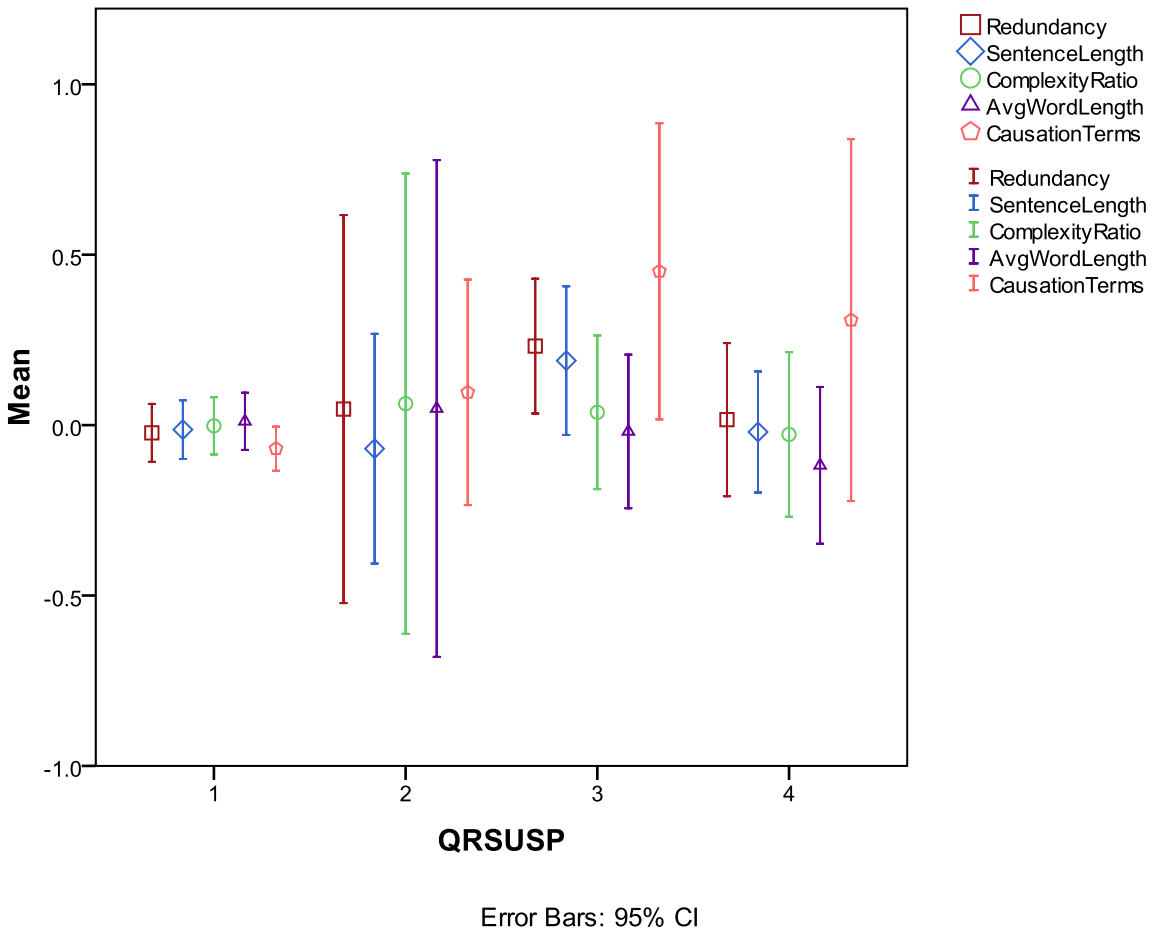


Figure 31, Clarity Construct

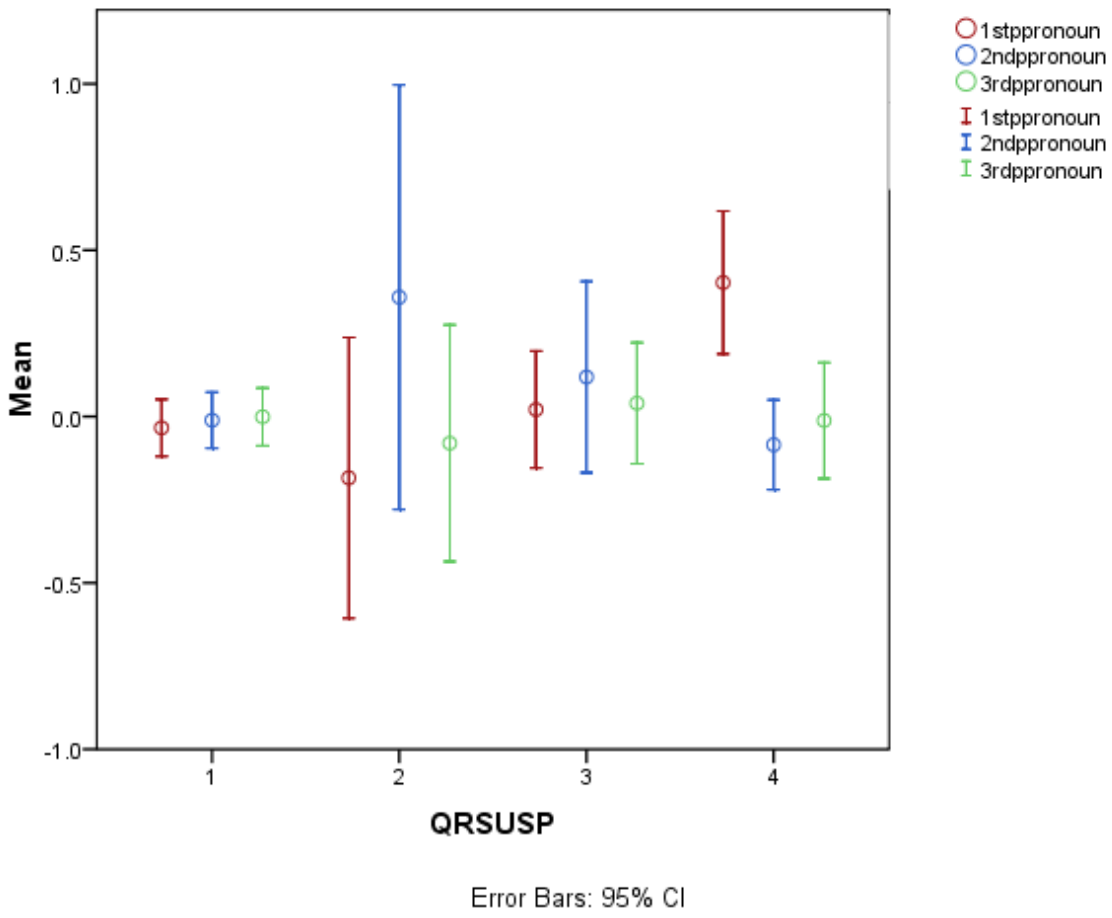


Figure 32, Immediacy Construct

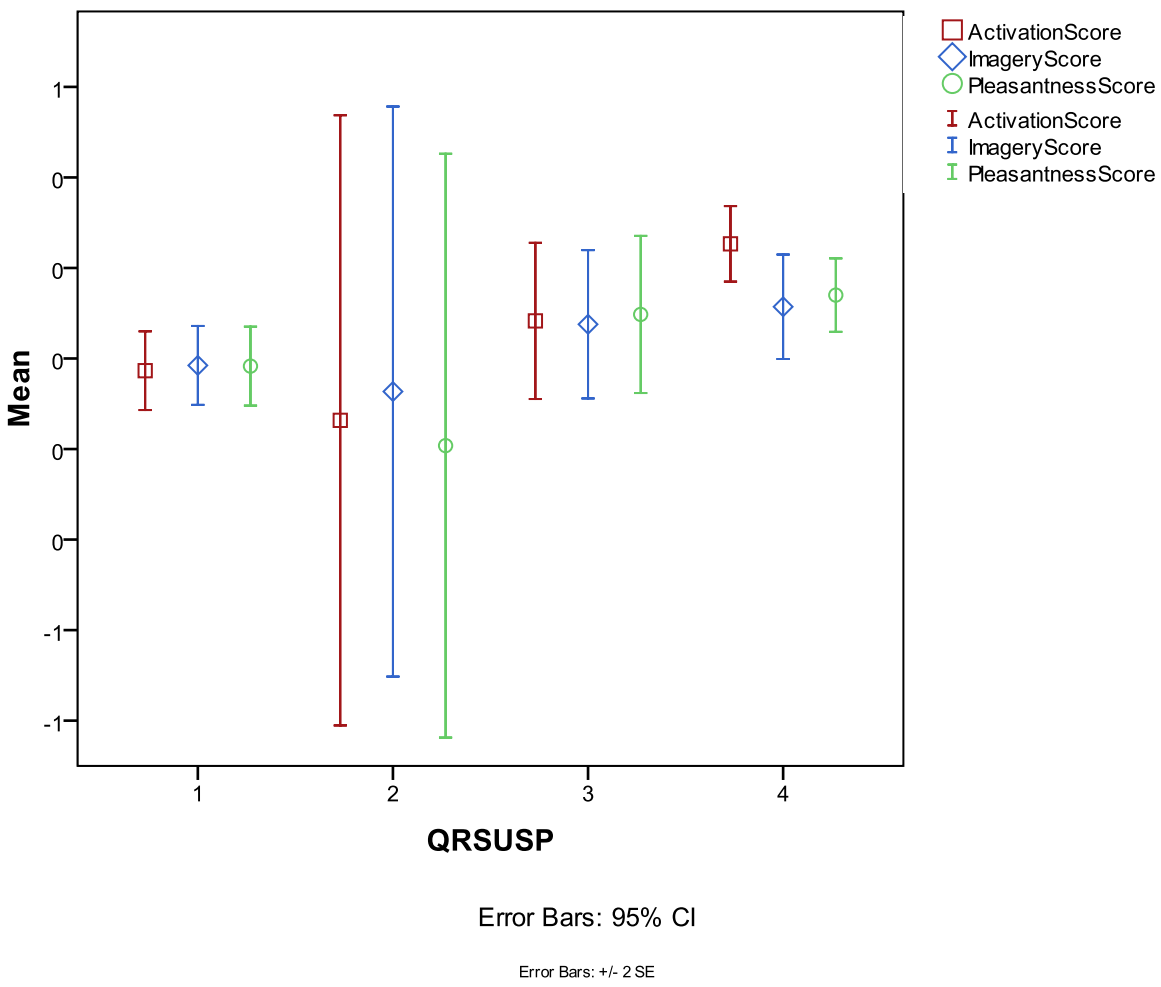


Figure 33, Affect Construct

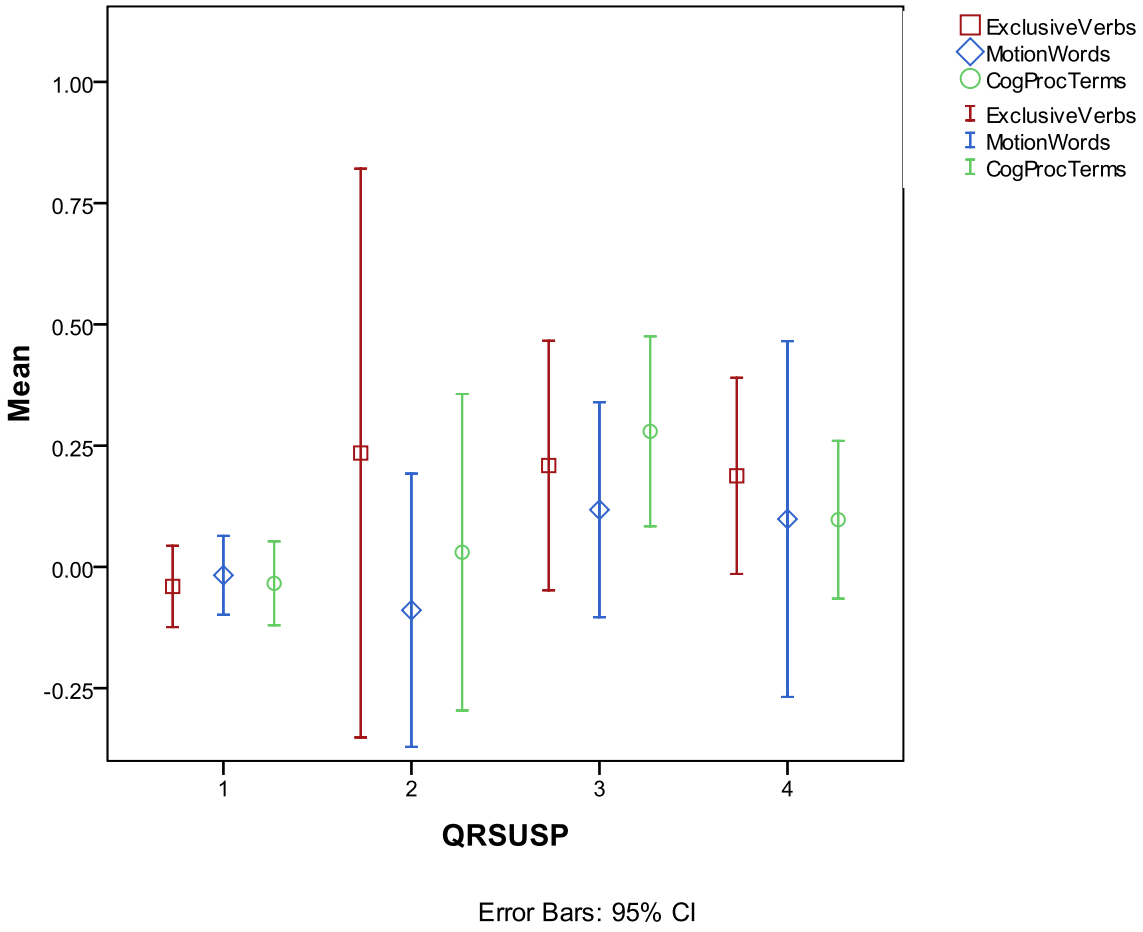


Figure 34, Cognitive Processing Construct

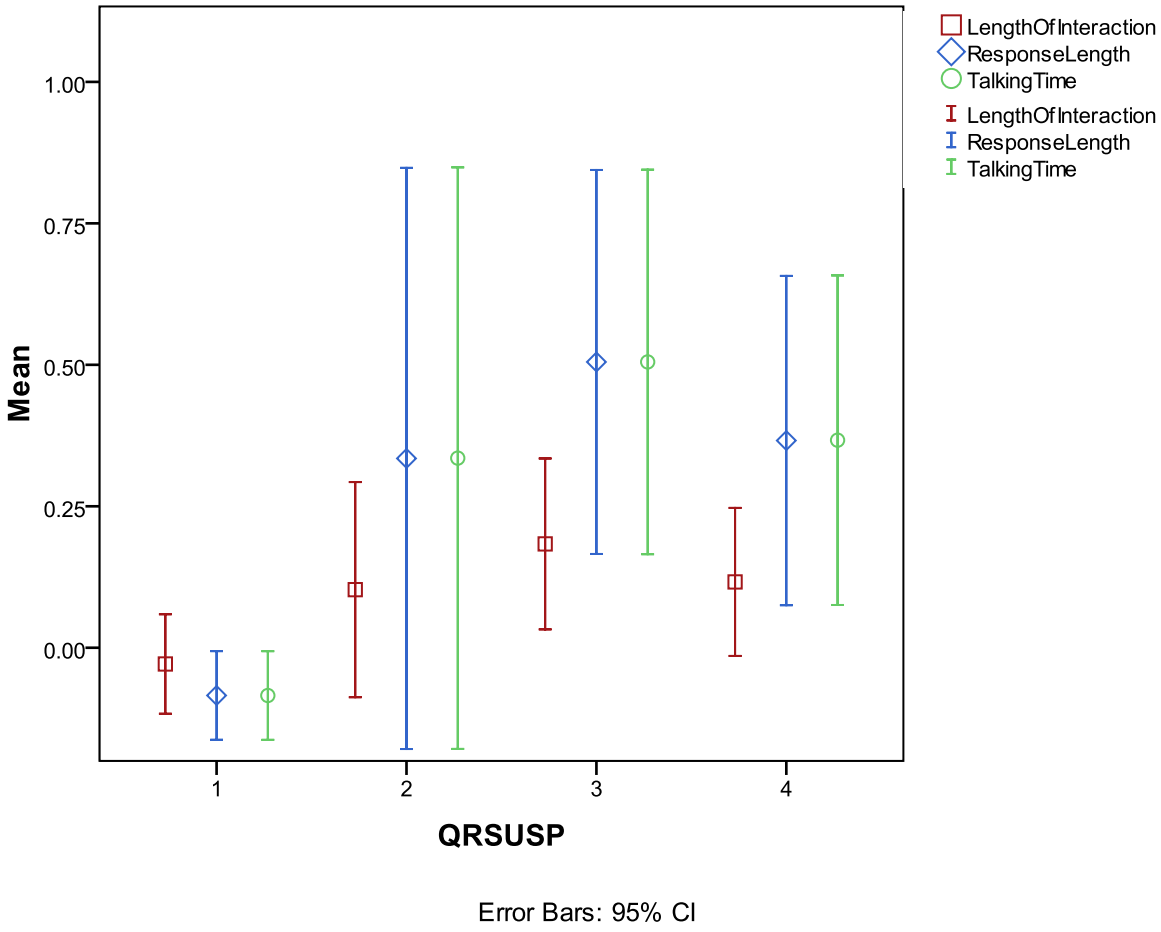


Figure 35, Time-Duration Construct

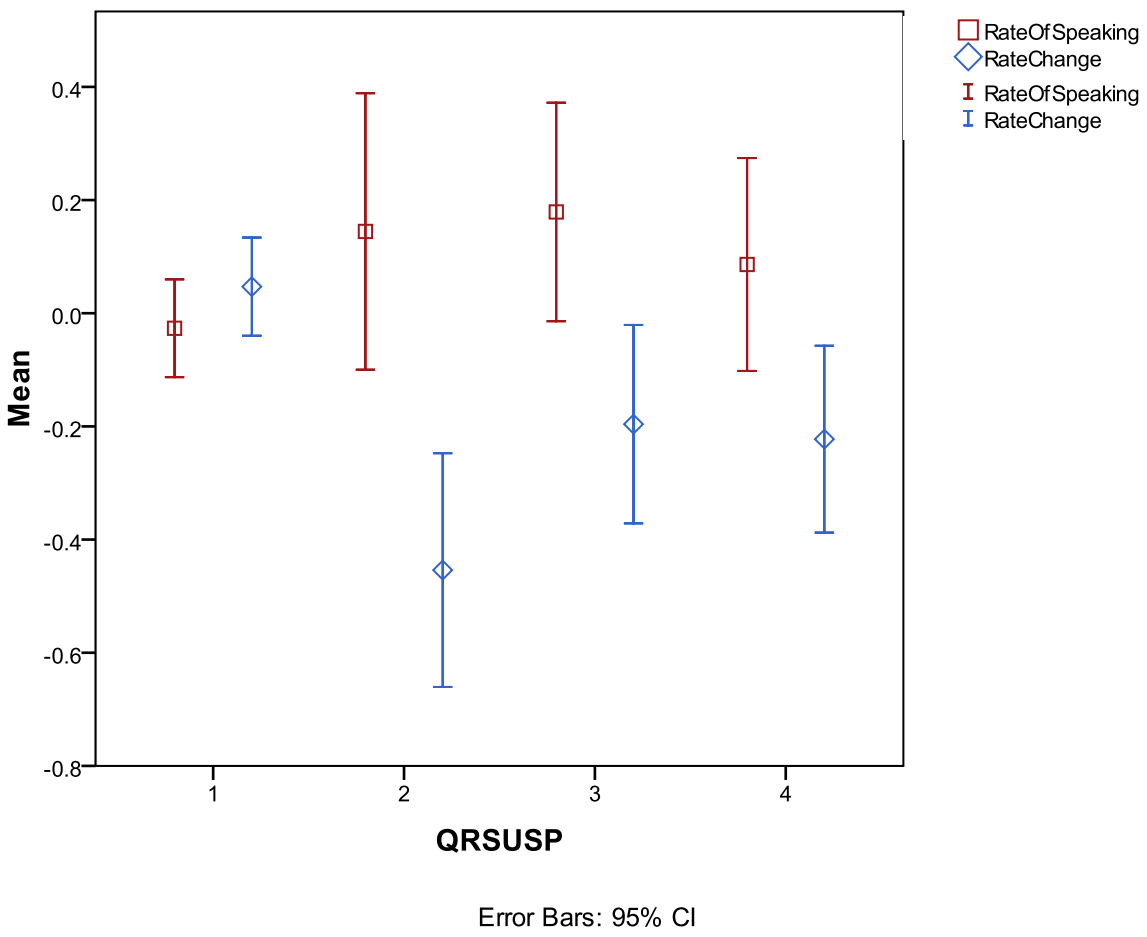


Figure 36, Time-Tempo Construct

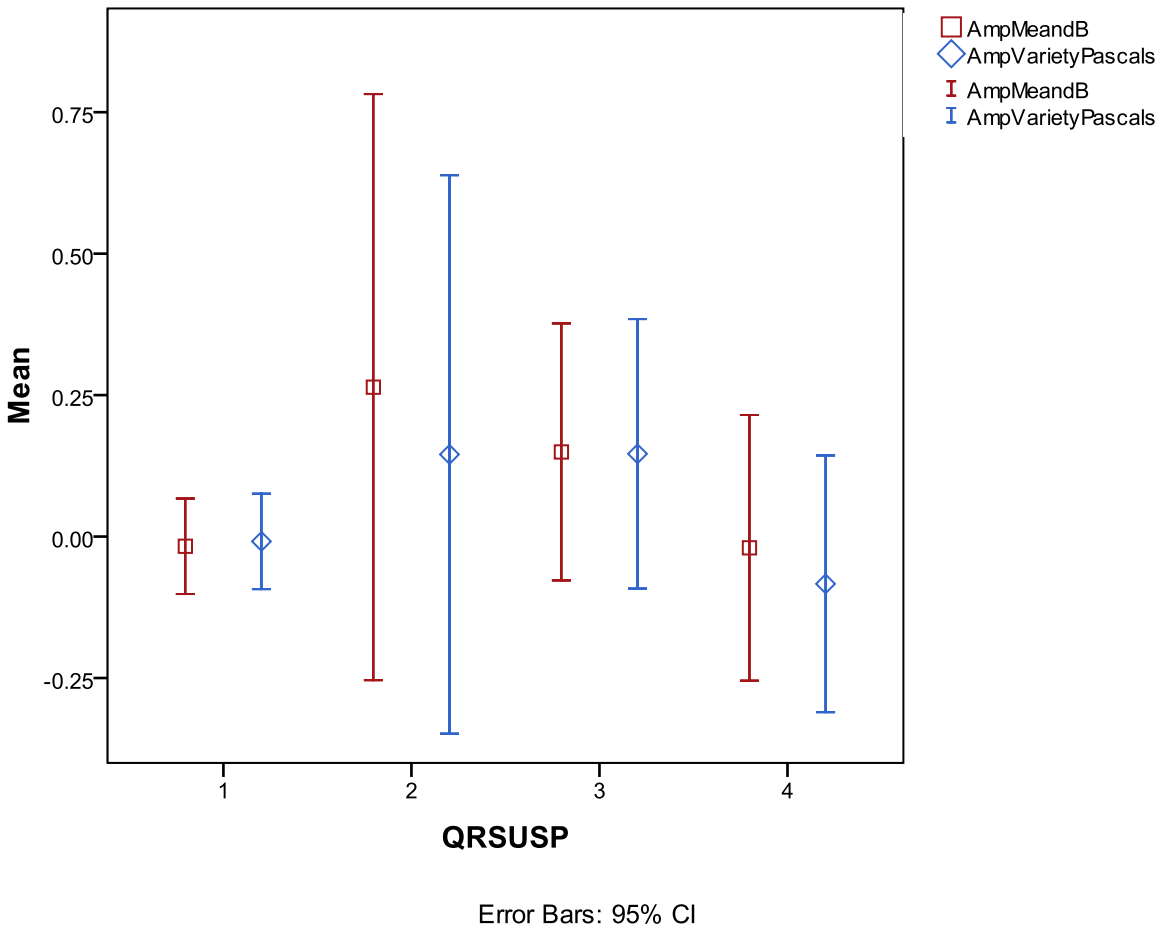


Figure 37, Intensity Construct

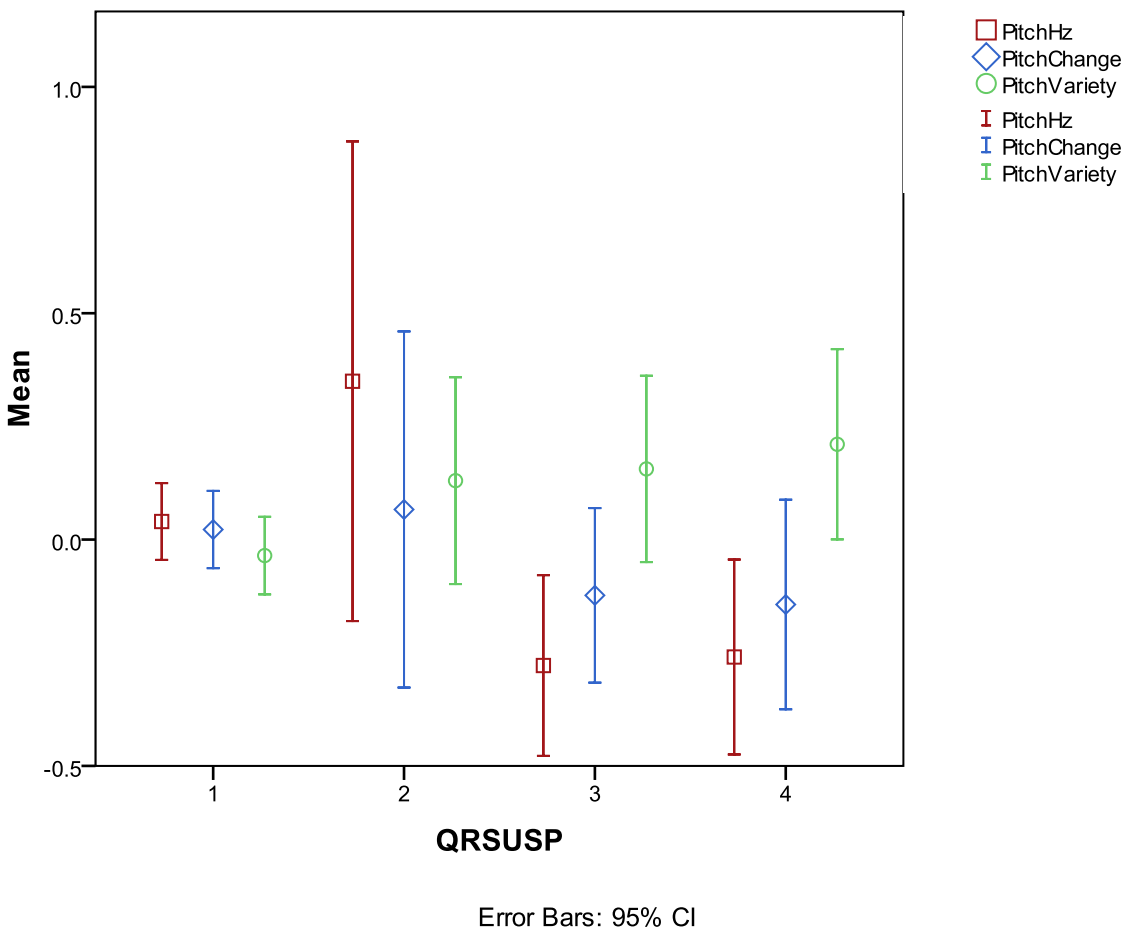


Figure 38, Frequency Construct

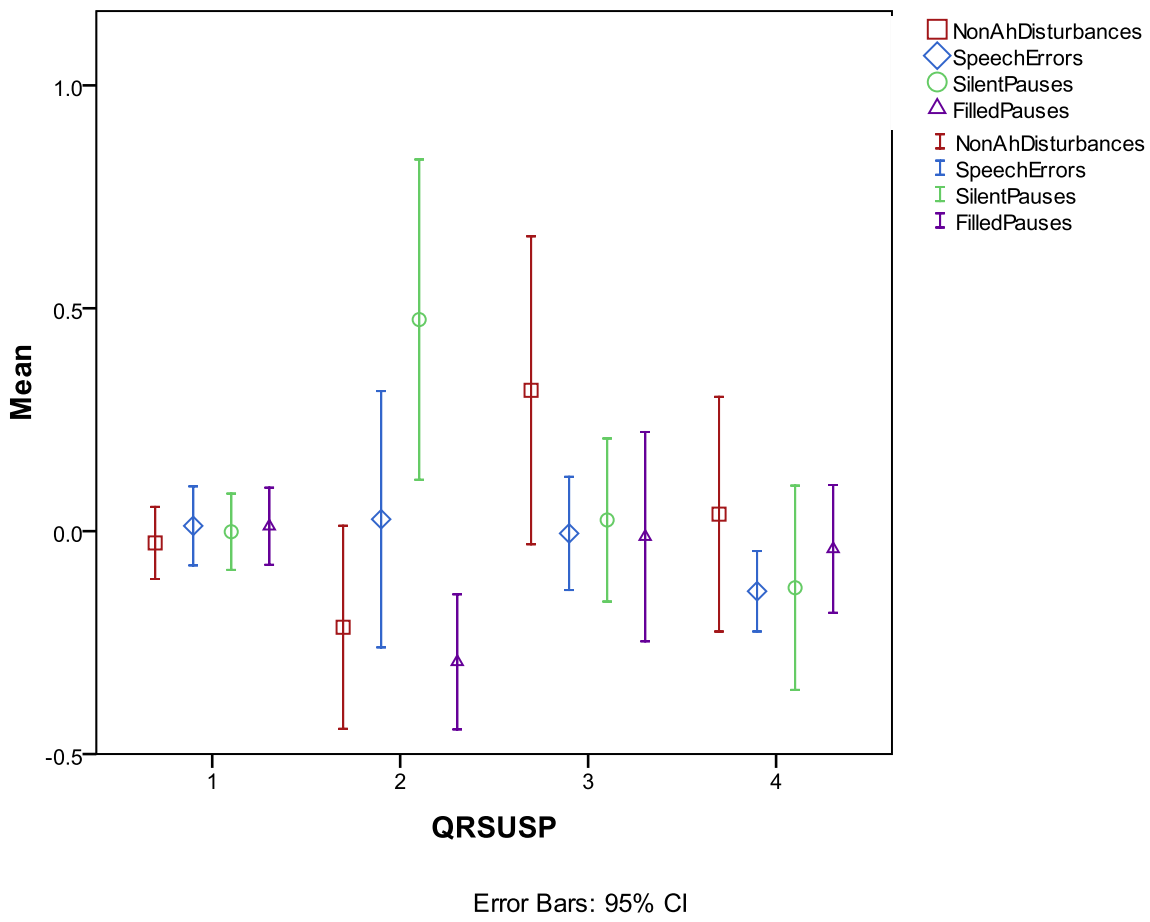
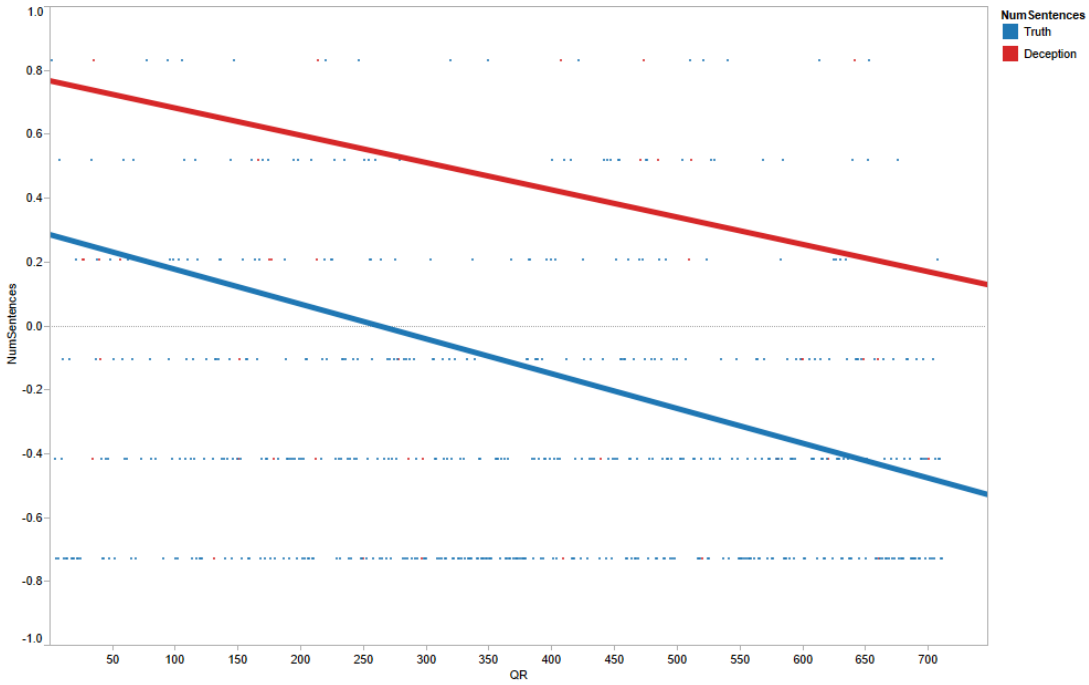


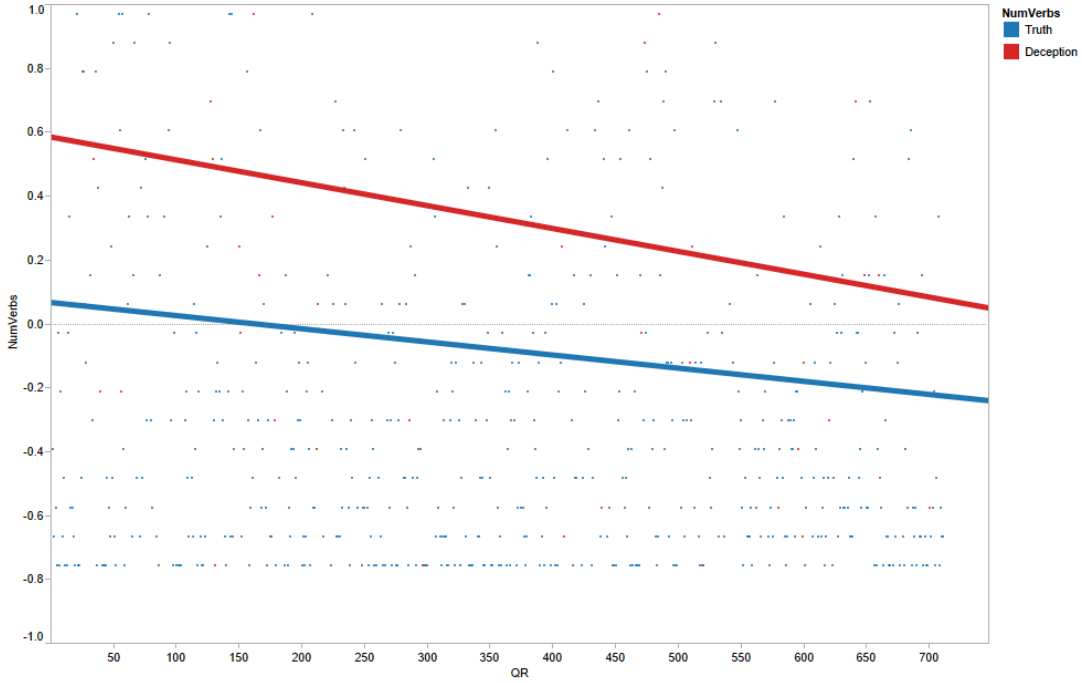
Figure 39, Fluency Construct

APPENDIX B

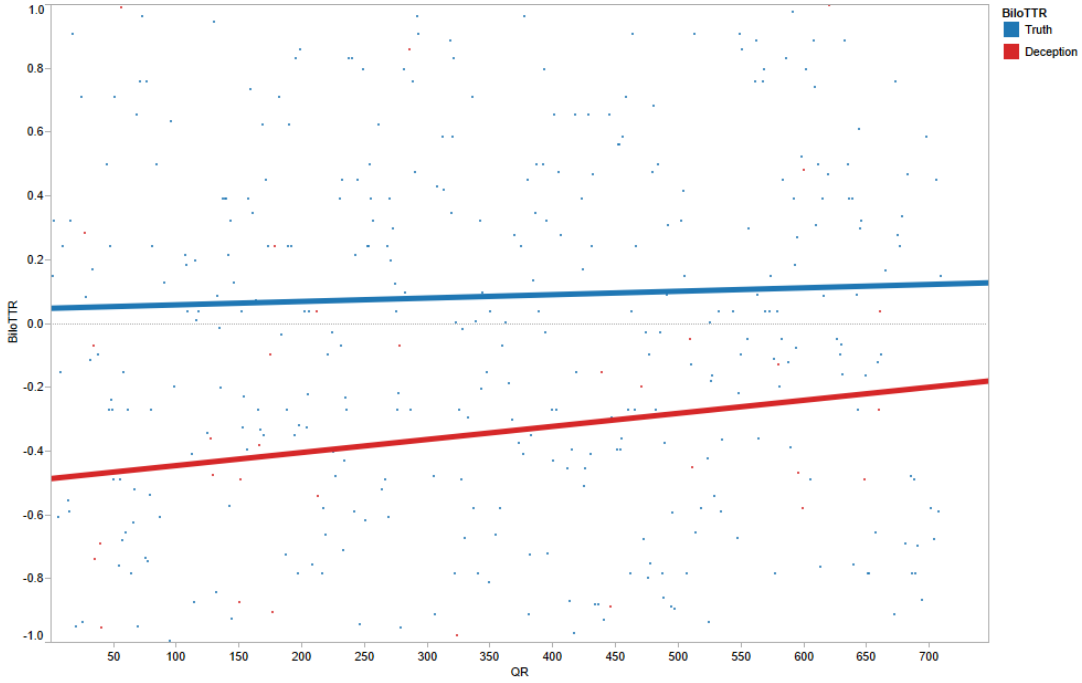
Construct Bar Graphs of Measures at Q/R Level



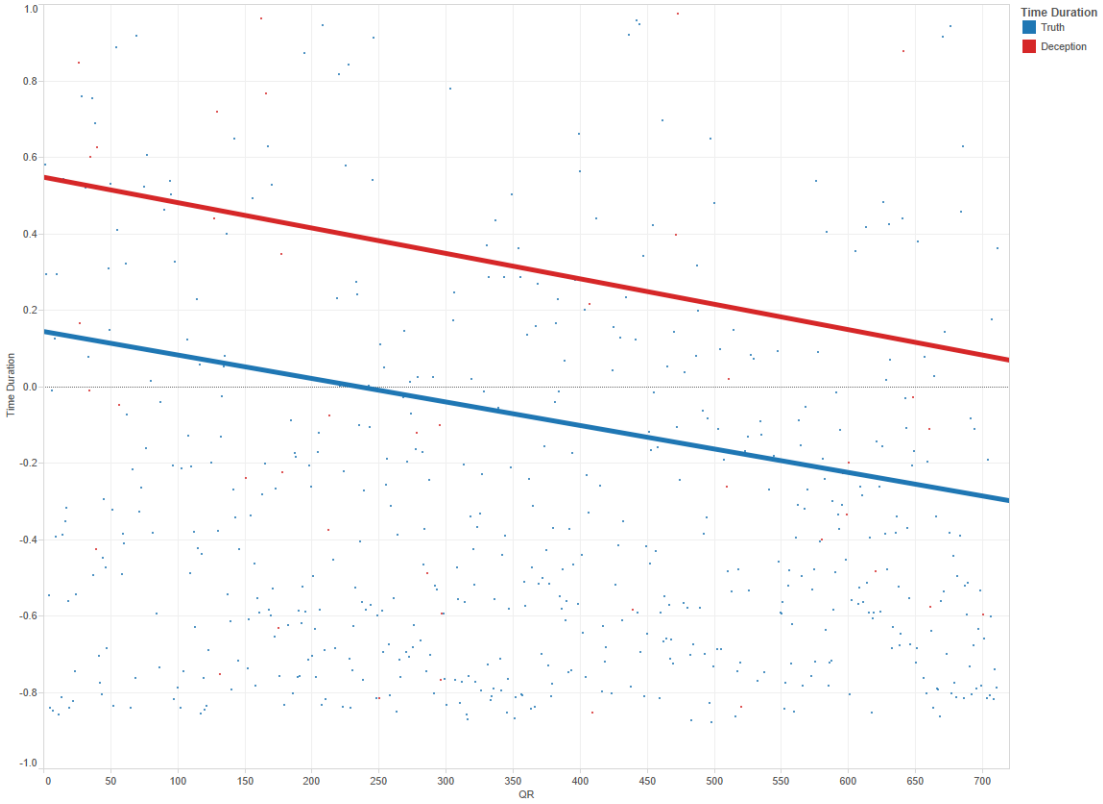
The plot of sum of NumSentences for QR. Color shows details about QRSUSP.



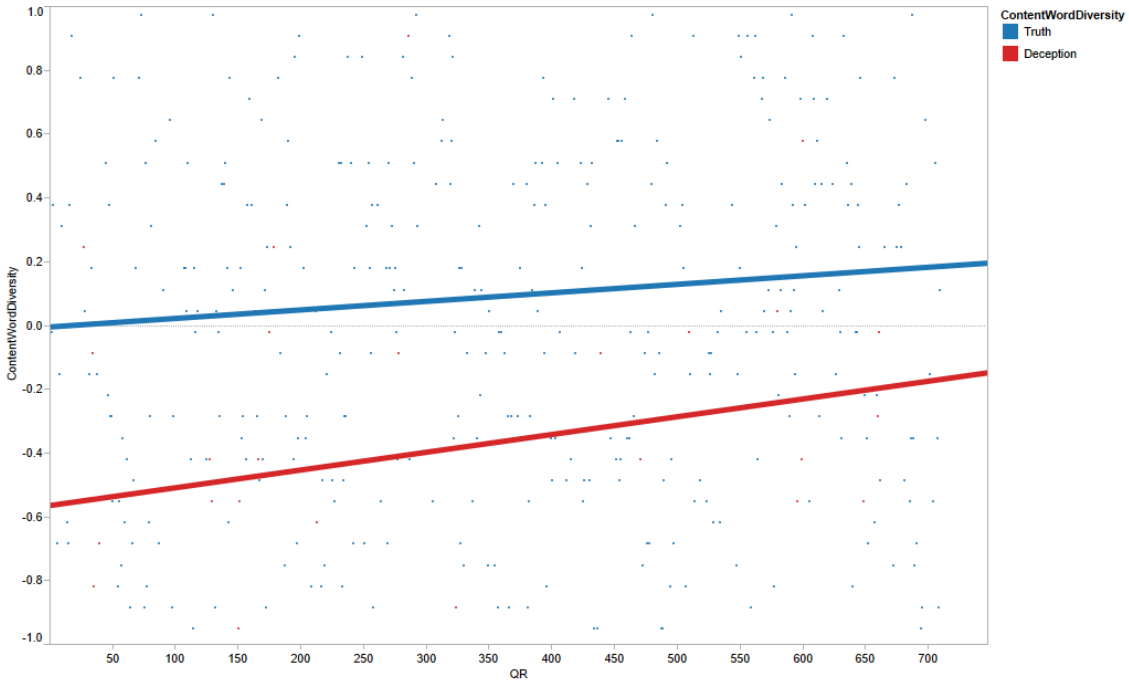
The plot of sum of NumVerbs for QR. Color shows details about QRSUSP.



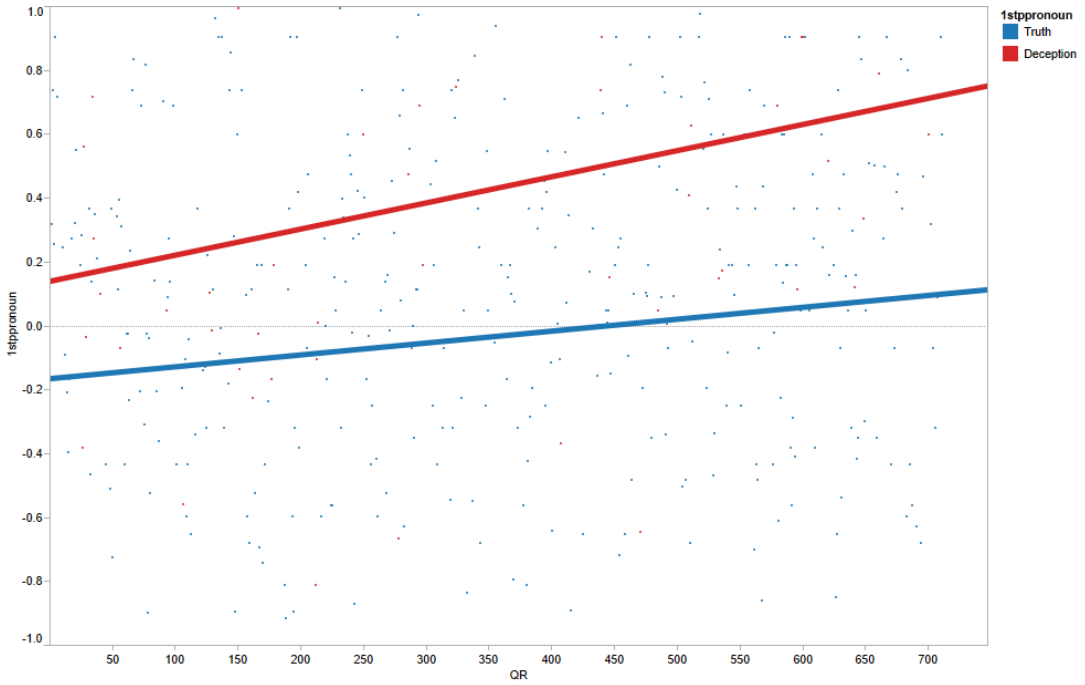
The plot of sum of BiloTTR for QR. Color shows details about QRSUSP.



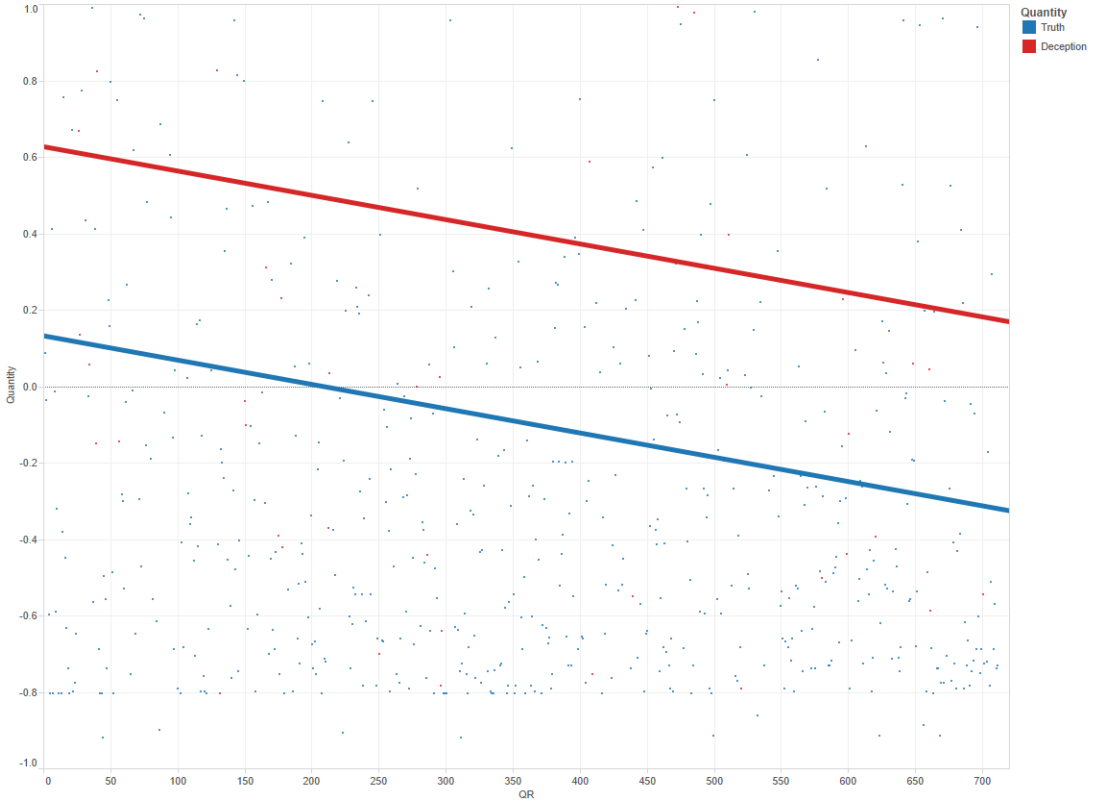
The plot of sum of Time Duration for QR. Color shows details about QRSUSP.



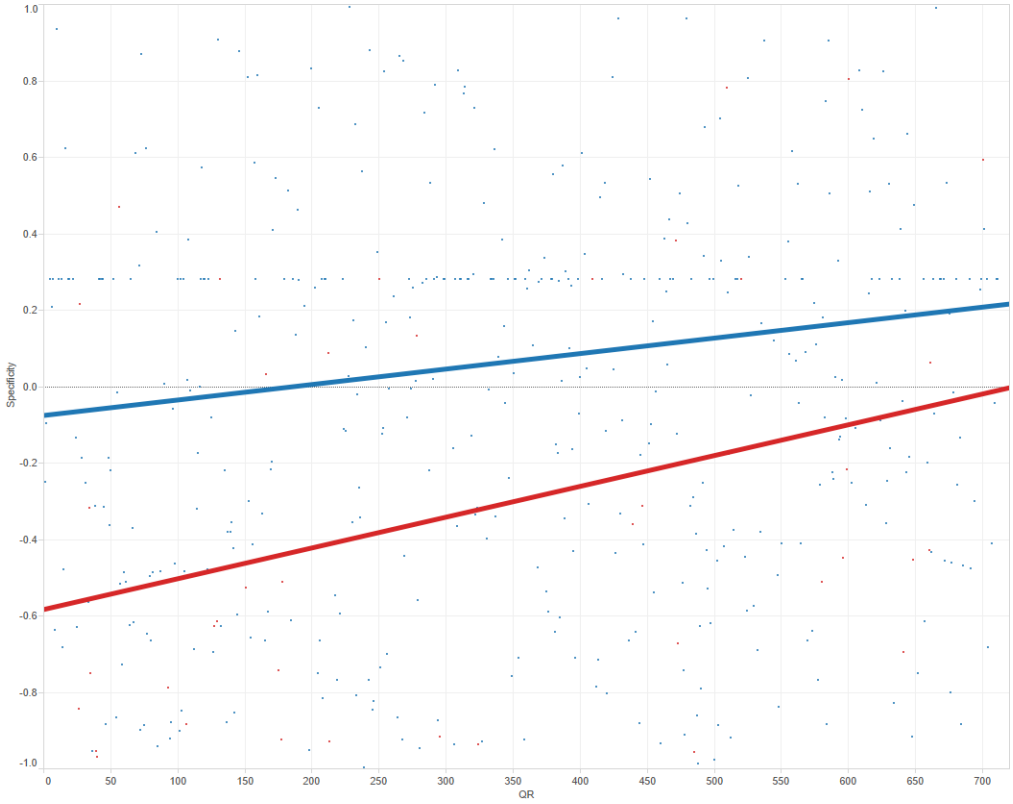
The plot of sum of ContentWordDiversity for QR. Color shows details about QRSUSP.



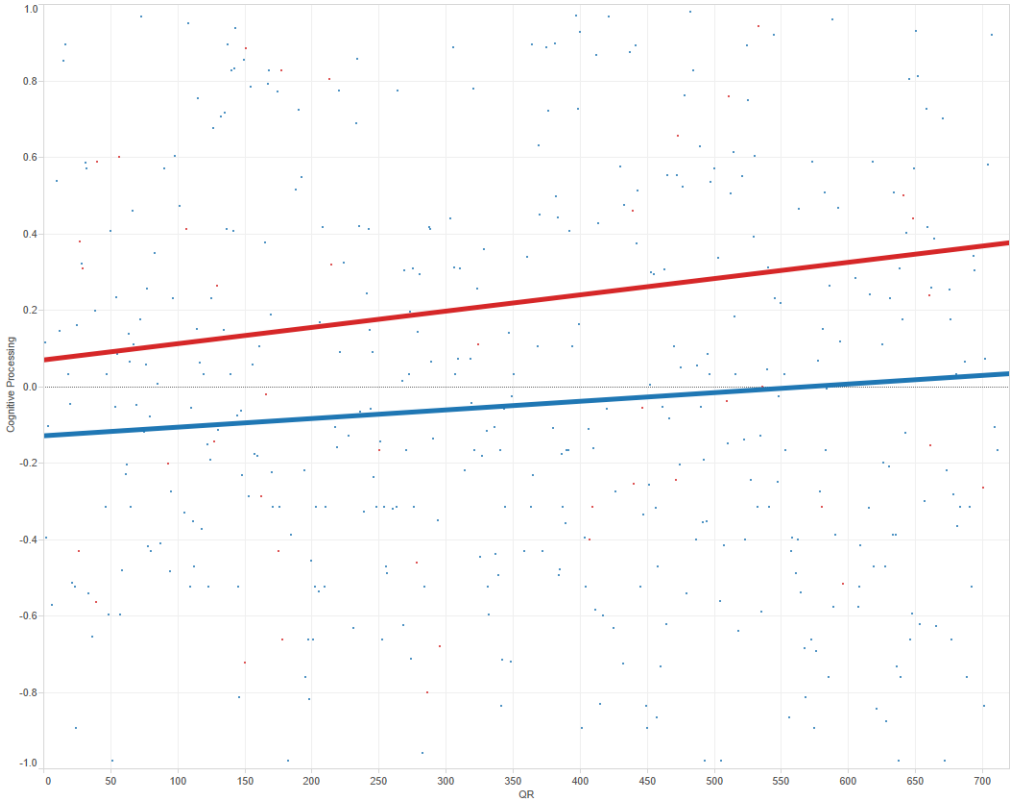
The plot of sum of 1stppronoun for QR. Color shows details about QRSUSP.



The plot of sum of Quantity for QR. Color shows details about QRSUSP.



The plot of sum of Specificity for QR. Color shows details about QRSUSP.



The plot of sum of Cognitive Processing for QR. Color shows details about QRSUSP.

APPENDIX C

Charges Against James Perry

What follows are the formal charges brought against James Perry that have been released to the public. Because of the nature of some of his crimes some of the details have been redacted or omitted.

COUNTS 1 AND 2

Your complainant has reviewed the police reports of Madison Police Officer Denise Marx who reports that on May 11, 2000, at 4:49 p.m., she was dispatched, along with numerous other officers, to the Learning Shop, 5225 High Crossing Blvd., in Madison, Dane County, Wisconsin. At that location she spoke to [redacted] who identified herself, who was crying and had an extremely red face. The Learning Shop is a retail establishment in a strip mall, which faces out onto a parking lot area. [redacted] made statements that included the following. She has worked at the store for quite some time and that she was at work on May 11, 2000 for the day and that at approximately 4:00 p.m., the man who assaulted her came into the store. He said hi; he went over to the jump rope area located approximately two aisles from in front of the store; he looked at some jump ropes; picked some up. After inspecting jump ropes for about 15 minutes he then left. Ten minutes later she observed him walk by the front of the Learning Shop on the outside sidewalk away from American TV, which is in the strip mall, towards Animart, which is also in the strip mall. He looked into the store at her and "grinned."

[redacted] told Officer Marx that at approximately 4:40 to 4:45 p.m., the same person came back into the store. She was standing back behind the checkout counter, and he stood on the opposite side of the counter from her. There was another customer, a woman, in the store. [redacted] asked the person who assaulted her if she could assist him, and he told her he had been in the store earlier and was looking at some globes. He told her he was looking for a normal globe, not a computer program globe. She exited from behind the cash register and began walking to the middle of the store to the east side where the globes were kept. He followed her. When they got to the globe section, she pointed out the larger and the smaller globes and then began pushing them back so that they were secure on the shelf. When she reached up well above her head to secure the globes, she could feel her assaulter standing

directly behind her. She said he was so close to her that she could feel his stomach in the small of her back. He put his left arm around her neck and held something sharp against the right side of her neck. The item felt cold and straight against her skin. He said, "Don't scream, do you feel this knife on your neck?" He said to her, "Don't make a sound."

With the knife to her neck, he said to her, "Walk with me, don't make a scene," and began walking [redacted] with him behind her, towards the front of the store. He was no longer holding the knife to her neck but holding her with his arm over her right shoulder across her collarbone, escorting her to the front of the store. They went up toward the front of the store and then went around one of the storage racks and again headed toward the rear of the store. It seemed as if he was looking around the store to see if there were any customers, and apparently he did not notice the other woman in the store. When they got to the back of the store to the last aisle where books are kept, he stood in the main aisle of the store, and she was directly to his right facing one bookshelves. He said to her, "Just calm down, don't do anything." She said to him, "I'll give you what you want. Do you want money?" He responded, "No," and told her to go to the back room.

[redacted] stated that the person who assaulted her then stood directly behind her so she could feel his stomach in the small of her back. He put his two hands on her collarbone area and attempted to push them down across the front of her body. She believed he was attempting to touch her breasts. He was unsuccessful because the entire time she had her two hands vertically against her chest. He was unable to touch her breasts for this reason. He again had the knife in his left hand and stated, "Let's go in the back room." [redacted] said the door to the back room was open. She responded, "No because there's a customer up there that needs help, and I can't leave her alone." He then told her the customer was not ready. She insisted she could not leave the customer out in the store by herself. At one point when he was trying to direct her to the back room she saw the knife in his left hand as he turned away from her. She described it as a folding knife with a dark red plastic handle. The knife and the blade were about the same size, and the blade folded into the handle. When the knife was fully extended, the length of the end of the handle to the end of the blade was approximately six inches.

After [redacted] refused to go into the back room, he said to her, "You're making this difficult for me, I have to leave." [redacted] said he then told her to go into the bathroom, which was located at the end of the aisle of the books. She stated, "No, no." He told her to go into the bathroom so he would have time to leave. He then began walking towards her as she backed away towards the bathroom. She did then walk into the unlit bathroom, and he then closed the door from the outside. He said, "Wait here and give me time to leave." She locked the bathroom door and waited in there for several minutes. When she came out of the bathroom, the other customer was still in the store, and [redacted] asked the customer if the man had left, and the customer stated he must have because she heard the front door close. Nothing appeared to be missing from the store; the cash register had not been opened.

Your complainant states she spoke with [redacted] on the date of the attack and noted that [redacted] appeared to be in a heightened state of distress and anxiety. [redacted] made statements that included the following. The suspect held the knife so close against her neck she thought for sure she would have an injury and he would kill her. The suspect repeatedly tried grabbing at her breasts and she struggled to keep his hands off her. The suspect repeatedly ordered her into the back bathroom where she believed he was going to sexually assault her [redacted] refused to go with him. Your complainant asked [redacted] if the suspect had

appeared to stalk her prior to the attack in the store. [redacted] said yes, and told your complainant she saw the suspect pass back and forth across the front of the store looking at her before he came in to attack her. Your complainant asked [redacted] if she noticed whether the suspect was smoking any cigarettes while he was outside the store or if she recalled the odor of cigarettes on his breath. [redacted] said she did not notice him smoking, and at this time could not remember if he smelled of cigarettes.

Your complainant and Special Investigator Meredyth Thompson of the Madison Police Department report that they were investigating the above-described crime together on May 11, 2000, after 5:00 p.m. Your complainant reports that the weather on this date was rain throughout the day up until approximately 4:00 p.m. In observing the garden areas which fronted the stores in the strip mall, your complainant noted that there were very few fresh looking cigarette butts which were not rain soaked which she and Investigator Thompson were able to locate. Your complainant directed Investigator Thompson to recover all three dry cigarette butts which appeared to be recent but untainted by rain located in the outside garden area in front of the Learning Store. Your complainant directed Investigator Thompson to convey these cigarette butts to the Wisconsin State Crime Laboratory for purposes of developing DNA profiles for future comparison of possible identification of DNA which may be present on the cigarette butts which would allow identification of the person(s) who had smoked those cigarettes. Investigator Thompson conveyed the cigarette butts to the crime laboratory where they were processed by crime lab Serologist Jill Hertzberg for DNA profiles. (See "Wisconsin State Crime Laboratory Results" section of the complaint below for a positive DNA match between the defendant James Perry and one of these butts.)

On April 8, 2004, your complainant, along with Detective Ann Turner, spoke with [redacted] who identified herself verbally and who stated that she is the wife of the defendant herein James Perry. They then shared a residence at 800 West South Street in Stoughton, Dane County, Wisconsin. Your complainant asked [redacted] if her husband had any pocket knives. She said yes and pulled out a red-handled Swiss Army knife and a couple of other pocket knives from among his belongings. She stated that the defendant also had a second similar red-handled pocket knife amongst his tools in his toolbox. Your complainant asked [redacted] whether or not her husband smoked cigarettes, and [redacted] stated, "Yes, he always smoked Marlboro reds."

On April 27, 2004, your complainant and Detective Bruce Frey of the Madison Police Department responded to the home of [redacted] where they showed her a photo array in a sequential-style photo presentation in which [redacted] was given only one photo at a time to view before moving on to view the next photo. Each of the photos was placed in a file folder. Your complainant shuffled the photos before giving them to Detective Frey to present to [redacted] One photo was that of James Perry, the defendant herein, but Detective Frey did not know which one this was, when he presented each of the file folders containing the photos to [redacted] for viewing. [redacted] was unable to identify her attacker as anyone from the photo array but later reported that she thought that seeing the offender in person, the way he walked, or hearing his voice, might help her better identify her attacker. There were six photos in the photo array made from records made and kept in the regular course of the business of the Dane County Sheriff's Office and regularly relied upon by law enforcement officers throughout the state of Wisconsin to be truthful, accurate, and reliable.

On May 11, 2004, your complainant and Detective Steve Reinstra of the Madison Police Department administered an in-person lineup. Detective Reinstra and your complainant spoke with [redacted] and explained the lineup procedure and the witness lineup identification form. Detective Bruce Frey accompanied [redacted] into the line-up room and witnessed her viewing the in-person lineup. At the conclusion of the lineup, [redacted] checked one of the boxes indicating she recognized a subject in the lineup. [redacted] passed the sheet to Detective Frey, at which time Detective Frey noted that she had written the word "yes" under the person identified as No. 2. Detective Frey reports that he presented this document to Attorney Charles Giesen, the defendant's attorney, who signed the document in front of Detective Frey, and Detective Frey in turn signed it and passed it back to [redacted] who signed and dated the form.

Detective Frey reports that he then escorted [redacted] to a side office room within the Madison Police Department and sat down with her and asked her specifically how she thought the lineup procedure went. She stated, "I picked out No. 2." She stated she came to that conclusion because "I just really had a gut feeling that was him." She explained that prior to being attacked in the Learning Shop store, he had walked past the front of the store while she was near the front counter area several times and she had noticed him and made eye contact with him prior to him entering the store. She stated that during the several passes he made, she had a distinct look at him and profile, and when she looked at the in-person lineup and the profile of this person, it also jarred her memory. She stated that getting a look at him physically standing sideways gave her the reaction that caused her to select No. 2 because she recognized him from that experience with him at the store and then seeing him again in the lineup. Your complainant confirms that the defendant herein James Perry had selected his own place in the lineup, and that was position No. 2. He is the person that [redacted] selected.

[redacted] told Officer Marx that she never consented to the defendant's attempts to touch her breasts and sexually assault her or to intentionally restain her.

COUNTS 3 AND 4

Your complainant reviewed the police reports of Madison Police Officer Alison Radzicki, who reports that on August 11, 1999, at 7:04 p.m., she was dispatched with other Madison police officers in regards to a sexual assault. Officer Radzicki spoke with [redacted] who identified herself verbally and who stated that she worked at the We Tan U Spa located in a strip mall at 4261 Lien Road in Madison, Dane County, Wisconsin [redacted] told Officer Radzicki that at the time of the offense, approximately 6:55 p.m. [redacted] was standing right outside of the door of the afore-described business smoking a cigarette when her assaulter approached her. He peered through the window of the building as if he was trying to see something. She asked him if she could help him. He said he was trying to see how much it would cost to get ten tans. She told him that he could come inside and she would explain it to him. He said no, that wasn't necessary. She then told him that it was all right, that she would go inside and get him the information. At that point they both went inside. After she gave him the information, he said, "I wasn't going to buy any anyway." They both then walked out the door together, her to finish her cigarette break. She stood just to the north of the entrance of the store. He walked away southbound toward Glamour Shots, another store in the strip mall. Suddenly he turned around and came back to her, pushed her up against the glass window with his right hand, and held her there while he grabbed her left breast with his left hand. He held onto her breast for at least 20 to 30

seconds. [redacted] was stunned. She said to him, "What are you doing?" He then stepped back from her approximately two to three feet, and she could then see that his penis was out of his pants, naked, in his hands, and he was stroking it. She said that it was "a little bit [erect] but it was not a full erection." He then looked at her and said to her, "I want to fuck you, that's what." She immediately ran in the store and locked the door behind her, and in the rush to do so, slammed the door on one of her fingers. He then walked casually away from the store in a southerly direction. While she was on the phone with police dispatch, she noticed that he was standing near the store Millennium that is directly across the parking lot from We Tan U in this same strip mall where she works. She saw him looking around in a casual fashion, and she gave that information to dispatch. He suddenly turned around, faced south, and started to walk away briskly. She told Officer Radzicki that during the assault she couldn't tell necessarily whether or not his penis ever touched her but she was paying more attention to the fact that he had grasped upon and was holding onto her breast.

Your complainant reports that on April 28, 2004, your complainant met with [redacted] at the North Madison Police District to view a photo array. Detective Kevin Linsmeier assisted your complainant as a lineup administration by shuffling six Dane County Jail mug photos and placing them randomly in eight file folders with folder seven and eight remaining empty. [redacted] viewed all of the photographs sequentially before returning to look at the photo in folder No. 2. She viewed No. 2 and then discarded it. She asked to view No. 3 again and then stated, "This guy looks like the guy who attacked me. He's got the same features. I specifically remember the vivid blue eyes that he had and something wrong with his face, like acne or something. This is the one but he looks like he's gained a little weight since that time." The defendant herein, James D. Perry, was photo No. 3 in this photo array, and thus was identified by [redacted] as being the person who assaulted her. She signed and dated the photograph. [redacted] had told your complainant during her investigation of this offense on January 31, 2001, "The thing I remember most about him is his very, very distinct bright blue eyes." Further, your complainant knows from speaking to [redacted] the defendant's wife, that her husband, the defendant herein, had a problem with acne that came and went depending on how stressed he was. She stated his acne became worse when he was under stress.

[redacted] told Officer Radzicki that she did not consent to the defendant throwing her up against the window and fondling her breast with his left hand.

COUNTS 5, 6, 7, AND 8

Your complainant has reviewed the police reports of Village of Marshall Police Officer David Viken, who reports that on May 17, 2000, he was in uniform, operating a marked squad car and was called to and did speak with [redacted] who identified herself verbally and who stated that for the past month or so her daughter [redacted], date of birth 04/02/1984, has been doing bad at school and been having bad headaches. She had been disciplining [redacted] for her actions and on the night of their conversation, May 17, 2000, [redacted] told her why she had been doing badly in school. [redacted] told her mother than on [redacted] birthday on April 2, 2000 at around 10:15 p.m., she was walking home on Hubbell Street in the Village of Marshall, Dane County, Wisconsin. She had left her friend's home at [redacted] Hubbell Street and was walking north. When she got to the lumber yard, a car pulled next to her. The lumber yard is located at 822 Hubbell Street in Marshall, Dane County, Wisconsin. A man in the car, a white male, looking like a businessman, stopped and asked her for directions to Sun Prairie, Wisconsin, and told her once he got to Sun Prairie he

would know where he was going and asked [redacted] if she could give him directions and she told him that she could. He then gave her a piece of paper and asked her to draw the directions for him. He opened the back door and told her to grab a pen that was laying on the back seat. When [redacted] when to grab the pen, he pushed her into the car. She tried to get out. The doors automatically locked. She kicked and screamed to get free. He punched her in the chest, knocked the wind out of her, slapped her in the face a couple of times, and told her to shut up or he would kill her. He then drove to Converse Park in Marshall about three blocks south of the lumber yard on the Best Built Parkway. Once at Converse, he raped her and told her if she told anyone, he would kill her, and then drove off, and she walked home.

Your complainant has also reviewed the police reports of Village of Marshall Police Lieutenant Brian Bailey, who reports that on May 24, 2000, he was working for the Village of Marshall Police Department, dressed in full police uniform, and operating an unmarked squad car. On that date, doing follow-up to Officer Viken's report, he spoke to [redacted] who identified herself verbally, and when asked why the police were here, she stated, "I know why you're here, I told my Mom something that happened on my birthday and she told you. I told her but I didn't want her telling anyone else." When asked if she would tell him about what happened on her birthday, she cried and cried and answered, "I don't want to talk about it." They asked her when it happened. She made statements that included the following. She said her birthday was Sunday, April 2nd. They asked her where she was beforehand, and she told them she was walking home from her friend's house at about 10:00 at night. She reiterated that she had told her Mom about it but didn't want to talk about it again. She then gave Lieutenant Bailey a description of the vehicle and the offender, describing him to be about 35 years old, looking like a businessman, in khaki pants, a white short-sleeved shirt with no tie, clean cut, a bigger guy, about 6'2" tall, a big build, blondish-brown hair longer than Bailey's, no facial hair, and, "not the kind of guy to have tattoos." She stated that she was walking by the railroad tracks north on Hubbell Street toward her house in Marshall, Dane County, Wisconsin, and then, "A car pulled up and a guy rolled down the window and asked directions to Sun Prairie." She told Lieutenant Bailey, "I started giving him directions, and he got out of his car and walked around to me and handed me a piece of paper and asked me to write down the directions. He didn't have a pen and told me that there was a pen on the back seat and asked me to get it. When I reached in to get the pen, he pushed me in the car and closed the door. I tried opening the door but it wouldn't open. The guy got in the car and drove through the lumber yard parking into Converse Park." She told Lieutenant Bailey, "He did what he wanted to do and grabbed me by the hair and threw me out of the car." She stated that during this time what he said was, "He has no problem killing me because he is not going to jail and people don't go to jail very long for rape, and 'I'll kill you when I get out if you tell anyone'."

[redacted] told Lieutenant Bailey that when she reached into the back seat of the car to get the pen and was pushed into the car, he closed the door behind her. She tried to open the door but it wouldn't open. She tried to open the other back seat door but it wouldn't open. He got in the car. She said, "What are you doing?" He didn't answer. He then drove through the parking lot of the lumber yard into the parking lot of Converse Park and stopped and parked near a sand pile. He got out of the car and opened the back door. She tried to push her way past him. He punched her in the face with a closed fist. She stated, "Let me go" but he didn't answer. He was standing at the back door. He pulled down his pants and underwear to his ankles. He reached in and grabbed both her hands together with his left hand and held her hands against her neck. He then unbuttoned her pants. She used her left foot and kicked him on the front of his right thigh. He then punched her in the chest, knocking the wind out of

her, and then pulling her closer to him he, she said, "did what he wanted to do." [redacted] cried to Lieutenant Bailey, "I was crying so much." [redacted] told Lieutenant Bailey that he had an erection but she had no idea whether or not he used a condom, "I was crying so much." She stated that he entered his penis into her vagina. She stated she had no idea whether or not he ejaculated. Lieutenant Bailey asked her if she would be able to identify the person, and she answered, "I don't think I'll be forgetting him anytime soon." When asked if he took off her shirt, she answered, "No, he took his right hand and went up under my shirt and over my bra and rested his hand on the base of my breast." She stated that when this happened, it was "in the back seat of the car, the guy left the door open, there was a dome light or inside light on." She stated when he was done sexually assaulting her, "he pulled me up by my hair and out of the car and sat me down on the ground next to the car and then left." When asked if she saw which way he went, she stated, "No, I was crying so hard I didn't see." When asked if he said anything after he threw her out of the car, she told Lieutenant Bailey, "He said people don't go to jail very long for rape and he'd kill me if I told anyone." When asked what her injuries were, she stated that her face and chest hurt afterwards from being punched.

Lieutenant Bailey spoke with [redacted] again on May 26, 2004. At that time he asked her whether or not she remembered anything else that had happened at that time or that she had failed to tell him at that time, and [redacted] stated that she had not previously told anyone that when she was sexually assaulted, the offender tore her vagina and that she had vaginal bleeding. She finally went to the doctor several weeks after Lieutenant Bailey had interviewed her back in the year 2000. Lieutenant Bailey asked her if she would yet still be able to identify the person who had sexually assaulted her, and she stated yes. She stated she would agree to view a line-up. Lieutenant Bailey reports that on May 26, 2004, at approximately 2:30 p.m., he made phone contact with your complainant in regards to this four-year-old unsolved sexual assault because the suspect's description, he believed, somewhat matched the description of the person known as the "mall rapist."

On June 1, 2004, Lieutenant Bailey reports that at approximately 2:30 p.m., he received the six-pack photo array consisting of Dane County Jail mug photos kept in the regular course of business at the Dane County Sheriff's Office. He received these from your complainant who recognized the description of this attacker matched that of the mall rapist. The photo array included a photograph of the defendant herein, James D. Perry, and five other men of similar age and physical appearance. At 3:20 he met with [redacted] and she viewed the photographs in a sequential photo array style, opening each envelope containing a photograph, one at a time. When she got to the image in the fifth envelope, which is that of the defendant, she got a concerned look on her face and stated, "What if I'm 85 percent sure?" She then started crying. He asked her if that was the person who had sexually assaulted her, and she said, "I'm 85 percent sure that's the person who sexually assaulted me." She initialed and dated the photograph. He asked her why she was 85 percent sure rather than 100 percent sure, and she stated that it has been four years and the only other reason was, "his hair is longer in the picture." [redacted] then started crying and said to Lieutenant Bailey, "Would you have given him directions?" She looked at the sixth envelope and stated that she didn't recognize that person. [redacted] told Lieutenant Bailey that she did not give the defendant consent to sexually assault or batter her or take her from one place to another without her consent.

EVIDENCE OF OTHER ACTS

Sears Assault - Madison

Your complainant has reviewed police reports of Madison Police Officer Scott Favour, who reports that on March 10, 1999, at 7:03 p.m., he was dispatched to Sears at the East Towne Mall in Madison, Dane County, Wisconsin, regarding a fourth degree sexual assault which had occurred there. Upon his arrival he made contact with [redacted] who identified herself verbally and who stated she was then employed as a clerk at the Sears store in the East Towne Mall in the junior's department. She stated that on March 10, 1999, at around 6:30 p.m., her assaulter approached her and asked her to help pick out a "sexy outfit" for his 21-year-old girlfriend. [redacted] told Officer Favour that she thought that was a strange request. She stated that the man then followed her around the junior's department and asked her several questions which she thought were strange. He asked her if she was married; he asked her if she always wore short skirts; he said to her that he wished his girlfriend would wear "short stuff" more often. [redacted] stated that the man remained in the junior's department for quite some time -- between 15 and 30 minutes. He continued wondering around and periodically came up to her and asked her questions. [redacted] stated that she then became distracted by another customer whom she was helping when the same man who had been asking her these questions came up behind her and fondled her buttocks using an open hand. She stated that the man made contact with her right buttocks and felt her buttocks through her clothing. She stated that she did not give consent to him touching her in any way and that was why she had contacted the police. [redacted] told Officer Favour that she would be able to identify him both in person and by way of photograph. Sears loss prevention video camera system produced a videotape of the offender. Officer John Summers and Officer Favour then watched the afore-described videotape and observed that the offender had an erection as he was entering the junior's department of the Sears store. [redacted] pointed out to Officer Summers that the person they were observing on the videotape was the same person who had assaulted her a short time earlier.

On April 30, 2004, your complainant met with [redacted] at her home with Detective Vicki Anderson, who assisted your complainant in administering a photo array of six photos made and kept in the regular course of the business of the Dane County Sheriff's Office to [redacted] to see if she could effect an identification. Detective Anderson shuffled the six Dane County Jail mug photos and randomly selected eight file folders. Your complainant advised [redacted] of how the identifications would proceed, sequentially. After [redacted] had viewed all of the photographs, she then returned to look at the photo in folder No. 4. She viewed photo No. 4 and stated, "This is the one. This is him." She signed and dated the picture, which was that of the defendant herein, James D. Perry.

Once Upon A Child Assault - Madison

Your complainant has reviewed the police reports of Madison Police Officer Scott Favour, who reports that on March 10, 1999, he responded to a call of a fourth degree sexual assault at a store called Once Upon A Child at the strip mall at 1625 Thierer Road, in Madison, Dane County, Wisconsin. Your complainant is aware that 1625 Thierer Road is located adjacent to East Towne Mall, so that a person could walk in minutes fro the Sears at East Towne to Once Upon A Child. Officer Favour spoke with [redacted] who stated that on March 10, 1999, at approximately 7:40 p.m., the person who assaulted her entered her

store and stated he was looking for clothing for a 2-1/2 year old. She began showing him clothing in the sizes 2 to 3 range and then the man began wandering around the store looking at various things but obviously just following her. She was placing clothes on the rack in the girl's size 10 area and he was standing next to her admiring a stuffed animal toy. When she turned her back on him, he then fondled her buttocks with his open left hand. She described it as "brief rubbing motion" but "clearly intentional." She stated the contact was made through her clothing, and when she turned around to look at him he placed his index finger across his lip and audibly said, "Shhhh." He then abruptly walked out of the store. She did not consent to him touching her in any way. Officer Favour further reports that [redacted] was somewhat similar in appearance to the victim in the previously reported case, [redacted] Both had longer brownish hair and were young white females. Each of them was 5'6 tall with slender builds, both wearing skirts at the time of the assaults.

Your complainant met with [redacted] on April 28, 2004 for purposes of showing her a photo array of photos made and kept in the regular course of the business of the Dane County Sheriff's Office. Detective Linsmeier assisted your complainant in administering the photo array by shuffling the six Dane County Jail mug shots and placing them randomly in eight file folders with folders 7 and 8 remaining empty. Your complainant instructed [redacted] as to the photo lineup instructions for the sequential lineup, and Detective Linsmeier then administered the photos. [redacted] viewed all the photos before promptly pointing to photo No. 4. She viewed photograph No. 4 at length and then stated, "This guy looks familiar from the attack on me." She signed her name next to the photo of the defendant, photo No. 4, and stated, "He has the same hair and face as the guy who attacked me." Your complainant noted photo No. 4 was that of James D. Perry.

Carol's II Laundromat - Monona

Your complainant has reviewed the police reports of City of Monona Detective Jack Jasensky, who reports that on February 2, 2000, at 5:21 p.m., he responded to Carol's II Laundromat, 5920 Monona Drive, in Monona, Dane County, Wisconsin, in reference to a complainant at that location stating that an offender had just exposed himself to her. When Detective Jasensky arrived at that location, he spoke with [redacted] who identified herself verbally and who stated that she would able to recognize the person who had offended against her in part because he previously was operating a golden colored beat-up type vehicle with no rear license plate in front of the store. She was in the laundromat washing her clothes when she observed him drive by in a golden colored type vehicle, beat-up condition, with no rear plate pass in front of Carol's II Laundromat. She observed that vehicle turn around in the Pizza Oven parking lot located just to the south of Carol's II Laundromat in the afore-described strip mall and then proceed slowly past the front of Carol's II Laundromat in a northerly direction. A short time later, the person who offended against her walked into Carol's II Laundromat through the front door on the west side and went into the only restroom of the laundromat. After approximately two to three minutes he then walked out of the bathroom and walked towards the rear of Carol's II Laundromat and opened the door to the outside. He never stepped outside the laundromat but did look around outside the door for a short time. He then closed the door and walked back into the laundromat around the north side of the washing machines. She was still the only person in the laundromat other than the person who offended against her. He then came around the front of the store and faced [redacted] estimated that she was about 15 feet away from him when she looked down and observed that he had his pants crotch zipper open and appeared to be wearing no underwear. He placed his hand over his penis and exposed it to

[redacted] He then began to stroke his penis. She asked him what he was doing. She stated she was going to call the police. She ran towards the rear of the laundromat and opened the back door and called for her mother, who lives behind the laundromat. At this time, he left the laundromat through the front doors and ran towards the north. She stated that she did not believe his penis was erect but that he did massage and stroke it in front of her for several seconds. She was sure it was the same person who had driven by the laundromat initially in the golden colored beat-up type vehicle. It appeared clear that he was already intending to expose himself to her or worse prior to coming into the laundromat in the mall in that by initially driving by the laundromat, he would be able to tell whether or not she was the only person in the business at the time because of the expanse of the windows at the front of the business, which allow a person to see into the business to see all of the space of the business and all who were present.

Monona Sergeant Richard O'Connor reports that on April 28, 2004, he received a photo array from your complainant who recognized that the description of this attacker matched that of the mall rapist. The photo array consisting of six photographs made and kept in the regular course of the business of the Dane County Sheriff's Office, including that of James D. Perry.

Detective Sergeant O'Connor reports that he contacted [redacted] and met with her at her place of employment. He asked her to view the photo array, and when she did so after three to five seconds she pointed to No. 5, the defendant herein, James Perry, and stated, "This face looks very familiar. I think that could be him. My god, that is him!" She signed her name just under photo No. 5 of the defendant James Perry, identifying him as being the person who had exposed himself to her on February 2, 2000.

Goodwill Industries - Monona

Your complainant has reviewed the police reports of Monona Police Sergeant Brian Redman, who reports that on February 24, 2000, at 8:21 p.m., he was dispatched Goodwill Industries at 2501 Royal Avenue, Monona, Dane County, Wisconsin, in regards to a female being assaulted in the parking lot area and currently inside the store waiting for the police to arrive. Upon his arrival at that location, Sergeant Redman spoke with [redacted] who stated that she had gone to Goodwill to purchase some clothing items. As she was shopping in the store in the mall she observed the person who subsequently assaulted her walking around inside the store. She stated that he was inside the store the entire time she was in the store. She went to the cash register, paid for her merchandise, and then exited through the front door. Her vehicle was parked directly in front of the store near the front door area. She went to the rear of her vehicle to place the items she had purchased in her trunk. As she was placing the merchandise in the trunk, she noticed that the same man who had been in the store the whole time she had been there now approach her from her left side. He would have been coming from an easterly direction. She stated that he walked directly behind her and she felt something touch her buttocks area. She stated that she believed it was his left hand brushing across the left side of her buttocks. He then grabbed her right buttocks area. He then continued to walk towards the entrance of Goodwill and then turned around towards her and stated, "Hey, that felt good, baby." He then started walking towards her direction once again and was fumbling with the front of his pants. She believed he was trying to unzip his pants. She picked up her cell phone and began to talk into it as if she was talking to someone. He then ran in an eastbound direction from her location. [redacted] told Sergeant Redman she was quite shaken up over this entire incident. She did not give him consent to have sexual contact with her in any way.

Detective Sergeant O'Connor of the Monona Police Department reports that on April 28, 2004, he received a photo array of six photographs from your complainant from photos made and kept in the regular course of business of the Dane County Sheriff's Office, and that on May 6, 2004, he contacted [redacted] for the purposes of having her view the photo array. The defendant herein, James Perry, was photograph No. 5 in the photo array and when she got to photograph No. 5 in the photo array, [redacted] stated that it was what the suspect looks like because she got a very unsettled feeling when she observed the person in photograph No. 5, and signed that photograph attesting to her identification of him as the person who assaulted her in this case.

Roscoe, Illinois

Further, your complainant has reviewed the police reports of Roscoe, Illinois Police Officer T. Bubrick, who reports that on April 10, 2000, she spoke to [redacted] date of birth [redacted] who reported that on April 10, 2000 at 9:15 p.m., she was leaving Sun Seekers in that city after a 15-minute tanning session, when a man came out of the bushes and started to approach her as she was walking to her car. He came up to her and put his hand over her mouth and put a knife, a carpenter's knife, to her throat and told her that if she didn't do what he told her to do he was going to cut her pretty little throat. He put her in his truck, an older model possibly '89 to '95 model, and put her on the floorboard and held her down with his hand. He asked her how old she was. She said 16. He repeatedly told her that she was a sweet little girl and that he just wanted to try "it." He drove a short distance and stopped. He had her pull her pants and underwear down to her knees. He lifted her skirt up and pulled up her bra. He sucked on her breasts. He put his mouth on her vagina and licked her there. He made her put his penis into her mouth and told her to suck on it. He then got on top of her and put his penis into her vagina. He then had her put her pants back on and told her to get out of the truck. She ran back into Sun Seekers and called the police. As a result of the assault, she had a tear-like abrasion on the inner wall of her vagina and physical marks on her left upper arm and lower left side of her stomach. She said that the person who assaulted her smelled like cigarette smoke and had a gold ring, like a wedding band, on his left finger. His hands were rough like he worked outside and the ring was worn and nicked. Officer Bubrick further reports that a sexual assault examination was conducted and that oral and vaginal smears and swabs were taken from the victim recovered and preserved for possible comparison at the later time.

Your complainant has reviewed the police reports of Winnebago County Sheriff's Detective Larry Marino, who reports that on April 27, 2004, he received a photo array of six photographs from your complainant, photos made and kept in the regular course of the business of the Dane County Sheriff's Office and regularly relied upon by law enforcement officers throughout the State of Wisconsin. Marino further reports that on Wednesday, May 12, 2004, at 10:15 a.m., he met with [redacted], who was given the photo array. She viewed the photos. When she was done she identified photograph No. 5 as the person who had sexually assaulted her. Photograph No. 5 is the photograph of the defendant herein, James D. Perry.

Cash Store - Monona

Your complainant also reviewed the police reports of Monona Police Officer Shawn Fogeltanz, who reports that on January 16, 2001, he was on duty in full police uniform while working 3:00 p.m. to 11:00 p.m. for the Monona Police Department assigned to a marked squad car. At 6:35 p.m. he was dispatched to the Tobacco Outlet located in the shopping mall at 2401-B West Broadway, in the City of Monona, Dane County, Wisconsin, regarding

a female at that address wanting to report an exposure that had just occurred in the parking lot near that business. At that Tobacco Outlet in the strip mall at that location, he spoke with [redacted] who reported that she is the manager of the Cash Store at 2401-A West Broadway, which is directly next to the Tobacco Outlet in the strip mall. [redacted] said that at 6:00 p.m. that day she had closed the store. She exited the store. She began to walk towards her truck parked in the parking lot in the third stall in the row of stalls directly in front of the Cash Store. As she began to walk towards her truck, the person who offended against her walked towards her from the east. The offender asked her if she knew where something was. She could not understand what he was asking for. She told him no and continued walking towards her truck. He followed her to her truck. When she neared the driver's side door of her truck, she turned around and that same person was standing near her. He said, "You've got pretty blue eyes." He then pointed towards his groin. She glanced down and saw that he was exposing himself to her. The zipper on his pants was open and his penis was exposed to her. He didn't yet have an erection. He said, "Why don't you go back in the store with me and suck my dick?" She then ran towards the store. He quickly walked towards a vehicle parked in the parking lot of Shopko nearby in the strip mall. She was very disturbed by his actions and called the police immediately. She did not give consent or permission to be exposed to, nor to his lewd and lascivious comments. She was very concerned for her safety and very concerned about returning to work the next day.

Detective Gary Buss of the Monona Police Department reports conducting additional investigation of this matter on January 17, 2001, and during his conversation with [redacted] at that time, she told him that when she saw the defendant's penis he was lifting up his sweatshirt with his right hand still in his pocket and exposing himself to her. The zipper in his blue jeans pants was open and he was holding his penis with his left hand and holding it out at her so that she could see it. He then leaned towards her and said, "Why don't you go back in that store and suck my dick?" He then raised his right hand towards her, and she then yelled, "Oh my god" and turned around from him and started to run between the vehicles heading east away from him, at which time she entered the Tobacco Outlet and made contact with the clerk at that location and yelled, "Call 911 now. Some guy just tried to do something to me in the parking lot!"

Detective Jack Jasensky of the Monona Police Department reports that on April 28, 2004, he received a photo array from your complainant, who recognized that the physical description of this attacker matched that of the mall rapist. The photo array was composed of six photographs made and kept in the regular course of the business of the Dane County Sheriff's Office and regularly relied upon by law enforcement officers throughout the State of Wisconsin to be truthful, accurate, and reliable. Detective Jasensky reports that he contacted [redacted] to view the photographs on April 28, 2004, and on that date he showed her the photo array and she viewed the photo array for just a few seconds, centered on photograph No. 5 and stated that the photograph of the subject in No. 5 looks like the person in this case that exposed himself to her. She initial and dated the photograph, identifying the defendant herein, photograph No. 5, James D. Perry, as being the person who had exposed himself to her on January 16, 2001.

Laundromat - Janesville

Your complainant has reviewed the police reports of Janesville, Wisconsin Police Officer Rodney Hirsch, who reports that on July 2, 2001, he was dispatched to the Kohl's grocery store at 2822 East Milwaukee Street in that city to make contact with a female subject there who had been the victim of a sexual assault that occurred at 2604 East Milwaukee Street.

The officer responded immediately across the street and made contact with the victim, [redacted], who stated that she was in the laundromat at 2604 East Milwaukee Street in Janesville, Wisconsin towards the rear portion of it, drying some clothes, when she observed the person who offended against her come into the laundromat through the back door and go into the bathroom. At the rear entrance to the laundromat on the southwest side of the laundromat. He exited the bathroom. She went back towards the back of the laundromat to sit down and read her paper. The person who offended against her then grabbed her by the back of her neck and held an object to her throat and said that she should not struggle with him or he would "slice her throat." He then took her into the bathroom at which time he locked the door behind them. He pushed her up against the wall in the bathroom and told her not to look at him and kept repeating that he was a black man but when she was able to turn around and look at him, she did observe that he had the features of a white man and not a black man. He then took her clothes off and vaginally penetrated her with his fingers from the rear while holding her up against the wall with her back to him. She struggled against him. He said he had had enough and he left. The next day, in conversation with Detective James Martin, she said that after the defendant forcibly pushed her into the bathroom and placed her hands up to the wall while he continued to hold the sharp object to her neck, he kept telling her to, "Keep your hands on the wall." She stated she did not completely do this and turned around three to four times and got a look at him. She stated that he then began to try push her face into the wall in an attempt to keep her from turning around, and he said to her, "I'll kill you, don't struggle." Prior to going into the bathroom, the defendant said to her, "Wait five minutes before you leave." She said that while she had her hands placed against the wall he fondled her breasts through her shirt and then finally pulled off her shirt and bra. He then continued to fondle her naked breasts, squeezing them, kissing her upper back area and saying, "Oh, baby." After he fondled her unclothed breasts, he then moved his hand down to her vaginal area and inserted his fingers into her vagina and then he kneeled down in an attempt to have oral sex with her, touching his tongue to her vagina. When she turned around while he was doing that, he kneed her in the abdomen and this blow, along with him pushing her, caused her to fall to the floor and she fell on her right side up against the door. She started to scream. He immediately put his hand over her mouth and told her that he was leaving. He also told her that if she contacted the police he would kill her. He then ran out the door very quickly.

At the request of your complainant who recognized that the physical description of this attacker matched that of the mall rapist, Detective Bruce Frey of the Madison Police Department joined in this investigation due to the connections between the Janesville offense and the offenses committed in Madison occurring in a similar manner in malls. Detective Frey reports that during his investigation he found that the laundromat at 2604 East Milwaukee Street in Janesville, Wisconsin is located in a strip mall with numerous other businesses, with a large parking lot, and East Milwaukee Street with two lanes of traffic with a high volume of traffic in the commercial district. In front of the Quick Wash was a vacant business next to it and further down, a restaurant.

Detective James Martin reports that on April 28, 2004, he received a six-photograph photo array from your complainant consisting of photographs of white male subjects made and kept in the regular course of the business of the Dane County Sheriff's Office and regularly relied upon by law enforcement officers throughout the State of Wisconsin to be truthful, accurate, and reliable. The defendant, James Daryl Perry, was suspect No. 5 in the six person photo array. On May 3, 2004, Detective Martin met with [redacted] and explained to her the photo array procedure and had her look through the photographs

carefully before making a decision or saying anything about who, if anyone, the person was. [redacted] pointed directly at No. 5, that being the photograph of the defendant herein, James Perry. She stated that she was about 75 percent sure that No. 5 was the person who sexually assaulted her. After looking at it a while more she stated, "I would say he's the one," again pointing to photograph No. 5, the photograph of the defendant herein.

WISCONSIN STATE CRIME LAB RESULTS

Your complainant reviewed a report of [redacted], dated June 12, 2000, Serology and DNA Analyst at the Wisconsin State Crime Laboratory, whom your complainant knows to be an expert in the Wisconsin State Crime Laboratory on serology and DNA and upon whose opinion law enforcement officers throughout the state of Wisconsin regularly rely. The report from Analyst Herzberg, regarding the assault of [redacted] stated she analyzed the three cigarette filters recovered from the front landscaping nearest the front door of the Learning Store. The DNA was isolated from cigarette filter item B, analyzed using the PCR and typed for short tandem repeat genetic markers utilizing the Perkin-Elmer AmpF1STR Profiler Plus and COfiler amplification kits. A DNA profile was obtained from the evidence of the item B cigarette filter, and entered into the National CODIS data bank, and there was no match to any recorded DNA data at that time.

Upon learning of the Roscoe, Illinois mall sexual assault, your complainant recognized that the physical description of the Roscoe, Illinois perpetrator matched that of the mall rapist and made arrangements for the Winnebago County Sheriff's Department to convey their sexual assault kit from [redacted] to the Madison, Wisconsin State Crime Lab.

On May 30, 2000, Hertzberg analyzed the swabs and smears taken in the afore-enumerated case from [redacted] of Roscoe, Illinois, and semen was identified on the vaginal cervical swab and panties recovered from the victim in that case. DNA was extracted from the vaginal cervical swabs, one semen stain from the panties, and the blood standard reportedly from the victim. This DNA was analyzed using PCR and typed for short tandem repeat genetic markers utilizing the Perkin-Elmer AmpF1STR Profiler Plus and COfiler amplification kits. The DNA profiles were developed and listed. Based upon the DNA profiles developed, Expert Hertzberg concluded that [redacted] is eliminated as possible source of the DNA profiles developed from the sperm fractions of the vaginal/cervical swabs and panties.

Your complainant requested that Serologist Herzberg compare the DNA profile she obtained from the Roscoe, Illinois evidence, to the DNA profile obtained from the [redacted] case. The DNA profiles were then listed and compared by Expert Hertzberg. She concluded that the DNA profile developed from cigarette filter item B obtained in the [redacted] mall case, as compared to the DNA profile developed from the vaginal cervical swabs submitted as evidence in laboratory case number M00-947 by the Winnebago County Sheriff's Office in the Roscoe Illinois mall case of [redacted]., were a positive match The DNA taken from item B, the cigarette butt from outside of the Learning Shop, is DNA of the same person who left his sperm fractions in the vagina of [redacted] in Roscoe, Illinois.

In February 2004, your complainant developed James Perry as a possible suspect in these investigations due to his physical build and descriptors, as well as the fact that he owned a 1990 red Ford F150 pickup truck as described by victims and witnesses of unsolved attacks. Your complainant requested the assistance of the FBI in obtaining Perry's DNA for comparison with the unknown subject DNA identified in these investigations.

On 03/02/2004, FBI Special Agent Steve Paulson and Stoughton Police Detective Marsha Clark obtained buccal swab DNA standards from James Perry. These were conveyed to the Wisconsin State Crime Laboratory by Detective Marsha Clark. On 03/02/2004, your complainant contacted the crime lab and requested Perry's DNA standards submitted by Detective Clark be processed for a DNA profile, and compared to the unknown DNA profile recovered in these investigations.

Your complainant has reviewed the report of the Wisconsin State Crime Laboratory DNA Analyst James E. Andreas of April 5, 2004, wherein he analyzed a buccal cell standard obtained from the defendant herein, James D. Perry, for DNA using the PCR method and typed for fifteen short tandem repeat genetic markers and the gender marker amelogenin using the Promega PowerPlex 16 amplification kit. The DNA profile obtained for the defendant herein, James D. Perry, was then listed in the table in the report so that the alleles could be compared.

Further, your complainant, on April 5, 2004, spoke to Marie Varriale of the Wisconsin State Crime Laboratory. She is the supervisor of all expert DNA serological analysts at the Wisconsin State Crime Laboratory including but not limited to Jill Hertzberg and Jim Andreas. Varriale advised your complainant that Andreas had conducted the DNA profile analysis of the defendant herein, James Perry's DNA and compared it with the samples submitted in the Roscoe [redacted] case and the Madison [redacted] case as described above. Maria Varriale verified to your complainant that Perry's DNA was a positive match with the two samples submitted from the crime scenes of the [redacted] and [redacted] cases. She stated the comparison was a one-on-one comparison. Your complainant notes that each of the two DNA profiles exactly matched Perry's DNA standard.

VITA

Joseph York Thomas

Candidate for the Degree of

Doctor of Philosophy

Thesis: REAL-WORLD, HIGH-STAKES DECEPTIVE SPEECH: THEORETICAL VALIDATION AND AN EXAMINATION OF ITS POTENTIAL FOR DETECTION AUTOMATION

Major Field: Business Administration, Management Information Systems

Education:

PhD, Business Admin, Mgmt Info Sys: Oklahoma State University, OK (Dec 2014)
MA, Information Systems Management (w/ Honors): Webster University, MO (2003)
BS, Computer Science (w/ Teacher Certification): Texas State University, TX (1993)

Experience:

Director IT Research, May 2013 – Present: Responsible for AF-level rapid research for the Institute for Information Technology Applications in direct support of operational missions.
Director of Technology & Instructor of Management, Oct 2011 – May 2013: Course Director responsible for content, policy, assessment and instruction for 280 cadets and 6 faculty in core management course.
PhD Student, Jan 2008 – present: MSIS, Oklahoma State University, Stillwater, OK.
Commander, Apr 2005 – Dec 2007: Responsible for effective development, protection, and utilization of more than 2,500 military and civilian personnel executing 140+ programs valued at more than $3.1B.
Director of Operations, AF Network Ops & Security Center, Jul 2004 – Apr 2005: Led 47 personnel responsible for all operations affecting $7M network control operations center.

Professional Memberships: Association for Information Systems (AIS), 2014-present

Publications:

Thomas, J. & Biros, D., (2014). "Theoretical Validation of IDT in Real-World, High-Stakes Deceptive Speech". *HICSS*, *2014*, 48th Hawaii International Conference on System Science 2014. In Press.

Thomas, J., Chongwatpol, J., Pengnate, F., & Hass, M., (2011). "Data Mining in Higher Education: University Student Declaration of Major". *MWAIS 2011 Proceedings.* Paper 15.

Thomas, J. & Biros, D., (2011). "A Conceptual Model of Real World, High Stakes Deception Detection", *HICSS*, 2011, 47th Hawaii International Conference on System Sciences 2011, pp. 1-10.

Hass, M., Nichols, J., Biros, D., Weiser, M., Burkman, J., & Thomas, J., (2009). "Motivating Knowledge Sharing in Diverse Organizational Contexts: An Argument for Reopening the Intrinsic vs. Extrinsic Debate". *AMCIS 2009 Proceedings.* Paper 285.