

Linguistic Cues to Deception

Caroline Amie Connell

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in  
partial fulfillment of the requirements for the degree of

Master of Arts  
In  
Communication

Douglas F. Cannon  
Robert G. Magee  
John C. Tedesco

April 27, 2012  
Blacksburg, VA

Keywords: Deception detection, natural language processing, linguistic cues to deception

# Linguistic Cues to Deception

Caroline Amie Connell

## ABSTRACT

This study replicated a common experiment, the Desert Survival Problem, and attempted to add data to the body of knowledge for deception cues. Participants wrote truthful and deceptive essays arguing why items salvaged from the wreckage were useful for survival. Cues to deception considered here fit into four categories: those caused by a deceivers' negative emotion, verbal immediacy, those linked to a deceiver's attempt to appear truthful, and those resulting from deceivers' high cognitive load. Cues caused by a deceiver's negative emotions were mostly absent in the results, although deceivers did use fewer first-person pronouns than truth tellers. That indicated deceivers were less willing to take ownership of their statements. Cues because of deceivers' attempts to appear truthful were present. Deceivers used more words and more exact language than truth tellers. That showed an attempt to appear truthful. Deceivers' language was simpler than that of truth tellers, which indicated a higher cognitive load. Future research should include manipulation checks on motivation and emotion, which are tied to cue display. The type of cue displayed, be it emotional leakage, verbal immediacy, attempts to appear truthful or cognitive load, might be associated with particular deception tasks. Future research, including meta-analyses, should attempt to determine which deception tasks produce which cue type.

# Table of Contents

Table of Contents .....	iii
List of Tables .....	vi
Chapter 1: Justification .....	1
Chapter 2: Introduction .....	4
Chapter 3: Literature Review .....	6
Theories on Cues to Deception .....	6
Causes of Cues to Deception .....	6
Theories on Verbal Cues to Deception.....	9
Theories of Deception in Computer-Mediated Communication .....	11
Automation of Deception Cues in Computer-Mediated Communication.....	14
Review of Verbal Cues to Deception and Hypotheses.....	15
Categories of Verbal Cues to Deception .....	15
Verbosity .....	17
Verbal Immediacy .....	19
Emotion Words.....	24
Linguistic Complexity.....	26
Qualitative Differences between Real and Imagined Events .....	30
Summary of Hypotheses and Research Questions .....	34
Chapter 4: Method .....	36
Participants.....	36
Measures .....	37
Demographics.....	37
Personality .....	37
Experimental Manipulation.....	37
Awareness and Contact Information .....	38
Counterbalancing.....	38
Dependent Measures .....	39
Procedure .....	42

Chapter 5: Results .....	44
Data Compilation .....	44
Descriptive Statistics .....	44
Hypothesis Testing .....	45
H1: Verbosity .....	46
H2: Verbal Immediacy .....	46
RQ1: Verbal Immediacy .....	46
H3: Emotion Words.....	47
H4: Linguistic Complexity .....	48
Post-Hoc Analyses .....	48
H1: Verbosity .....	49
H2: Verbal Immediacy .....	49
RQ1: Verbal Immediacy .....	50
H3: Emotion Words.....	51
H4: Linguistic Complexity .....	51
Chapter 6: Discussion .....	52
Participant Characteristics.....	52
Discussion .....	52
H1: Verbosity .....	52
H2: Verbal Immediacy .....	53
RQ1: Verbal Immediacy .....	54
H3: Emotion Words.....	55
H4: Linguistic Complexity .....	57
Limitations .....	57
Chapter 7: Conclusion.....	60
References.....	63
Appendices.....	73
Appendix A: Institutional Board Approval Letters.....	73
Appendix A.1: Approval Letter.....	73
Appendix A.2: Amendment Approval Letter.....	74

Appendix B: Experimental Measures .....	75
(Introduction and Instructions).....	75
Part 1: Basics .....	75
Part 2: (Personality).....	75
Part 3(A): Essay (Desert Survival Problem, TRUTH Condition) .....	77
Part 3(B): Essay (Desert Survival Problem, DECEPTION Condition).....	78
Part 4: Ending Questions.....	79
Appendix C: LIWC 2007 Output Variable Information .....	80
Appendix D: Experiment Announcement, Recruitment and Consent Form, and Debriefing Statement.....	84
Appendix D.1: Announcement.....	84
Appendix D.2: Recruitment and Consent Form.....	84
Appendix D.3: Debriefing Statement.....	86

## List of Tables

Table 1: Verbosity.....	17
Table 2: Verbal Immediacy .....	20
Table 3: Emotion Words.....	24
Table 4: Linguistic Complexity .....	27
Table 5: Qualitative Differences between Real and Imagined Events .....	31
Table 6: Operationalization of Variables .....	39
Table 7: Results for All Essays.....	45
Table 8: Results for Longer Essays .....	49

## Chapter 1: Justification

Deception is pervasive in social interaction. College students tell on average two lies a day, and community members tell one (DePaulo, Kirkendol, Kashy, Wyer, & Epstein, 1996). Deception has damaged the relationships and finances of a great many people. An accurate and usable deception-detection scheme could help countless individuals avoid heartache and save money. This project worked toward that goal by adding to the body of deception-detection data.

Deception is the intent to convey a belief to a receiver that the sender considers false. (Zuckerman, DePaulo, & Rosenthal, 1981). Modern deception-detection research is rooted in the work of Ekman and Friesen (Ekman & Friesen, 1969, 1971, 1974). They discovered that facial expressions could signal deception and that human facial expressions and emotions were universal across cultures (Ekman & Friesen, 1971; Ekman et al., 1987). Expressions could be either intended, or “leaked,” because of the deceiver’s underlying emotion (Ekman, Friesen, & O’Sullivan, 1988). In addition, cognitive load and attempts to appear truthful could cause deception cues (Zuckerman et al., 1981).

Deception-detection schemes measure a vast array of behaviors, including facial expressions, vocal pitch, gestures, and heart rate. These measures attempt to identify emotion, cognitive load, or a person’s attempts to appear truthful. Average deception-detection accuracy is about 54% (DePaulo et al., 2003). Some systems can correctly categorize 78% of truthful and deceptive messages, but they require extensive training and are not feasible for lay use (Vrij, Edward, Roberts, & Bull, 2000).

Most cues to deception require the deceiver to be present or on video, but computer-mediated communication is more and more common. Burgoon (Burgoon, Blair, Qin, &

Nunamaker, 2003; Burgoon & Nunamaker, 2004a, 2004b) has tackled the challenge of computer-mediated deception. Burgoon and Zhou (Zhou, Burgoon, Nunamaker, & Twitchell, 2004a; Zhou, Burgoon, & Twitchell, 2003) have tried to identify how deceptive messages change over time. Their work has contributed to interpersonal deception theory. It states that interpersonal deception, like interpersonal communication, is dynamic and adaptive (Buller & Burgoon, 1994; Buller, Burgoon, Buslig, & Roiger, 1996). Nonverbal cues are unavailable in computer-mediated communication, so Burgoon and Zhou's methodology examines verbal cues. Verbal cues occur for the same reasons as nonverbal cues – emotion on a liar's part, cognitive load, or attempts to appear truthful. In researching interpersonal deception theory, Burgoon et al. and Zhou et al. have used natural-language processing software to identify differences in truth tellers' and deceiver's statements (Burgoon et al., 2003; Zhou et al., 2003).

Zhou and Burgoon (2003, 2004a, 2004b) used difficult-to-run software to detect verbal cues to deception, but user-friendly programs are available. For example, Linguistic Inquiry Word Count (LIWC, pronounced "Luke") categorizes words like first-person pronouns or emotion words (Pennebaker, Francis, & Booth, 2001). The user can compare the percentage of such words in deceptive and truthful statements. Laypeople can use LIWC.

Not enough data exist on verbal cues to deception. Results regarding some cues are conflicting. For instance, few dispute that deceivers use more negative-emotion words than truth tellers, probably because of guilt or fear (DePaulo et al., 2003; Pennebaker, Newman, Berry, & Richards, 2003; Zhou et al., 2004a; Zhou et al., 2003; Zhou & Zhang, 2008). Research results are split, though, on whether liars use more or fewer words than truth tellers (Burgoon et al., 2003; Burgoon & Qin, 2006; DePaulo et al., 2003; Hancock, Curry, Goorha, & Woodworth, 2008; Zhou et al., 2004a; Zhou et al., 2003; Zhou & Sung, 2008; Zhou & Zhang, 2008). Guilt may

cause a deceiver to distance herself from her statement and utter fewer words. Alternately, desire to appear forthcoming may cause a deceiver to use more words. More data are needed regarding how word count and other cues predict deception for different individuals and in different contexts.

This study aimed to add to the body of data on verbal cues to deception. Participants made truthful and deceptive statements via computer-mediated communication. The study followed Zhou, Burgoon et al's methodology (Zhou et al., 2004a; Zhou et al., 2003; Zhou, Burgoon, Twitchell, Qin, & Nunamaker, 2004b; Zhou, Burgoon, Zhang, & Nunamaker, 2004c). Burgoon and Zhou used a complex program to analyze verbal cues to deception, but this project used LIWC to help validate this user-friendlier program for lay use. More data should help researchers develop an automated and accessible deception-detection scheme.

## Chapter 2: Introduction

Deception is common, and can be harmful to relationships and personal finance. Adults most often lie to protect and promote themselves or to protect another (Seiter, Brusckke, & Bai, 2002). Adults are more likely to lie to strangers than friends (Ennis, Vrij, & Chance, 2008). Those with psychological disorders, like psychopathy, may deceive to seek thrills (Siegel, 1986). Consequently, when interacting with a stranger, a person may be lied to because the stranger has something to gain or because she enjoys deceiving.

Although deception is common, humans are poor deception detectors. A meta-analysis of 120 deception detection studies published from 1943 to 2000 concluded that humans could discriminate between truth and lies only 54% of the time (DePaulo et al., 2003). Deception detection was poor partly because people tended to believe that what others said was true. Although that “truth bias” may fulfill a psychological need to trust those around us and reduce cognitive load, the tendency decreases our ability to detect deception (Ekman, 2001; Feeley & Young, 1992; Levine, Kim, Park, & Hughes, 2006). Many people believe in cues that do not indicate deception. Those non-indicative cues include gaze direction, the most commonly reported cue to deception among laypeople (Global Deception Research Team, 2006). Belief in non-indicative cues hampers deception-detection accuracy (Global Deception Research Team, 2006; Vrij et al., 2000). Further, many people have a demeanor bias. That is the tendency to think some people look like deceivers and others look like truth tellers (Zuckerman, Defrank, Hall, Larrance, & Rosenthal, 1979). Demeanor bias can interfere with perception of actual cues to deception. For example, people judged some individuals as deceivers and others as truth tellers irrespective of how those individuals behavior (C. F. Bond, Omar, Mahmoud, & Bonser, 1990).

Truth bias, reliance in non-indicative cues, and demeanor bias all hamper laypeople's ability to detect deception. Computer programs, though, can automatically detect some cues. Automated deception detection removes human judgment from the equation and may increase a layperson's ability to detect deception. Seeing if computer programs could aid in deception detection was the goal of this project.

This research report describes theories about why cues to deception exist, theories specific to verbal cues to deception, and theories about computer mediation in communication. The report discusses automating detection of verbal deception cues and reviews prior studies on verbal cues to deception. The following literature review explains how each deception cue was operationalized, research results, and possible theoretical reasons for those results. The review lays the foundation for four hypotheses and one research question. The methods section describes the experiment performed in this study and how variables were operationalized. The results section reports descriptions of participants, results of hypothesis testing, and responses to the research question. The discussion section gives possible explanations for unexpected results, applications of outcomes to future research, and study limitations.

## Chapter 3: Literature Review

### Theories on Cues to Deception

Theories regarding deception and deception detection are divided here into three groups. The first group explains causes for cues to deception and how those cues vary in different contexts (Buller & Burgoon, 1994; DePaulo, Kirkendol, Tang, & Obrien, 1988; DePaulo, Lanier, & Davis, 1983; Zuckerman et al., 1981). Theories in this group are four-factor theory, the motivation impairment effect, and interpersonal deception theory. The second group deals with verbal cues to deception (Johnson & Raye, 1981; Mehrabian, 1966; Steller & Kohnken, 1989). These theories describe why cues occur and how to categorize and code for them. The theories and coding schemes are verbal immediacy, criteria-based content analysis, and reality monitoring. The last group is the most specific. It describes effects of computer-mediated communication on verbal cues to deception (Daft & Lengel, 1984; Walther, 1992; Walther, Anderson, & Park, 1994). That group consists of media richness theory, social presence theory, and social information procession theory. Those theories led to conclusions about synchronous computer-mediated communication, like instant messaging, and asynchronous computer-mediated communication, like email.

#### Causes of Cues to Deception

*Four-factor theory.* Zuckerman et al.'s (1981) meta-analysis of deception-detection research stated that no cue or cues to deception could be accurate all the time because deception was an individual psychological process. Researcher said four psychological factors contributed to cues to deception: generalized arousal, felt emotion, cognitive load, and strategy to appear

truthful (Zuckerman et al., 1981). The first factor, generalized arousal, involved heightened level of emotion not linked to a specific task or statement that the deceiver performed. The broader act of deceiving heightened arousal. Generalized arousal was observed through greater pupil dilation, higher vocal pitch, and more pauses in speech (Zuckerman et al., 1981). Higher emotional intensity caused more facial expression leakage (Porter, ten Brinke, & Wallace, 2011). The second factor encompassed feelings associated with deceiving, or felt emotion. Deceivers might feel guilt or fear. Evidence of that might leak out in gestures, pauses, and fidgeting or in unpleasant or unfriendly behavior. Zuckerman et al.'s (1981) third factor was cognitive load. Deceivers had a more cognitively demanding task than truth tellers because deceivers monitored their statements for consistency. Longer response latencies, hesitations, and more filler words, such as "um" and "ah," could signify a high cognitive load. The fourth factor was a deceiver's attempts to appear truthful. Deceivers had a stronger desire than truth tellers to appear truthful. Truth tellers might assume that they were believed. The desire to appear truthful could lead to overly controlled behavior, such as formal speech or too few natural gestures.

This study examined verbal cues to deception caused by each of Zuckerman et al.'s (1981) four factors.

*Motivation impairment effect.* The motivation impairment effect states that a liar's motivation is tied to his or her level of arousal and felt emotion. Higher motivation leads to more nonverbal leakage (DePaulo et al., 1988; DePaulo et al., 1983).

Many researchers have manipulated motivation to amplify cues to deception. For example, Ekman, O'Sullivan and Frank (1999) offered participants monetary rewards for being able to deceive interviewers. deTurck, Harszлак and Bodhorn (1990) and deTurck, Feeley and Roman (1997) encouraged participants to cheat and then lie about it. Ekman told deceivers that

lying would be important in their future careers (Ekman & Friesen, 1974; Ekman, Friesen, & Osullivan, 1988). DePaulo et al.'s (2003) meta-analysis supported the motivation-impairment effect by finding that higher motivation resulted in a greater number of deception cues being detected.

This study took deceiver's motivation into account. Participants were given the chance to win a prize. Such an offer has been shown to increase motivation and lead to more leaked cues to deception (DePaulo et al., 2003).

*Interpersonal deception theory.* Interpersonal deception theory states that interpersonal deception, like interpersonal communication, is dynamic and adaptive (Buller & Burgoon, 1994; Buller et al., 1996). While communicating, deceivers use feedback from message recipients to modify deception strategy. Deceivers' tactics include reducing the quantity of false statements, reducing the quality and clarity of statements by adding modifiers and qualifiers to decrease certainty, using irrelevant statements, depersonalizing statements to distance themselves from the lie, and bolstering their self-image (Zhou et al., 2004a). Researchers have found that deceivers' statements are not stable over time. Deceivers' number of words, verbal self-distancing tactics, and use of adjective and adverb increased during a conversation while complexity decreased. In general, deceivers' conveyed more cues to deception at the beginning of an interaction than at the end because they had not modified their behavior (Burgoon, Buller, White, Afifi, & Buslig, 1999; Burgoon & Qin, 2006).

This project did not examine deceivers' and truth tellers' statements over time, so interpersonal deception theory was not considered. A number of studies have looked at interpersonal deception theory and shed light on verbal cues to deception. Relevant results of those studies are reviewed below.

## Theories on Verbal Cues to Deception

Cues to deception fall into three categories: verbal, nonverbal, and physiological. Verbal cues can be subdivided into two groups: (1) content-based cues, which can be proved wrong when compared to the truth, and (2) linguistics-based cues, which can be detected when the truth is unknown. Because text-only computer-mediated communication does not convey nonverbal cues to deception and the truth is often difficult to verify, most studies on deception in computer-mediated communication focus on linguistics-based verbal cues. A number of theories and coding schemes for linguistics-based verbal cues to deception have been developed.

This study on deception in text-only computer-mediated communication, henceforth referred to as “computer-mediated communication,” investigated linguistics-based cues verbal to deception.

*Verbal immediacy.* Verbal immediacy is one basis for linguistics-based verbal deception cues, henceforth referred to as “verbal deception cues.” Immediacy is the directness and relevance of a person’s communication behavior and can be verbal or nonverbal (Mehrabian, 1966). Immediacy is a sign of caring and closeness and leads to perceptions of trustworthiness in others. Much of the research on immediacy has been on teacher immediacy, and high immediacy is linked with student impressions of trustworthiness and competence (Gorham, 1988; Neuliep, 1997; O’Sullivan, Hunt, & Lippert, 2004). Verbal immediacy has proved difficult to code (Robinson, 1995), but some indicators have been developed. They include using personal examples, asking questions, using humor, using others’ names, praising others, using detail and personal language, and avoiding qualifiers (Montgomery, 1981; O’Sullivan et al., 2004). Liars’ guilt and fear may cause them to distance themselves from a lie and/or communication partner.

Such distancing results in less verbal immediacy. Subsequent sections of this report identify cues falling into that category.

*Criteria-based content analysis.* Criteria-based content analysis is a coding scheme used to verify the veracity of verbal statements. It is based on the Undeutsch hypothesis, which states that a statement of an experience sourced from memory will differ in content and quality from a statement of a fabricated event (Blandon-Gitlin, Pezdek, Lindsay, & Hagen, 2009; Undeutsch, 1967). True statements have more of the criteria outlined by criteria-based content analysis, and, thus, have higher scores than deceptive ones. Originally developed to evaluate children's claims of sexual abuse (Steller & Kohnken, 1989), criteria-based content analysis has been used in a number of academic and forensic settings (Blandon-Gitlin et al., 2009; Sporer, 1997; Vrij, Akehurst, Soukara, & Bull, 2004; Vrij et al., 2000). Criteria-based content analysis criteria fall into three categories:

- (1) General characteristics. A true statement will be typical, coherent and have natural digressions and a normal level of detail.
- (2) Specific content. A true statement will embed the reported events in a time and place, will reproduce speech, mention complications, and contain superfluous and subjective content.
- (3) Control. Similar to Zuckerman et al.'s (1981) fourth factor, deceivers may overly manage their behavior resulting in cues to deception. Deceivers' statements will be overly controlled, have less natural language, include fewer self-corrections, and have fewer admissions of memory lapses.

Using criteria-based content analysis, Vrij et al. (2000) correctly classified 72.6% of truthful and deceptive statements, and Sporer (1997) classified 65%.

Criteria-based content analysis was designed to differentiate truthful from deceptive reports of past events. This project examined persuasive statements regarding a current situation. Criteria-based content analysis criteria were not used to identify verbal cues to deception. However, cues in the criteria-based content analysis schema have contributed greatly to verbal deception-detection research and were reviewed.

*Reality monitoring.* Reality monitoring, like criteria-based content analysis, is a coding scheme developed to test verbal statements in forensic settings. True statements coded for reality-monitoring criteria will contain more criteria, and have higher scores, than deceptive ones. Reality monitoring rests on the theory that memories of true events differ from memories of fabricated or imagined events. Statements about experienced events will include more sensory, spatial, temporal and emotional content because the memory was encoded in a real context. Imagined events lack that context. Inventing a memory is more cognitively demanding than reporting one (Johnson & Raye, 1981), so imagined events will contain more indicators of cognition. Using reality monitoring, Vrij et al. (2000) identified 67.12% of truthful and deceptive statements, and Sporer (1997) classified 71.3%. When verbal statements were coded with both reality-monitoring and criteria-based content analysis criteria, as well as for nonverbal cues to deception, researchers were able to classify about 80% of liars (Vrij et al., 2004; 2000).

Like criteria-based content analysis, reality monitoring was designed to differentiate truthful from deceptive reports of past events. Although cues-based reality-monitoring criteria were not identified in this experiment, research reporting these cues was reviewed.

Theories of Deception in Computer-Mediated Communication

Computer-mediated communication is growing in importance. Internet use increased more than 600-fold, from 2.6 million people in 1990 to more than 1.8 billion people in 2010 (World Bank, 2011). Researchers have developed theories that inform the study of verbal cues to deception in computer-mediated communication. Two forms of computer-mediated communication are discussed here: synchronous, real-time communication like chat and asynchronous, delayed-response communication like email. Media richness theory, social presence theory and social information processing theory are described below.

*Media richness theory.* Media richness theory states that media are defined by four characteristics: immediacy of feedback, presence of interpreting cues, ability to use natural language in the medium, and personal focus of the medium (Daft & Lengel, 1986). In face-to-face communication, all four characteristics are available, so the medium is rich. Communication via telephone removes visual interpreting cues, so the interaction is less rich than face-to-face communication. In synchronous computer-mediated communication, vocal interpreting cues, like voice pitch, are removed, so communication is less rich than via telephone. Asynchronous computer-mediated communication, the focus of this study, does not provide immediate feedback, visual or vocal cues. This medium is very low in richness.

According to media richness theory, people will be better able to detect deception in the presence of more defining characteristics, i.e. via richer media. Thus, deception detection in asynchronous computer-mediated communication would be poorer than in face-to-face interaction. A meta-analysis of mostly spoken messages found the average deception-detection rate was 54% (DePaulo et al., 2003), so media richness theory suggests that deception detection in asynchronous computer-mediated communication would be very close to, if not exactly at, chance level.

*Social presence theory.* Social presence is the salience of another during interaction. Social presence depends on the number of channels available in communication (Walther et al., 1994). Because asynchronous computer-mediated communication does not have a nonverbal channel, one would expect impressions of presence and warmth to be less than in face-to-face communication. Like media richness theory, social presence theory suggests that cues to deception would be less available in computer-mediated communication and that deception-detection accuracy would be less accurate than in face-to-face communication.

*Social information processing theory.* Social information processing theory states that communicators use whatever means possible to enhance expression. They adapt their style to make up for a less-rich medium. This adaptation includes changing verbal behavior to compensate for the lack of a nonverbal channel (Walther, 1992). Research shows that computer-mediated communication users include more immediacy in their messages than those communicating face-to-face (Walther, Loh, & Granka, 2005). Thus, computer-mediated messages may contain more verbal cues to deception than spoken ones because computer users compensate verbally for the lack of nonverbal social cues.

Social information processing theory applies to deception detection. Research suggests that deceivers may express more verbal social cues, including deception cues, when using computer-mediated communication than when communicating face-to-face. One study found that deception detection in computer-mediated communication was as accurate as in face-to-face communication (Hancock, Woodworth, & Gorrha, 2010). That finding implies that verbal cues in computer-mediated communication compensated for the lack of nonverbal cues. The same study found that motivated liars were easier to detect in both channels. That result supports the motivational-impairment effect in computer-mediated communication.

## **Automation of Deception Cues in Computer-Mediated Communication**

Verbal linguistics-based deception cues sent via computer-mediated communication are well-suited to automated deception detection. The verbal messages are already digitized and easily imported into software programs. Natural-language analyzing programs, such as LIWC, can categorize words and phrases by comparing them to a dictionary (Pennebaker et al., 2001). Other natural-language processing programs exist, such as the one designed by Voutilainen (2000), but in general they are not as user-friendly and accessible as LIWC.

In LIWC's dictionary words are divided into categories that can indicate the speaker's psychological state. LIWC's dictionaries categorize words hierarchically. For example, "happy" and "crying" would be placed in a number of categories. LIWC would categorize both as affective-processes words. The affective-processes dictionary contains sub-dictionaries for positive and negative emotion words. Therefore, LIWC would categorize "happy" as a positive-emotion word and "crying" as a negative-emotion word, in addition to their being categorized in the parent, affective-process dictionary. The negative-emotion dictionary has additional sub-dictionaries for anxiety, fear, and sadness. Thus, LIWC would categorize "crying" as an affective-processes word, a negative-emotion word, and a sadness word. "Crying" would be categorized according to part of speech and identified as a present-tense verb. LIWC's categories are not exhaustive. For instance, the program has no category for adjectives, like "happy." Although LIWC reports words per sentence, it does not report average word length.

Categories available in LIWC informed the choice of hypotheses in this study. The methods section discusses how each variable was operationalized and how LIWC's dictionaries were used to measure results.

## **Review of Verbal Cues to Deception and Hypotheses**

Researchers have identified cues for differentiating truthful from deceptive statements. Researchers examined different cues, and some researchers operationalized the same cues differently. Instead of reviewing the literature study-by-study, this section examines each verbal cue to deception, how it was operationalized, research results, and possible theoretical reasons for the results.

### **Categories of Verbal Cues to Deception**

From an examination of research literature, five categories of verbal cues to deception emerged. The categories depend on the underlying theory of why the cue occurs: generalized arousal and felt emotion, cognitive load, and the differences between real and imagined events (Johnson & Raye, 1981; Steller & Kohnken, 1989; Zuckerman et al., 1981). The first three categories, verbosity, emotion words and verbal immediacy, contain cues measuring a deceivers' generalized arousal and felt emotion but approach measurement in different ways. The second category, linguistic complexity, includes cues that measure cognitive load. People with a higher cognitive load tend to use less complex language (Burgoon & Qin, 2006; DePaulo et al., 2003; Pennebaker et al., 2003; Zhou et al., 2004a). The last category, qualitative differences between real and imagined memories, pulls from criteria-based content analysis and reality-monitoring criteria on how deceptive event descriptions differ from truthful ones. Each category contains multiple cues, and in many instances, researchers operationalized a single cue differently.

Verbosity, verbal immediacy and emotion words attempt to ascertain a communicator's affective state. Many researchers measure verbosity, the average length of messages, because it is easy to operationalize with word or sentence counts. Two theories influence verbosity. First, a

deceiver's guilt may cause him to make fewer deceptive utterances and be reticent (DePaulo et al., 2003). Second, a deceiver who wants to appear more truthful may be more verbose than a truth teller (Zhou & Zhang, 2008). Verbal immediacy cues are related to a communicator's directness and relevance. Deceivers' guilt, fear, and anxiety may cause them to be less direct (DePaulo et al., 2003; Ebesu & Miller, 1994; Pennebaker et al., 2003). Directness can be difficult to operationalize, and a verbal immediacy scale has yet to be developed and validated. Cues in this broad category include first-person pronoun use, demonstrated directness and ownership of a statement, and passive verb use, indicating indirectness. Emotional cues are more directly tied to a communicator's affective state than either verbosity or verbal immediacy. Deceivers experience more fear, guilt, and, perhaps, more emotion in general than truth tellers (Zuckerman et al., 1981). Such emotion is conveyed in verbal communication. Emotion words can be positive, like happy, and negative, like fear. They are easily identified, categorized and counted by natural-language processing software.

Cues indicating linguistic complexity, the fourth category, attempt to operationalize a speaker's cognitive load. Deceivers manage their behavior more than truth tellers and have a higher cognitive load (Zuckerman et al., 1981). Their higher cognitive load reduces deceivers' ability to use complex language. Some cues of linguistic complexity include sentence length, word length, and punctuation use.

The fifth category encompasses cues reflecting the qualitative differences between descriptions of real and fabricated events. Real events were encoded into memory in a real context, rich with detail and perceived through the speaker's senses. Imagined events were not recorded through the senses and did not occur in a detailed, genuine context. Consequently, reports of real events contain more sensory and contextual information and more detail than

reports of fabricated events (Johnson & Raye, 1981; Steller & Kohnken, 1989). Cues in this category include amount of sensory information and detail in a statement.

Results from studies on verbal cues to deception in each of the five categories are reported in this section.

### Verbosity

Verbosity measures how lengthy a person’s verbal messages are. It can be operationalized as number of words (Burgoon et al., 2003; Burgoon & Qin, 2006; DePaulo et al., 2003; Hancock et al., 2008; Zhou et al., 2004a; Zhou et al., 2003; Zhou & Sung, 2008; Zhou & Zhang, 2008), number of sentences (Twitchell, Forsgren, Wiers, Burgoon, & Nunamaker, 2005; Zhou et al., 2004a; Zhou & Sung, 2008; Zhou & Zhang, 2008), and number of various linguistic features (Zhou et al., 2003). Table 1 divides verbosity into two frequently reported cues: number of words and number of sentences.

**Table 1: Verbosity**

A-CMC = asynchronous computer-mediated communication  
 S-CMC = synchronous computer-mediated communication

Verbal cue	Liars have	Medium	Possible cause	Operationalized as	Source
<b>Number of words</b>	fewer	spoken and A-CMC	emotion/arousal; fear; desire to produce fewer deceptive messages	number of words	(Burgoon et al., 2003)
		multiple			(DePaulo et al., 2003)
		S-CMC, Chinese			(Zhou & Sung, 2008)
		Spoken			(Burgoon & Qin, 2006)
	more	A-CMC	strategy of self-promotion, preparation time	number of words, verbs, modifiers, noun phrases and sentences	(Zhou et al., 2003)
		S-CMC	strategy of self-promotion	number of words	(Hancock et al., 2008)
		A-CMC	strategy of self-promotion, preparation time		(Zhou et al., 2004a)
		A&S-CMC	strategy of self-promotion, preparation time		(Zhou & Zhang, 2008)

<b>Number of sentences</b>	fewer	S-CMC, Chinese	emotion/arousal; fear; desire to produce fewer deceptive messages	number of sentences	(Zhou & Sung, 2008)
		S-CMC			(Twitchell et al., 2005)
	more	A-CMC	strategy of self-promotion, preparation time		(Zhou et al., 2004a)
		A&S-CMC	strategy of self-promotion		(Zhou & Zhang, 2008)

Studies considering deceivers’ and truth-tellers’ numbers of words yield conflicting results. Two studies found that in spoken communication, deceivers used fewer words than truth tellers (Burgoon et al., 2003; Burgoon & Qin, 2006). The same result was found when data from multiple (mostly spoken) media were pooled (DePaulo et al., 2003). Fear, guilt, and anxiety could cause liars to distance themselves from deceptive statements and use fewer words. In contrast, three studies and one summary of research in deception in computer-mediated communication found that deceivers used more words than truth tellers (Hancock et al., 2008; Zhou et al., 2004a; Zhou et al., 2003; Zhou & Zhang, 2008). The seeming conflict could indicate that deceivers have more time when using computers to prepare their messages than in face-to-face encounters. Deceivers can take time to use strategy to appear truthful and forthcoming, one aspect of which is using more words, resulting in higher word counts. Being face-to-face with a receiver may cause more anxiety and reticence than communicating via computer.

Results of the single Chinese-language computer-mediated communication study do not support the trend of deceivers using fewer words when speaking but more words via computer-mediated communication. The authors acknowledged that word-count in Chinese could be problematic because Chinese is a monosyllabic language. Consequently, their analysis may not be comparable to English-language findings (Zhou & Sung, 2008).

Results on deceivers' and truth tellers' number of sentences in synchronous and asynchronous computer-mediated communication followed a similar trend. In synchronous computer-mediated communication, deceivers wrote fewer sentences than truth tellers, likely because of anxiety and arousal (Twitchell et al., 2005; Zhou & Sung, 2008). However, in asynchronous computer-mediated communication, deceivers tended to write more sentences than truth tellers, probably because they had time to strategize (Zhou et al., 2004a; Zhou & Zhang, 2008).

Those results led to the following hypothesis:

**H1) Deceivers will be more verbose than truth tellers.**

### Verbal Immediacy

Verbal immediacy reflects a communicator's directness. Researchers operationalize verbal immediacy in a number of ways: by counting first- and third-person pronouns (G. D. Bond & Lee, 2005; Ebesu & Miller, 1994; Hancock et al., 2008; Lee, Welker, & Odom, 2009; Pennebaker et al., 2003; Zhou et al., 2004a; Zhou et al., 2003; Zhou & Sung, 2008; Zhou & Zhang, 2008), group references (Ebesu & Miller, 1994; Zhou et al., 2004a; Zhou et al., 2003), causation words (Hancock et al., 2008; Lee et al., 2009), qualifiers (Ebesu & Miller, 1994), present-tense verbs (Ebesu & Miller, 1994), modal verbs (Zhou, 2005; Zhou et al., 2004a; Zhou et al., 2003; Zhou & Zhang, 2008), and exclusion words (Pennebaker et al., 2003). Some studies pool data from one or more of those cues to measure verbal immediacy (DePaulo et al., 2003). Research results on verbal-immediacy cues are summarized in Table 2.

**Table 2: Verbal Immediacy**

Verbal cue	Liars have	Medium	Possible cause	Operationalized as	Source
<b>First-person pronouns</b>	fewer	multiple	emotion/arousal, distance self from deceptive statements	number or percent of first person singular pronouns	(Pennebaker et al., 2003)
		spoken			(Ebesu & Miller, 1994)
		A-CMC			(Zhou et al., 2003)
		A-CMC			(Zhou et al., 2004a)
		A&S-CMC		(Zhou & Zhang, 2008)	
	S-CMC	number or percent of first person pronouns	(Hancock et al., 2008)		
more	A-CMC	self-promotion, attempt to appear truthful	number of first person singular pronouns	(Lee et al., 2009)	
<b>Third-person pronouns</b>	fewer	multiple	emotion/arousal leads to less immediacy	number or percent of third person pronouns	(Pennebaker et al., 2003)
		spoken			(G. D. Bond & Lee, 2005)
	more	S-CMC	strategy to shift focus away from self		(Hancock et al., 2008)
		S-CMC, Chinese			(Zhou & Sung, 2008)
<b>Group references</b>	more	spoken	strategy to dilute effect of first person statements	number or percent of first person plural pronouns	(Ebesu & Miller, 1994)
		A-CMC			(Zhou et al., 2003)
		A-CMC			(Zhou et al., 2004a)
<b>Causation words</b>	fewer	A-CMC	Emotion/arousal leads to inexact language	number of causation words	(Lee et al., 2009)
	more	S-CMC	strategy to appear exact		(Hancock et al., 2008)
<b>Qualifiers</b>	More	Spoken	emotion arousal leads to inexact language	percent of qualifying terms	(Ebesu & Miller, 1994)
<b>Present tense verbs</b>	fewer	spoken	emotion/arousal leads to less immediacy	percent present tense verbs	(Ebesu & Miller, 1994)
<b>Modal verbs</b>	more	A&S-CMC	emotion/arousal leads to less immediacy	number or percent of modal verbs	(Zhou & Zhang, 2008)
		A-CMC			(Zhou et al., 2003)
		A-CMC			(Zhou et al., 2004a)
		S-CMC			(Zhou, 2005)
<b>Exclusion words</b>	fewer	multiple	emotion/arousal leads to desire to be inexact	percent exclusion words	(Pennebaker et al., 2003)
<b>Verbal immediacy</b>	less	multiple	emotion/arousal leads to less immediacy	multiple factors	(DePaulo et al., 2003)

*First-person pronouns.* Five studies operationalized first-person pronouns and self-references in the singular only. These studies counted occurrences of “I,” “me,” “my,” and “myself” in statements (Ebesu & Miller, 1994; Lee et al., 2009; Pennebaker et al., 2003; Zhou et

al., 2004a; Zhou et al., 2003). Hancock et al.'s (2008) study combined first-person singular and plural pronouns, as did Zhou and Zhang's (2008) summary of verbal deception cue data.

Despite operationalizing first-person pronouns differently, six of the seven studies found deceivers used fewer first-person pronouns and self-references than truth tellers. Lee, Welker and Odom (2009) found deceivers used more first-person pronouns than truth tellers. Deceivers likely used fewer first-person pronouns as a distancing tactic because of anxiety. They did not want to be seen as taking ownership of deceptive statements.

*Group references.* Group references are operationalized as number or percentage of first-person plural pronouns. The three studies reporting this cue found deceivers used more group references than truth tellers in both spoken and computer-mediated communication (Ebesu & Miller, 1994; Zhou et al., 2004a; Zhou et al., 2003). Using first-person plural instead of the singular dilutes the effect of statement ownership. Deceivers' guilt and fear could cause them to avoid taking sole ownership of false statements.

*Modal verbs.* Modal verbs include words like "may" and "could" that complement the main verb and are included in a verb phrase. Modals reduce the strength of a main verb. Researchers count the occurrence of modals in deceivers' and truth tellers' statements. All four studies reporting on modal verbs in computer-mediated communication found that deceivers used more modal verbs than truth tellers (Zhou, 2005; Zhou et al., 2004a; Zhou et al., 2003; Zhou & Zhang, 2008). Modals may reduce the certainty, directness and immediacy of a statement. These tactics may distance the fearful deceiver from the words.

Findings on pronouns, group references and modal verbs led to the following hypothesis:

## **H2) Deceivers will use:**

- a) fewer first-person pronouns and self-references,**
- b) more group references, and**
- c) more modal verbs,**

**than truth tellers.**

*Causation words.* Words like “because” and “so” show causation. They link cause and effect, a sign of candidness. Of the two studies that reported causation words, one found deceivers used more causation words (Hancock et al., 2008), the other fewer (Lee et al., 2009). Deceivers were expected to be anxious and indirect and to use fewer causation words, but that was not conclusively supported by research.

*Qualifiers.* Qualifiers are adverbs like “maybe” and “perhaps.” They weaken a statement by modifying and reducing the strength of the verb. Researchers count qualifiers in statements. One study on spoken deception found deceivers used more qualifiers than truth tellers (Ebesu & Miller, 1994). Deceivers probably used more qualifiers to weaken their deceptive statements and reduce feelings of fear and guilt.

*Present-tense verbs.* Present-tense verbs describe current actions. One study on spoken deception found deceivers used fewer present-tense verbs than truth tellers (Ebesu & Miller, 1994). Relevance, including temporal relevance indicated by present-tense verbs, showed verbal immediacy. Only one study reported data on present-tense verbs.

*Exclusion words.* Exclusion words make statements more exact by narrowing definitions and pointing out negative instances. Exclusion words include “but” and “except,” which are counted in statements. One study on deception across multiple media found deceivers used fewer exclusion words than truth tellers (Pennebaker et al., 2003). Deceivers probably wanted to be

vague and, therefore, used fewer exclusion words than truth tellers. Only one study reported data on exclusion words.

*Third-person pronouns.* Researchers counted the number of third-person pronouns, singular and plural, in statements tellers (G. D. Bond & Lee, 2005; Hancock et al., 2008; Pennebaker et al., 2003; Zhou & Sung, 2008). Two studies, one using multiple media and the other spoken, found deceivers used fewer third-person pronouns than truth tellers (G. D. Bond & Lee, 2005; Pennebaker et al., 2003). Third-person pronouns referred directly to objects and could reflect verbal immediacy. Deceivers were expected to be less direct. Two other studies, however, both on deception in computer-mediated communication, found liars used more third-person pronouns than truth tellers (Hancock et al., 2008; Zhou & Sung, 2008). The reason could be that using a computer, even in synchronous communication, gave deceivers more time to strategize than when speaking. Time to plan might have overcome anxiety and arousal. Use of more third-person pronouns could be a tactic to divert attention from deceptive statements.

The data on deceivers' use of third-person pronouns and other references are unclear. Different theories support liars using both more and fewer third-person pronouns than truth tellers. That confusion and the limited data on causation words, qualifiers, present-tense verbs and exclusion words in deception led to the following research question:

**RQ1) How will deceivers use:**

- a) **causation words,**
- b) **qualifiers,**
- c) **present-tense verbs,**
- d) **exclusion words, and**
- e) **third-person pronouns**

**when compared to truth tellers?**

Emotion Words

Emotion words express affect in a statement and reflect the emotional state of the speaker. Five groups of researchers operationalized emotion words by counting how many negative emotion words appeared in statements (DePaulo et al., 2003; Pennebaker et al., 2003; Zhou et al., 2004a; Zhou et al., 2003; Zhou & Zhang, 2008). Two groups of researchers counted the number of positive emotion words as well (Zhou et al., 2004a; Zhou & Zhang, 2008). One study compared the number of negative statements and complaints in deceiver's and truth tellers' statements (DePaulo et al., 2003). Two studies operationalized emotiveness, the relative number of adjectives and adverbs compared to nouns and verbs in a statement (Zhou et al., 2004a; Zhou & Zhang, 2008). Table 3 summarizes research results on emotion words.

**Table 3: Emotion Words**

Verbal cue	Liars have	Medium	Possible cause	Operationalized as	Source
<b>Negative emotion words</b>	more	multiple	emotion/arousal	number or percent of negative emotion words	(Pennebaker et al., 2003)
		multiple			(DePaulo et al., 2003)
		A-CMC			(Zhou et al., 2004a)
		A-CMC			(Zhou et al., 2003)
		A&S-CMC			(Zhou & Zhang, 2008)
<b>Negative statement/complaint</b>	more	multiple	emotion/arousal	number of negative statements	(DePaulo et al., 2003)
<b>Positive emotion words</b>	more	A-CMC	strategy for self-promotion	number of positive emotion words	(Zhou et al., 2004a)
		A&S-CMC			(Zhou & Zhang, 2008)
<b>Expressivity &amp; Emotiveness</b>	more	A&S-CMC	emotion/arousal	frequency of adjectives and adverbs	(Zhou & Zhang, 2008)
		A-CMC		adjectives & adverbs / nouns & verbs	(Zhou et al., 2004a)

*Negative emotion words.* Negative emotions are expressed in words like “fear” and their adjectival (fearful) and adverbial (fearfully) forms. Researchers compare the number of negative emotion words in deceivers’ and truth tellers’ statements. All five studies and data summaries in Table 3—covering spoken, synchronous and asynchronous computer media—found deceivers used significantly more negative emotion words than truth tellers (DePaulo et al., 2003; Pennebaker et al., 2003; Zhou et al., 2004a; Zhou et al., 2003; Zhou & Zhang, 2008). Deceivers were thought to experience negative emotions, which leaked out and colored their speech.

*Positive emotion words.* Words like “happiness,” “happy,” and “happily” show positive emotions. Researchers compare how many positive emotion words occur in deceivers’ and truth tellers’ statements. One study and one data summary found deceivers used more positive emotion words than truth tellers (Zhou et al., 2004a; Zhou & Zhang, 2008). Several reasons could explain that result. Deceivers using computer-mediated communication have more time to strategize and create positive emotional statements in an attempt to appear truthful. Deceivers may genuinely experience positive emotion because of duping delight or because of generalized arousal.

*Emotiveness.* Emotiveness is operationalized as the number of adjectives and adverbs compared to total word count or the number of adjectives and adverbs compared to the number of nouns and verbs (Zhou et al., 2004a; Zhou & Zhang, 2008). One study and one data summary found that deceivers were more emotive than truth tellers (Zhou et al., 2004a; Zhou & Zhang, 2008). Deceivers might have been more emotional because of arousal, and their emotionality leaks out in a higher number of description words. Deceivers might have used more adjectives and adverbs as a strategy to appear more descriptive and truthful.

Findings on effects of emotional words in deception led to the following hypothesis:

### **H3) Deceivers' statements will have:**

- a) more positive emotion words,**
- b) more negative emotion words, and**
- c) be more emotive**

**than those of truth tellers.**

*Negative statements/complaints.* Some researchers counted the number of times participants complained during a deception task. Depaulo et al.'s (2003) meta-analysis found that deceivers complained significantly more than truth tellers, but most of the studies she examined were of spoken data. This project was not time-consuming, and participants were not engaged in free dialog, so few, if any, complaints were expected. Negative statements and complaints were not be analyzed.

### Linguistic Complexity

A statements' linguistic complexity can indicate the communicator's cognitive load. People with a high cognitive load have fewer resources to dedicate to creating complex messages. Those people tend to produce simpler statements. Linguistic complexity can be operationalized a number of ways. Two studies counted the number of motion verbs (G. D. Bond & Lee, 2005; Pennebaker et al., 2003). Motion verbs like "go" are simple compared to abstract verbs like "love." A higher occurrence of simple verbs indicates less linguistic complexity. Two studies measured sentence length (Burgoon et al., 2003; Zhou & Sung, 2008). Shorter sentences are less linguistically complex. Six studies reported content diversity, the percentage of unique and different words in a statement (Burgoon & Qin, 2006; Zhou, 2005; Zhou et al., 2004a; Zhou et al., 2003; Zhou & Sung, 2008; Zhou & Zhang, 2008). More unique content indicates more

complex language. Redundancy, the number of repeated words and phrases in a statement, is a cue of linguistic simplicity (Burgoon & Qin, 2006; DePaulo et al., 2003). One study (Burgoon et al., 2003) operationalized and reported language complexity as the Flesh-Kincaid readability grade level of a statement (Burgoon et al., 2003), a measure of language use. Another study assumed that more punctuation indicated more complex language and counted punctuation in a statement (Burgoon & Qin, 2006). Yet another study factored average sentence length, average word length and punctuation count (Zhou & Zhang, 2008) to create a linguistic complexity construct. Burgoon et al. (2003) reported sentence complexity, but how the authors operationalized the concept was unclear. Table 4 summarizes research results on linguistic complexity.

**Table 4: Linguistic Complexity**

Verbal cue	Liars have	Medium	Possible cause	Operationalized as	Source
<b>Motion verbs</b>	more	spoken	less complex due to cognitive load	percent motion words	(G. D. Bond & Lee, 2005)
		multiple			(Pennebaker et al., 2003)
<b>Sentence length</b>	shorter	spoken	less complex language due to higher cognitive load	Number of long sentences	(Burgoon et al., 2003)
		S-CMC, in Chinese		message length	(Zhou & Sung, 2008)
<b>Content diversity</b>	less	Spoken (deception first)	less complex language due to higher cognitive load	percent or number of different words or terms	(Burgoon & Qin, 2006)
		A&S-CMC			(Zhou & Zhang, 2008)
		A-CMC			(Zhou et al., 2003)
		S-CMC			(Zhou, 2005)
	A-CMC	(Zhou et al., 2004a)			
more	S-CMC in Chinese	strategy to appear truthful or misdirect	(Zhou & Sung, 2008)		
<b>Redundancy</b>	more	Spoken (deception first)	Less complex language due to cognitive load	number of repeated words	(Burgoon & Qin, 2006)
		Multiple		number or repeated word and phrases	(DePaulo et al., 2003)
<b>Casual language</b>	more	A-CMC	Higher cognitive load and/or emotional arousal	Percent or number of typos	(Zhou et al., 2003)
		A-CMC			(Zhou et al., 2004a)
		A&S-CMC			(Zhou & Zhang, 2008)

			leads to more mistakes and less exact language		
<b>Language complexity</b>	more	spoken	Self-promotion strategy	multiple factors	(Anolli, Balconi, & Ciceri, 2003)
	less	Spoken (deceptive statement first)	Less complex language due to higher cognitive load	amount of punctuation	(Burgoon & Qin, 2006)
		Spoken and S-CMC		sentence complexity	(Burgoon et al., 2003)
		Spoken and S-CMC		Flesh-Kincaid grade level	(Burgoon et al., 2003)
		A&S-CMC		average sentence and word length, punctuation	(Zhou & Zhang, 2008)

*Motion verbs.* Motion verbs and their conjugations, like “walk” and “talking,” refer to actions. Motion verbs can be counted and compared across deception condition. Two studies, one on spoken communication, the other in multiple media, found that deceivers’ statements had significantly more motion verbs than those of truth tellers’ (G. D. Bond & Lee, 2005; Pennebaker et al., 2003). Motions are concrete and easy to describe, unlike abstract actions, like “love” and “ponder.” Since deceivers try to manage the impressions have of them, liars have a higher cognitive load than truth tellers and might use less abstract and complex language.

*Sentence and word length.* Sentence length can be operationalized as the number of long sentences in a statement (Burgoon et al., 2003) or the length of a message (Zhou & Sung, 2008). Word length is operationalized as the average length of a word in a statement (Zhou & Zhang, 2008). One study found that found that in spoken communication, deceivers uttered shorter sentences than truth tellers (Burgoon et al., 2003). A second study, on Chinese users of synchronous computer-mediated communication, reported the same trend (Zhou & Sung, 2008). Zhou and Zhang’s (2008) summary of verbal deception reported that deceivers’ statements were less complex than truth tellers. That finding was based on a linguistic complexity construct

incorporating average sentence and word lengths and punctuation. Shorter sentences are grammatically simpler because they incorporate fewer clauses. Shorter words, like “a,” “an,” and “the,” tend to be function words that do not convey meaning. The presence of function words reduces the average word length in a statement. Deceivers whose cognitive load is high may be more likely to utter simpler sentences and more abstract statements than truth tellers. Simpler statements will have shorter sentences and more function words. Those qualities will reduce the average sentence and word lengths.

*Content diversity and redundancy.* Content diversity has been operationalized as the number or percentage of different words or terms in a statement (Burgoon & Qin, 2006; Zhou, 2005; Zhou et al., 2004a; Zhou et al., 2003; Zhou & Sung, 2008; Zhou & Zhang, 2008). Redundancy has been operationalized as the number of repeated words or phrases in a statement (Burgoon & Qin, 2006; DePaulo et al., 2003). Both approaches measure the linguistic complexity of the words in a statement. Regarding content diversity, four studies—one in spoken communication (Burgoon & Qin, 2006), three in computer-mediated communication (Zhou, 2005; Zhou et al., 2004a; Zhou et al., 2003), and one summary of data on computer-mediated communication (Zhou & Zhang, 2008)—found that deceivers’ statements had less content diversity than those of truth tellers. Zhou and Sung’s (2008) study of Chinese computer-communication users found the opposite, but the mono-syllabic nature of Chinese could make word counts inaccurate. Regarding redundancy, two studies found that in spoken communication and in multiple media, liars repeated themselves more than truth tellers (Burgoon & Qin, 2006; DePaulo et al., 2003). Thinking of new words and phrases to say required more cognition than repeating oneself. Deceivers had a higher cognitive load and fewer cognitive resources to dedicate to creating unique messages, so they were more redundant than truth tellers.

Research results on motion verbs, sentence length and word length, led to the following hypothesis.

**H4) Deceivers' statements will have:**

- a) more motion verbs,**
  - b) shorter sentences, and**
  - c) shorter words,**
- than those of truth tellers.**

Qualitative Differences between Real and Imagined Events

Some verbal cues to deception are based on the theory that reports of real events will be qualitatively different from those of fabricated events. Five studies and one meta-analysis reported cues in this category (G. D. Bond & Lee, 2005; DePaulo et al., 2003; Hancock et al., 2008; Sporer, 1997; Vrij et al., 2004; Vrij et al., 2000). All but two studies analyzed spoken data. One that did not was a meta-analysis of multiple communication media (DePaulo et al., 2003), and the other examined synchronous computer-mediated communication (Hancock et al., 2008). Spatial words, temporal words, and visual perception and auditory perception words were operationalized according to reality-monitoring criteria and coded by trained individuals (G. D. Bond & Lee, 2005; Vrij et al., 2004; Vrij et al., 2000). Reproduction of conversation, amount of detail, and contextual embedding were operationalized according to criteria-based content analysis criteria and coded by trained individuals (Sporer, 1997; Vrij et al., 2004). Sensory words were counted in statements (G. D. Bond & Lee, 2005; Hancock et al., 2008), as were instances of admitted memory flaws (DePaulo et al., 2003). Cues based on the differences between real and

imagined events do not apply to the present study on deception in persuasion. Nevertheless, these concepts have value in research on verbal cues to deception, so they are summarized in Table 5.

**Table 5: Qualitative Differences between Real and Imagined Events**

RM = Reality monitoring

CBCA = Criteria-based content analysis

Verbal cue	Liars have	Medium	Possible cause	Operationalized as	Source
Spatial words	fewer	spoken	difference between actual and imagined events	RM scoring	(Vrij et al., 2000)
		spoken			(Vrij et al., 2004)
	more	spoken	strategy to appear truthful		(G. D. Bond & Lee, 2005)
Temporal words	fewer	spoken	difference between actual and imagined events	RM scoring	(Vrij et al., 2000)
		spoken			(Vrij et al., 2004)
Reproduction of conversation	less	spoken	difference between actual and imagined events	CBCA scoring	(Vrij et al., 2004)
Amount of detail	less	multiple	difference between actual and imagined events	Varied among studies	(DePaulo et al., 2003)
		spoken		CBCA scoring	(Vrij et al., 2004)
Contextual embedding	less	spoken	difference between actual and imagined events	CBCA scoring	(Vrij et al., 2004)
		spoken			(Sporer, 1997)
Visual perception words	fewer	spoken	difference between actual and imagined events	RM scoring	(Vrij et al., 2000)
		spoken			(Vrij et al., 2004)
Auditory perception words	fewer	spoken	difference between actual and imagined events	RM scoring	(Vrij et al., 2000)
		spoken			(Vrij et al., 2004)
Sensory words	more	S-CMC	difference between actual and imagined events	number or percent sense words	(Hancock et al., 2008)
	fewer	spoken	strategy to appear truthful		(G. D. Bond & Lee, 2005)
Cognitive complexity	more	spoken	difference between real and imagined events	RM scoring or number of cognition words	(Vrij et al., 2000)
		spoken			(Vrij et al., 2004)
		A&S-CMC		percent cognition words	(Zhou & Zhang, 2008)
		S-CMC			(Zhou, 2005)
Admitted memory flaws	fewer	multiple	strategy to appear truthful	number of self-doubting statements	(DePaulo et al., 2003)

*Spatial words.* Spatial words are words like “next to” and “near.” Reality monitoring coders identify the presence of spatial words in a statement. Two studies found that deceivers’ spoken statements had fewer spatial words than those of truth tellers (Vrij et al., 2004; Vrij et al.,

2000), and one found the opposite (G. D. Bond & Lee, 2005). The theory behind reality monitoring states that memories are laid down in a concrete context. That context includes objects, people, and their relative positions. Fabricated events lack such context, so descriptions of real events will contain more spatial-contextual cues than descriptions of fabricated events. On the other hand, deceivers could add spatial-contextual detail in an attempt to appear truthful.

*Temporal words.* Temporal words sequence events in a statement temporally, like “before” and “during.” Reality monitoring coders identify the presence of temporal words in a statement. Two studies found that deceivers’ statements had fewer temporal words than those of truth tellers (Vrij et al., 2004; Vrij et al., 2000). The theory behind reality monitoring states that real memories will be laid down as events unfold. Descriptions of real events will contain information about what happened in what order and, thus, have more temporal words than descriptions of fabricated events.

*Visual and auditory perception words and sensory words.* Visual and auditory perception words describe events as they were seen or heard and include words like “glimpsed” and “heard.” Reality-monitoring coders identify the presence of visual and auditory words in statements. Sensory words combine words referring to all of the senses. These words can be counted in statements. Two studies on spoken statements found that deceivers’ statements had fewer visual perception words and fewer auditory perception words than those of truth tellers (Vrij et al., 2004; Vrij et al., 2000). One study reported that in spoken communication, deceivers used fewer sensory words than truth tellers (G. D. Bond & Lee, 2005). A second study reporting sensory words in computer-mediated communication found the opposite (Hancock et al., 2008). The theory behind reality monitoring is that real events are perceived and that imagined events are not. Descriptions of real events will have more visual and auditory perception words as well

as more sensory words than those of fabricated events. On the other hand, deceivers could add more sensory words to their statements in an attempt to appear truthful.

*Contextual embedding.* Contextual embedding is the extent to which reported events are embedded in a spatial and temporal context. Trained criteria-based content analysis coders identify the presence of contextual embedding in a statement. Two studies on spoken communication found that deceivers' statements were less likely to contain contextual embedding than those of truth tellers (Sporer, 1997; Vrij et al., 2004). That is probably because real events occurred in real contexts. Statements about real events will refer to those contexts more so than statements about fabricated events.

*Amount of detail and reproduction of conversation.* A trained criteria-based content analysis coder can judge the amount of detail in a statement. Reproduction of conversation, a specific instance of detail in a statement, is either present or absent in a statement. One study on spoken statements found that deceivers' used less detail and that their statements had fewer instances of reproduction of conversation than those of truth tellers (Vrij et al., 2004). A meta-analysis of multiple modes of communication found that deceivers' statements had less detail than those of truth tellers (DePaulo et al., 2003). That is probably because real events occur in contexts full of detail and conversation. Fabricated events do not. Reports of real events will likely contain more instances of detail and reproductions of conversation than descriptions of fabricated events.

*Admitted memory flaws.* Admitted memory flaws are present when a speaker says he is unsure of his memory. DePaulo et al.'s (2003) meta-analysis found that deceivers' statements had fewer admitted memory flaws than those of truth tellers. Deceivers wanted to appear truthful. They were less likely to admit that they were not sure. That admission could cast doubt

on their statements. So, their statements were likely to contain fewer instances of admitted memory flaws than those of truth tellers, who usually took for granted that they would be believed.

#### Summary of Hypotheses and Research Questions

The preceding literature reviewed led to the following hypotheses and research question:

**H1) Deceivers will be more verbose than truth tellers.**

**H2) Deceivers will use:**

- a) fewer first-person pronouns and self-references,**
- b) more group references, and**
- c) more modal verbs,**

**than truth tellers.**

**H3) Deceivers' statements will have:**

- a) more positive emotion words,**
- b) more negative emotion words, and**
- c) be more emotive**

**than those of truth tellers.**

**H4) Deceivers' statements will have:**

- a) more motion verbs,**

**b) shorter sentences, and**  
**c) shorter words,**  
**than those of truth tellers.**

**RQ1) How will deceivers use:**

**a) causation words,**  
**b) qualifiers,**  
**c) present-tense verbs,**  
**d) exclusion words, and**  
**e) third-person pronouns**  
**when compared with truth tellers?**

## Chapter 4: Method

### Participants

Volunteer participants were recruited from a pool of undergraduate students in a large southeastern university in the United States. These volunteers were asked to complete an experiment for course credit. Effects on linguistic cues to deception have been detected in message comparisons between 14 truthful and 16 deceptive participants (Zhou et al., 2004a). This study envisioned up to 400 participants. The study was approved by Virginia Tech's Institutional Review Board. See Appendix A.

This research used one-tailed *t*-tests to detect directional effects and two-tailed *t*-tests to detect bi-directional effects, both with a significance level of  $p < .05$ , in the four hypotheses and the research question. To detect medium-sized effects between two groups at that significance level, groups of 64 individuals are needed (Cohen, 1988). Cohen, however, assumes good measurement tools to detect those medium-sized effects. This study's primary measure, LIWC, has categories that have internal reliabilities (Cronbach's Alpha) reported in Table 6. LIWC's categories, other than first-person pronouns, group references and exclusion words, have reliabilities of greater than 0.80, and are considered good. This experiment was a replication of prior work with a new, less-tested measurement tool. The reliabilities lower than the standard of 0.80 were accepted because the project's exploratory nature. To improve the effects size, more than 64 participants were assigned to each condition. Up to 80 people were asked to participate in the pilot study (40 in each condition). The main study called for up to 320 participants (160 in each condition).

In both the pilot and main studies, participants were randomly assigned to either the truth or deception condition. Since no procedural changes were made after the pilot study, data from those groups were combined in the final analysis.

## **Measures**

### Demographics

Participants were asked their age, sex, major, and level of education. No hypotheses or research questions related to demographics. These measures were used to describe the participants.

### Personality

Participants completed a five-factor personality test (Big Five Personality Inventory). It measures extraversion, agreeableness, conscientiousness, neuroticism, and openness (John, Donahue, & Kentle, 1991; John, Naumann, & Soto, 2008). Neuroticism has been linked with a deceptive interpersonal communication style (Gudjonsson & Sigurdsson, 2004; Weaver III, 2005). Psychoticism, a correlate of neuroticism (Gudjonsson & Sigurdsson, 2004), and extraversion (Kashy & DePaulo, 1996) were linked to deception as well. No hypotheses or research questions dealt with personality and deception. Personality measures familiarized participants with the computer interface used in this study.

### Experimental Manipulation

Both conditions were presented a Desert Survival Problem, versions of which were used by Burgoon et al. (2003), Burgoon, Chen and Twitchell (2010), and Zhou et al. (2003, 2004a) to

create truth and deception conditions. Participants were asked to imagine that a plane crash had stranded them in a desert with certain items. Participants were asked to rank which items would be most important for survival. Here the experimental conditions diverged. In the truth condition, participants were asked to write an essay justifying their rank ordering of the items. In the deception condition, participants were asked to write an essay justifying the *opposite* rank ordering of items. In both cases, participants were asked to write their essays to “someone else in the study.” That instruction was to create the impression that participants were communicating with a peer. Such an interaction is often the case in computer-mediated communication.

Participants were encouraged to be as persuasive as possible. They were told that the person writing the most persuasive essay, as judged by someone else in the study, would receive a \$50 gift card to Best Buy. Increasing motivation has been shown to lead to more deception cues (DePaulo et al., 2003). The winner would be the participant writing the essay with the highest word count. This experiment involved misleading participants, and a monetary prize was awarded, so participants were fully debriefed after all data were collected and the winner chosen.

#### Awareness and Contact Information

Participants were asked what they thought the study was about. They were asked to provide their email addresses so the experimenter could contact the winner.

#### Counterbalancing

Participants completed the demographics section first and the awareness and contact information section last. The administration of the personality testing and manipulation was counterbalanced. Thus there were four questionnaire types: truth condition, manipulation first;

truth condition, manipulation second; deception condition, manipulation first; and deception condition, manipulation second.

Appendix B contains all measurement instruments.

### Dependent Measures

Each hypothesis and research question was operationalized according to LIWC’s ability to identify and categorize linguistic features. Table 6 lists each linguistic feature investigated, the LIWC category or categories used to operationalize it, and the internal reliability of those categories as reported by LIWC’s creators. See Appendix C.

**Table 6: Operationalization of Variables**

<b>Cue Category</b>	<b>Verbal Cue to Deception</b>	<b>Operationalized as LIWC Category</b>	<b>Internal Reliability (Cronbach Alpha)</b>
<b>Word Count</b>	Verbosity	Word count	
<b>Verbal Immediacy</b>	First-person pronouns and self-references	First person singular	.62
	Group references	First person plural	.66
	Modal verbs	Discrepancy	.80
	Causation words	Causation	.88
	Qualifiers	Tentative	.87
	Present-tense verbs	Present tense	.91
	Exclusion words	Exclusive	.67
<b>Emotion Words</b>	Third-person pronouns	He/She	.75
	Positive emotion words	Positive emotion	.97
	Negative emotion words	Negative emotion	.97
<b>Linguistic Complexity</b>	Emotiveness	Adverbs	.84
	Motion verbs	Motion	.96
	Sentence length	Words per sentence	
	Word length	Words with less than six letters	

H1 concerns verbosity. Other researchers have operationalized verbosity as number of words (Burgoon et al., 2003; Burgoon & Qin, 2006; DePaulo et al., 2003; Hancock et al., 2008; Zhou et al., 2004a; Zhou et al., 2003; Zhou & Sung, 2008; Zhou & Zhang, 2008). In this study

verbosity was operationalized as LIWC's word count category, which, likewise, reports the number of words in a text.

H2 concerns (a) first-person pronouns and self-references, (b) group references, and (c) modal verbs. Other researchers have operationalized the H2a variables as the number or percentage of first person singular pronouns in a statement (Ebesu & Miller, 1994; Hancock et al., 2008; Lee et al., 2009; Pennebaker et al., 2003; Zhou et al., 2004a; Zhou et al., 2003; Zhou & Zhang, 2008). Others have operationalized the H2b variable as the number or percentage of first person plural pronouns (Ebesu & Miller, 1994; Zhou et al., 2004a; Zhou et al., 2003). Previous research operationalized the H2c variable as number or percentage of modal verbs (Zhou, 2005; Zhou et al., 2004a; Zhou et al., 2003; Zhou & Zhang, 2008). This study operationalized (a) first-person pronouns and self-references as percentage of first person pronouns as reported by LIWC's first-person singular category. The (b) group references were operationalized as LIWC's first-person plural category, which reports percentage of first-person plural pronouns. The (c) modal verbs were operationalized as LIWC's discrepancy category, which reports percentage of modal verbs.

RQ1 concerns (a) causation words, (b) qualifiers, (c) present-tense verbs, (d) exclusion words, and (e) third-person pronouns. Other researchers operationalized (a) causation words as number of causation words (Hancock et al., 2008; Lee et al., 2009). LIWC does not report number of causation words but does report percentage of causation words. Therefore, that causation category was used to operationalize causation words in this study. Previous research operationalize (b) qualifiers as the percentage of qualifying terms (Ebesu & Miller, 1994). LIWC's tentative category, which reports percentage of qualifying terms, was used to operationalize qualifiers here. Earlier studies operationalized (c) present-tense verbs as

percentage of present-tense verbs (Ebesu & Miller, 1994). LIWC's present-tense category reports percentage of present-tense verbs and was used to operationalize present-tense verbs here. Other researchers operationalized (d) exclusion words as the LIWC exclusive category (Pennebaker et al., 2003). This study did the same. For (e) third-person pronouns, previous research operationalized this variable as the number or percentage of third-person pronouns (G. D. Bond & Lee, 2005; Hancock et al., 2008; Pennebaker et al., 2003; Zhou & Sung, 2008). LIWC's third-person singular category reports percentage of third-person singular pronouns, and LIWC's third-person plural category reports percentage of third-person plural pronouns. In this study third-person pronouns were operationalized as the sum of LIWC's third-person singular and third-person plural categories.

H3 concerns (a) positive emotion words and (b) negative emotional words. Previous researchers have operationalized (a) positive emotional words as number of positive emotion words (Zhou et al., 2004a; Zhou & Zhang, 2008). LIWC does not report number of positive emotion words, but LIWC's positive emotion category reports percentage of positive emotion words. That category was used to operationalize positive emotion words here. Earlier studies operationalized (b) negative emotion words as the number or percentage of negative emotion words. LIWC's negative emotion category reports percentage of negative emotion words and was used to operationalize negative emotion words here. Studies reviewed here report emotiveness as the number of adjectives and adverbs divided by the number of nouns and verbs or by the total word count (Zhou et al., 2004a; Zhou & Zhang, 2008). LIWC does not identify nouns or adjectives, so emotiveness was operationalized here as the number of words in LIWC's adverb category divided by total word count.

H4 concerns (a) motion verbs, (b) sentence length, and (c) word length. Other researchers have operationalized (a) motion verbs as the percentage of motion words (G. D. Bond & Lee, 2005; Pennebaker et al., 2003). LIWC's motion category reports percentage of motion words and was used to operationalize motion words here. Earlier studies operationalized (b) sentence length as number of long sentences (Burgoon et al., 2003) or as message length in synchronous computer-mediated communication (Zhou & Sung, 2008). LIWC does not report number of long sentences, and this study does not concern synchronous computer-mediated communication. Therefore, LIWC's words/sentence category, which reports average number of words per sentence, was used to operationalize sentence length. Other studies have reported average word length (Zhou & Zhang, 2008), which LIWC does not calculate. Therefore, (c) word length was operationalized as LIWC's words less-than-six-letters category, which reports the percentage of words having fewer than six letters.

Appendix C contains a table of the word categories LIWC identifies, example words in each category, internal reliabilities, and how those reliabilities were calculated.

### **Procedure**

The experiment was conducted in two phases. A pilot study was conducted to identify possible problems in experiment implementation. The main study followed. The pilot study had two experimental sessions, and the main study had 24 sessions. Participants who signed up for the pilot sessions were excluded from the main study.

Participants in an online human-subjects research pool received an announcement that a study called "Persuasive Writing" was available. Potential volunteers signed on to the research portal, read the study description and consent form, and registered for an experiment session.

Sign-up indicated consent. Up to 20 participants were allowed per session because of the number of workstations in the computer laboratory where the experiments were done.

After acknowledging informed consent and signing up for the study, participants were asked to attend their scheduled sessions. The computer lab was set up with one of four letters (A through D) posted on workstations. Participants were asked to sit at the workstation of their choice. They were then directed to log on to their Scholar accounts and directed to the Scholar page where the study was administered. Scholar is a course management software platform approved by the Institutional Review Board for human subjects research. Scholar has survey functionality and is familiar to students in the research pool. Students were asked to complete a questionnaire through Scholar corresponding to the letter on their workstation.

Attendance was taken, and participants were granted credit while they completed the online questionnaires. Participation was voluntary. All participants who arrived on time were given full credit for attending whether they completed the measures or not. After all sessions were completed, participants were debriefed and notified of the prizewinner via email sent through the online human-subjects research pool portal. The prizewinner was contacted individually via an email address he provided in the questionnaire and awarded the prize.

Appendix D contains the experiment announcement, consent form, and debriefing statement.

## **Chapter 5: Results**

### **Data Compilation**

Experimental conditions in the pilot study and main study were not significantly different, so data were combined. Data were downloaded from Scholar and imported into a Microsoft Excel spreadsheet. The contest winner was identified by word count. After that, columns with information that could identify participants were deleted. Columns were created for condition and counterbalance condition. Responses to the question asking if participants guessed the nature of the experiment were read. No participant accurately determined the purpose of the study.

The essays were copied into separate text files and input into LIWC, which output a spreadsheet file. The spreadsheet cells contained values for word count and the proportion of words in each text falling into LIWC's categories.

LIWC's output data for truthful and deceptive texts were imported into JMP (SAS Institute, 1989-2007), a statistical analysis software package, for further analysis.

### **Descriptive Statistics**

Participants totaled 138 individuals, with 70 in the deception condition and 68 in the truth condition. Of those in the deception condition, 36 answered the Desert Survival Problem before the Big Five Inventory (BFI) personality test and demographic questions, and 34 answered the Desert Survival Problem after the BFI and demographic questions. Of those in the truth condition, 36 answered the Desert Survival Problem before the BFI and demographic questions, and 32 answered the Desert Survival Problem after the BFI and demographic questions. In all

conditions, participants provided their contact information and answered the question about the purpose of the study last.

Participants' mean age was 20.1 ( $SE = 1.58$ ). The oldest participant was 29, and the youngest was 18. The pool included 72 female and 65 male participants, with one participant not wanting to answer the sex question. Participants' most common major was communication ( $n = 63$ ) followed by business ( $n = 24$ ) and computer science ( $n = 9$ ). Participants had spent an average of 2.22 years ( $SE = 1.34$ ) studying at university.

### Hypothesis Testing

This research used one-tailed  $t$ -tests to detect directional effects and two-tailed  $t$ -tests to detect bi-directional effects, both with a significance level of  $p < .05$ , in the four hypotheses and the research question. Table 7 shows test results for each hypothesis.

**Table 7: Results for All Essays**

Category	Cue	Deception Mean	Truth Mean	$p$ -value	Direction
<b>Word Count</b>	Verbosity	557	530	0.497	Expected
<b>Verbal Immediacy</b>	First-person pronouns	0.580	1.38	0.004 (one-tailed)	Expected
	Group references	0.64	1.06	0.084 (one-tailed)	Expected
	Modal verbs	3.65	3.21	0.094 (one-tailed)	Expected
	Causation words	3.07	2.92	0.448	n/a
	Qualifiers	3.36	3.46	0.594	n/a
	Present-tense verbs	9.11	8.63	0.352	n/a
	Exclusion words	2.80	2.56	0.226	n/a
<b>Emotion Words</b>	Third-person pronouns	0.78	0.87	0.581	n/a
	Positive emotion words	2.99	2.82	0.389	n.s.
	Negative emotion words	1.41	1.62	0.156	Unexpected
<b>Linguistic Complexity</b>	Emotiveness	3.71	3.61	0.655	Expected
	Motion verbs	1.46	1.42	0.781	ns
	Sentence length	19.7	22.5	0.005 (one-tailed)	Expected
	Word length	15.5	16.4	0.079 (one-tailed)	Expected

## H1: Verbosity

H1 said deceivers would be more verbose than truth tellers. H1 was not supported. For each essay, LIWC output the number of words. Using *t*-tests, the word count of truthful essays was compared to that of deceptive essays. Participants in the deception condition did not write essays with significantly more words ( $M = 557, SE = 27.9$ ) than those in the truth condition ( $M = 530, SE = 28.3, p = .497$ ).

## H2: Verbal Immediacy

H2 said deceivers would use (a) a lower proportion of fewer-first person pronouns and self-references, (b) a lower proportion of group references, and (c) a higher proportion of modal verbs than truth tellers. H2a was supported. H2b and H2c were not supported at the  $p < .05$  level but approached significance. LIWC output the proportion of words in each H2 category for each essay. *T*-tests compared the proportions of cues in the truthful essays to those in the deceptive essays. Participants in the deception condition used a significantly lower proportion of first-person pronouns ( $M = .580, SE = .205$ ) than those in the truth condition ( $M = 1.38, SE = .208, p = .004$ , one-tailed). Participants in the deception condition used a higher proportion of group references ( $M = .637, SE = .216$ ) than those in the truth condition ( $M = 1.06, SE = .219$ ). The difference was marginally significant ( $p = .084$ , one-tailed). Participants in the deception condition used a higher proportion of modals ( $M = 3.65, SE = .229$ ) than those in the truth condition ( $M = 3.21, SE = .232$ ). The difference was marginally significant ( $p = .094$ , one-tailed).

## RQ1: Verbal Immediacy

RQ1 asked how deceivers and truth tellers used (a) causation words, (b) qualifiers, (c) present-tense verbs, (d) exclusion words and (e) third-person pronouns. No conclusions could be made. LIWC output the proportion of words in each RQ1 category for each essay. *T*-tests compared proportions of cues in the truthful essays to those in the deceptive essays. Participants in the deception condition did not use a significantly different proportion of causation words ( $M = 3.07, SE = .140$ ) from those in the truth condition ( $M = 2.92, SE = .142, p = .448$ ). Participants in the deception condition did not use a significantly different proportion of qualifiers ( $M = 3.36, SE = .135$ ) from those in the truth condition ( $M = 3.46, SE = .137, p = .594$ ). Participants in the deception condition did not use a significantly different proportion of present-tense verbs ( $M = 9.11, SE = .356$ ) from those in the truth condition ( $M = 8.63, SE = .361, p = .352$ ). Participants in the deception condition did not use a significantly different proportion of exclusion words ( $M = 2.80, SE = .138$ ) from those in the truth condition ( $M = 2.56, SE = .140, p = .226$ ). Participants in the deception condition did not use a significantly different proportion of third-person pronouns ( $M = .776, SE = .126$ ) from those in the truth condition ( $M = .874, SE = .127, p = .581$ ).

### H3: Emotion Words

H3 said deceivers' statements would (a) have a greater proportion of positive emotion words, (b) have a greater proportion of negative emotion words, and (c) be more emotive than those of truth tellers. H3 was not supported. LIWC output the proportion of words in each H3 category for each essay. *T*-tests compared the proportions of cues in the truthful essays to those in the deceptive essays. Participants in the deception condition did not use a significantly different proportion of positive emotion words ( $M = 2.99, SE = .137$ ) from those in the truth condition ( $M = 2.82, SE = .139, p = .389$ ). Participants in the deception condition used a lower

proportion of negative emotion words ( $M = 1.41$ ,  $SD = .106$ ) than participants in the truth condition ( $M = 1.62$ ,  $SE = .106$ ). This result was not significant ( $p = .156$ ) and opposite the expected direction. Writing in the deception condition did not show significantly different emotiveness ( $M = 3.71$ ,  $SE = .160$ ) from that in the truth condition ( $M = 3.61$ ,  $SE = .163$ ,  $p = .655$ ).

#### H4: Linguistic Complexity

H4 said deceivers' statements would have (a) a greater proportion of motion verbs, (b) shorter sentences and (c) shorter words than those of truth tellers. Results were mixed. H4a and H4c were not supported. H4b was supported. For each essay, LIWC output the proportion of words in each H4 category for each essay. *T*-tests compared the proportions of cues in the truthful essays to those in the deceptive essays. The proportion of motion verbs was not significantly different between the deception condition ( $M = 1.46$ ,  $SE = .0952$ ) and the truth condition ( $M = 1.42$ ,  $SE = .0966$ ,  $p = 0.781$ ). Participants in the deception condition produced significantly shorter sentences ( $M = 19.7$ ,  $SE = .756$ ) than those in the truth condition ( $M = 22.5$ ,  $SE = .767$ ,  $p = .005$ , one-tailed). Participants in the deception condition used shorter words ( $M = 15.5$ ,  $SE = .400$ ) than those in the truth condition ( $M = 16.4$ ,  $SE = .406$ ). The word-length difference was not significant at the  $p < .05$  level ( $p = .079$ , one-tailed), but the trend was in the expected direction.

#### Post-Hoc Analyses

Significance was not achieved for many of the hypothesized cues. Motivation causes deceivers to produce more cues to deception (DePaulo et al., 1988). Participants ranged from those who appeared highly motivated—based on how much they wrote—to those who seemed to

put very little effort into their submission. Longer essays might indicate more motivated participants. So, in the longer essays, cues to deception might be more pronounced. *T*-tests compared truthful and deceptive essays for each cue to deception for essays longer than 544 words, the mean essay length. That group comprised 63 essays, 30 truthful and 33 deceptive. Table 8 shows the results.

**Table 8: Results for Longer Essays**

Category	Cue	Deception Mean	Truth Mean	<i>p</i> -value	Direction
<b>Word Count</b>	Verbosity	766	718	0.258	Expected
<b>Verbal Immediacy</b>	First-person pronouns	0.55	1.38	0.014 (one-tailed)	Expected
	Group references	0.87	1.07	0.688	Unexpected
	Modal verbs	4.14	3.07	0.003 (one-tailed)	Expected
	Causation words	3.02	2.61	0.088	Expected
	Qualifiers	3.75	3.54	0.431	n.s.
	Present-tense verbs	8.24	7.94	0.587	n.s.
	Exclusion words	3.08	2.46	0.161	n/a
<b>Emotion Words</b>	Third-person pronouns	0.92	1.04	0.667	n.s.
	Positive emotion words	3.06	2.67	0.041 (one-tailed)	Expected
	Negative emotion words	1.54	1.65	0.586	n.s.
<b>Linguistic Complexity</b>	Emotiveness	3.98	3.41	0.016 (one-tailed)	Expected
	Motion verbs	1.44	1.45	0.952	n.s.
	Sentence length	20.3	22.4	0.031 (one-tailed)	Expected
	Word length	15.8	16.3	0.509	n.s.

### H1: Verbosity

H1 was again not supported. Of the participants writing longer essays, those in the deception condition did not use significantly more words ( $M = 766, SE = 29.3$ ) than those in the truth condition ( $M = 718, SE = 30.8, p = .258$ ). But in the longer essays, the difference in means increased from 27 words to 48 words, and the *p*-value was nearly halved from .497 to .259. The trend of deceivers using more words than truth tellers was stronger in the longer essays.

### H2: Verbal Immediacy

Analyzing only the longer essays had a mixed effect on the hypothesized verbal immediacy measures. Two of the hypothesized verbal immediacy cues were more significant for longer essays, but one was less significant for longer essays. Deceivers writing longer essays used a lower proportion of first person pronouns than truth tellers, but the significance decreased from  $p = .004$  (one-tailed) for all essays to  $p = 0.014$  (one-tailed) for the longer essays. Deceivers writing longer essays did not use group references significantly differently from truth tellers ( $p = .688$ , one-tailed). The group references results for all essays was marginally significant, with deceivers using a smaller proportion of group references ( $p = .084$ , one-tailed). Deceivers writing longer essays used a greater proportion of modal verbs than truth tellers. The significance increased from  $p = .094$  (one-tailed) for all essays to  $p = 0.003$  (one-tailed) for the longer essays.

#### RQ1: Verbal Immediacy

For all essays, none of the verbal immediacy cues in RQ1 were significantly different between truth and deceptive conditions. Repeating the  $t$ -tests with just the longer essays produced different results. Two of the verbal immediacy cues in the research question were significant or approached significance for the longer essays, but three remained insignificant. For longer texts, condition produced no significant effect on use of qualifiers (RQ1b), present-tense verbs (RQ1c), or third-person pronouns (RQ1e). For causation words (RQ1a) and exclusion words (RQ1d), differences in longer essays approached significance. Those in the deception condition ( $M = 3.02$ ,  $SE = .163$ ) used a greater proportion of causation words (RQ1a) than those in the truth condition ( $M = 2.61$ ,  $SE = .171$ ). The difference was marginally significant ( $p = .088$ ). Those in the deception condition ( $M = 3.08$ ,  $SE = .154$ ) used significantly a greater proportion of exclusion words (RQ1d) than those in the truth condition ( $M = 2.46$ ,  $SE = .140$ ,  $p = .161$ ,  $p = .007$ ).

### H3: Emotion Words

When all essays were analyzed, none of the emotion-word cues were significantly different between truthful and deceptive conditions. Repeating the *t*-tests on emotion-word cues for the longer essays resulted in two cues reaching significance. Deceivers writing longer essays ( $M = 3.06$ ,  $SE = .125$ ) used a significantly greater proportion of positive emotion words (H3b) than truth tellers ( $M = 2.67$ ,  $SE = .160$ ,  $p = .041$ , one-tailed). Deceivers writing longer essays did not use negative emotion words differently from truth tellers. Deceivers writing longer essays were significantly more emotive ( $M = 3.98$ ,  $SE = .179$ ) than truth tellers ( $M = 3.41$ ,  $SE = .188$ ,  $p = .016$ , one-tailed).

### H4: Linguistic Complexity

For all essays, two of three linguistic complexity cues were significantly different between truth and deception conditions. Repeating the analysis for only the longer essays produced different results. All three linguistic complexity cues were less significant in longer essays. For longer essays, condition produced no significant effect on the use of motion verbs ( $p = .781$ ). Nor did it produce a significant effect for all essays ( $p = .952$ ). For the longer essays, average sentence length was greater for truthful essays than for deceptive ones ( $p = .031$ ). That difference, however, was less significant than the one for all essays ( $p = .005$ ). The proportion of short words was not significantly different between truthful and deceptive conditions the longer essays ( $p = .509$ ), but it was for all essays ( $p = .079$ ).

## **Chapter 6: Discussion**

### **Participant Characteristics**

Participants fit the expected demographic characteristics a communication research pool: mostly communication majors, more female than male, in their late teens or 20s. Most participants completed all measures. Each condition included nearly the same number of participants. The distribution of participants gives little reason to suspect the data are invalid because of assignment issues.

### **Discussion**

H1: Verbosity

The experiment yielded both expected and unexpected results. H1 said deceivers would be more verbose than truth tellers because they would attempt to appear truthful and forthcoming. But results showed that deceivers averaged only 27 more words per essay than truth tellers. Because of a high variance in essay word counts (around 28 words), the difference was not significant. Previous research had been split on whether deceivers and truth tellers would use more words. In spoken or synchronous computer-mediated communication, deceivers tended to use fewer words than truth tellers because the fear of being caught lying caused deceivers to communicate less (Burgoon et al., 2003; Burgoon & Qin, 2006; DePaulo et al., 2003). In the few studies on asynchronous communication, that trend was reversed (Zhou et al., 2004a; Zhou et al., 2003). In asynchronous communication, deceivers were isolated from the listener and have time to prepare statements, so they were less anxious. Their attempts to appear truthful became more influential than fear. Using more words was one way to appear forthcoming. Results here

followed the trend of deceivers using asynchronous computer-mediated communication typing more words than truth tellers. Post-hoc analyses on the longer essays supported that trend. Participants writing longer essays showed a stronger tendency of deceivers to use more words than truth tellers.

Verbosity results may not have reached significance for either group of essays because the experiment design assumed one hour would be enough time for participants to complete all measures. But a small, unrecorded number of participants had to rush to finish. If the group of slower writers had more time, their contributions to the data pool might have increased word-count averages. That increase might have been stronger for deceivers than for truth tellers and pushed results into significance.

## H2: Verbal Immediacy

Not all verbal immediacy measures in H2 were supported at the  $p < .05$  level of statistical significance. Nevertheless, the direction of findings hinted that deceivers were somewhat less immediate than truth tellers. Deceivers used significantly fewer first-person pronouns than truth tellers. That finding supported H2a. Deceivers' use of more group references (H2b) and modals (H2c) approached significance. Previous research supported H2a more strongly than H2b and H2c. Six studies performed by three different groups of researchers found deceivers using fewer first-person pronouns than truth tellers (Ebesu & Miller, 1994; Hancock et al., 2008; Pennebaker et al., 2003; Zhou et al., 2004a; Zhou et al., 2003; Zhou & Zhang, 2008). The effect of the deception condition might be greater, in general, on first-person pronoun use than on group references and modals. This study was not designed to consider that possibility. Furthermore,

this study might not have had significant results for group references and modals because of the small sample size.

For longer essays, condition had a stronger effect on modal verbs than for all essays but a weaker effect on first-person pronouns and group references. Those results might be a meaningless statistical artifact because a data subset was analyzed. Alternatively, the results might show a meaningful trend for those cues. More data are needed.

#### RQ1: Verbal Immediacy

Prior research showed inconclusive effects of deception condition on causation words, qualifiers, present-tense verbs, exclusion words, and third-person pronouns (Ebesu & Miller, 1994; Hancock et al., 2008; Lee et al., 2009; Pennebaker et al., 2003). This study also found no significant difference between conditions for those cues. This replication might not have produced enough data to determine an effect, or no effect might be present.

The immediacy concept suggests that deceivers would use less exact language—and thus a smaller proportion of causation words, present-tense verbs, and exclusion words alongside more third-person pronouns and qualifiers—than truth tellers. Prior results on immediacy in deception have not supported that suggestion, nor do the results from this study. Deceivers writing longer essays in this study did use a greater proportion causation words and exclusion words, both signs of more immediate language. Other cues were insignificant in both the longer essays and in all essays. Those contradictory findings might indicate that cues in the research question did not all reliably indicate verbal immediacy. The cues in the research question had face validity, but different speakers might express warmth in different ways. For example, education researchers discovered immediacy cues were not the same for Chinese and American

teachers. Those researchers developed a new teacher immediacy scale for use in China (Zhang & Oetzel, 2006; Zhang, Oetzel, Gao, Wilcox, & Takai, 2007). A great deal of verbal immediacy research was on teacher immediacy (Baringer & McCroskey, 2000; Gorham, 1988; McCroskey, Richmond, Sallinen, Fayer, & Barraclough, 1995; Richmond, Gorham, & McCroskey, 1987). Teachers might display immediacy differently from other individuals. Teacher-immediacy research might not be applicable to deception situations. Cues that seemed like valid indicators of immediacy for teachers might not be accurate for deceivers.

Cues to deception are caused by competing psychological goals. Some participants may have been less immediate because of anxiety. Others could have been more immediate in an attempt to sound direct and persuasive. Participants prone to displaying cues caused by anxiety, or desire to sound truthful, would have been divided between conditions. That distribution would cause a lack of significance in the immediacy results between conditions. Personality testing, however, might allow deception-detection researchers to pinpoint which cause of cues to deception was dominant in an individual. For example, deceivers high in neuroticism might be likely to display anxiety cues and be less immediate. On the other hand, deceivers low in neuroticism could be more immediate than truth tellers. Sub-selecting for personality type when analyzing verbal immediacy deception cues would inform future studies. Finally, the results for verbal immediacy cues in RQ1 could be a statistical artifact because a subset of data was analyzed. More research on those cues is needed.

### H3: Emotion Words

Deceivers were expected to be anxious and to use a greater proportion of negative emotion words (H3a). Deceivers were expected to try to sound truthful and persuasive and use a

greater proportion of positive emotion words (H3b) and be more emotive (H3c) than truth tellers. Motivation increases the occurrence of cues to deception (DePaulo et al., 1988). The experimental design attempted to motivate participants in two ways. The first was monetary: Participants were offered a \$50 prize. The second was social: Participants were told a peer would be read the essays. The intent was to encourage participants to consider the task important and socially stressful. Analysis of truthful and deceptive essays, however, showed no significant difference in use of emotion word cues to deception. The motivation manipulations might not have succeeded.

More motivated participants were assumed to produce longer essays. Results analyzing only the longer essays showed no significant difference between conditions in the use of negative emotion words (H3a). Negative emotion in the form of negative emotion words (H3a), even for those writing more, did not leak. The first motivational factor, desire for a prize, was unlikely to cause fear and negative emotion. The second motivational factor, fear of peer judgment, might have been weak. The hypothetical “peer” who participants were told would evaluate the essays might not have been tangible enough to produce anxiety and lead to negative emotion word leakage.

For the longer essays, the effects of condition on the other two emotion cues, positive emotion words and emotiveness, were enlightening. When words were higher—possibly the result of sub-selecting the most motivated participants—the proportion of positive emotion words and emotiveness became significantly different between the truth and deception conditions. Deceivers’ attempts to sound truthful and persuasive appeared to have influenced the more motivated participants to produce essays with a greater proportion of positive emotion words and emotiveness.

A major limit in interpreting the H3 results was that the experiment did not include a manipulation check on motivation nor did it include a measure of emotion. Either of those would have allowed for sub-selection of participants motivated by money or fear. Then effects of motivation on emotion cues to deception could be tested.

#### H4: Linguistic Complexity

Deceivers used less complex language than truth tellers. Deceivers wrote both shorter sentences and words. The difference was probably the result of deceivers having a more demanding task and a higher cognitive load than truth tellers. Deceivers were, therefore, less able to produce complex language than truth tellers.

Evidence of cognitive load appeared even though participants had enough time to prepare and write their essays. Cognitive load cues appeared despite participant's apparent lack of anxiety regarding the deception task. Thus, cognitive load cues appeared to be robust in asynchronous communication, even though participants might not have felt pressured to respond quickly or felt close to the reader.

Those writing more words showed fewer differences in linguistic complexity cues between truth and deception conditions. That result could represent a ceiling effect. People writing more might have used more complex language in general, irrespective of condition, and diminished differences in complexity between conditions. Without more data, though, complexity results from those writing more words remain unexplained.

### **Limitations**

A major hindrance to data analysis in this study was the lack of a motivation-manipulation check. Essay length was used as a proxy for motivation. Time to complete

measures could have been another proxy but was not used here. A single question at the end of the measures about how much each participant wanted the prize would have measured monetary motivation. Likewise, a question about how anxious participants were about having a peer read essays would have measured social motivation. Then, instead of using a proxy like essay length, data from participants who reported being more motivated by desire for a prize or by anxiety could have been analyzed.

Another hindrance was the lack of an instrument to measure emotion. A scale, like the Positive and Negative Affect Schedule, administered at the end of the questionnaire would have shown whether differences in emotion existed between those in the truth and deception conditions (Watson, Clark, & Tellegen, 1988). If participants' emotional state were known, results on cues caused by emotional leakage could have been more readily understood.

A handful of participants thought they did not have enough time to complete the measures because their longer essays took more time to compose. Others, however, wrote very short essays and left the session before it was halfway finished. The word-count distribution was skewed. A large number of participants wrote short essays, but a small number wrote very long essays. Perhaps the task engaged and motivated only a minority of participants. Including a motivation manipulation check could help determine if that were the case.

Some hypotheses in this study were based on research that used Desert Survival Problem differently. Burgoon and Zhou used the Desert Survival Problem in interactive deception tasks. Participants knew a peer would read their responses, judge them for deception, and respond (Zhou et al., 2003). Future studies on cues to deception should consider the interactivity of deception task and the type of participant motivation, be it to win a prize or to avoid being caught lying. Other cues hypothesized in this study were based on results of research on verbal

statements or multiple statement types (DePaulo et al., 2003; Pennebaker et al., 2003; Zhou & Zhang, 2008). Future hypothesizing about cues to deception should more closely consider the task type—one-time statements or peer-interaction. Peer-interaction studies should be subdivided by how much anxiety the deceiver has about deceiving.

## Chapter 7: Conclusion

Deception is common. People lie to each other in business, friendship and romantic relationships. Deception is harmful. When institutions deceive clients, outcomes can be far-reaching. For example, the Great Recession of 2009 was caused in part by questionable lending practices. On an interpersonal level, many people have lost friends and partners because of deceit. Deception-detection research is a mature field. However, most methods analyze visual cues and/or require trained coders and a great deal of time. High-labor methods are suitable for law-enforcement and academe but not for everyday victims of deception. A tool that could help laypeople detect lies would be valuable.

Computer-mediated communication is on the rise, but deception-detection research has yet to catch up. More and more deceitful interactions take place online via synchronous and asynchronous computer-mediated communication. Much of the current deception-detection research is on visual cues, like posture and fidgeting, absent in text-only messages. Digitized text, however, lends itself well to automatic, linguistic-cue processing using natural-language processing software. LIWC, used here, is one such program. Research on cues to deception in computer-mediated communication is in its adolescence. This study sought to add to the data on linguistic cues to deception. A possible outcome of research like this study is an automatic deception-detection tool for computer-mediated messages. That tool would be popular among those suspecting they are being deceived. That group includes most of us at one time or another.

This study was a replication of a common experiment format, the Desert Survival Problem. Participants wrote truthful and deceptive essays arguing why items were useful in a survival situation. Cues to deception from the two conditions were compared. Results aligned with prior research. Deception cues were categorized based on cause: emotional leakage, verbal

immediacy, attempts to appear truthful, and cognitive load. Deception cues in the emotion category were, for the most part, absent in this study, possibly because participants were not anxious regarding the task and felt little guilt or stress. Past research identified a number of verbal immediacy cues indicating deception. In this study, those cues were significantly different between deceptive and truthful essays. Prior research was divided or lacked significant findings on other verbal immediacy cues indicating deception. This study had insignificant or unexpected results for those cues as well. Prior researchers discovered deception cues indicating deceivers' high cognitive load and attempts to appear truthful. Those cues were present here. No significant results were discovered in an unexpected direction on reliable cues to deception. Because results of this project corresponded to those of past studies, this project validated both prior research and the experimental procedures used here.

Longer essays, likely indicating more motivated writers, had more cues to deception, further supporting the existence of the motivation-impairment effect.

Linguistic cues to deception differ depending on communicator, context, motivation, anxiety, and possibly language and personality. Any replication that slightly modifies prior experimental condition sheds light on how people deceive differently in different contexts. In other experiments using the Desert Survival Problem, participants interacted with a deception-judge partner (Burgoon et al., 2010; Zhou et al., 2003). The task here had comparatively lower stakes, and was less anxiety-provoking, than those performed by Burgoon and Zhou. The difference in task type caused a difference in results. Here, cues linked to anxiety and guilt were not found, likely because the deceivers were not adequately aroused.

The results here can be explained using Zuckerman's theoretical framework. His four factors causing cues to deception were general arousal, felt emotion, cognitive load, and desire to

appear truthful (Zuckerman et al., 1981). Here, results indicated that two competing factors were at play: anxiety and desire to appear truthful. Cues caused by the desire to appear truthful were stronger in the longer essays. Those might have been written by more motivated participants. Cues caused by general arousal and felt emotion were weak, likely because participants were not anxious about the task. Cognitive load cues were significant because composing a deceitful statement is more cognitively demanding than composing a truthful one, even in the absence of anxiety.

The post-hoc analysis of longer essays was possibly new in deception-detection research. More prolific writers might display different cues to deception above and beyond differences caused by motivation. More research could determine the effect of motivation on word count and the effect of word count on linguistic cues to deception.

Both motivation and emotional state cause deception cues. More motivated people and those feeling more guilt are more likely to display cues but perhaps in different ways. Future research should include manipulation checks on motivation and emotion and should differentiate between motivation types.

The type of cue displayed, be it emotional leakage, attempts to appear truthful or cognitive load, might be associated with particular deception tasks. Future research, including meta-analyses, should attempt to determine which deception-task types produce which cues.

Only when the body of research encompasses as many deception situations and as many cues as possible will cues be reliable. This project was a step in that direction.

## References

- Anolli, L., Balconi, M., & Ciceri, R. (2003). Linguistic styles in deceptive communication: Dubitative ambiguity and elliptic eluding in packaged lies. *Social Behavior and Personality, 31*(7), 687-710.
- Baringer, D. K., & McCroskey, J. C. (2000). Immediacy in the classroom: Student immediacy. *Communication Education, 49*(2), 178-186.
- Blandon-Gitlin, I., Pezdek, K., Lindsay, D. S., & Hagen, L. (2009). Criteria-based content analysis of true and suggested accounts of events. *Applied Cognitive Psychology, 23*(7), 901-917. doi: 10.1002/Acp.1504
- Bond, C. F., Omar, A., Mahmoud, A., & Bonser, R. N. (1990). Lie detection across cultures. *Journal of Nonverbal Behavior, 14*(3), 189-204.
- Bond, G. D., & Lee, A. Y. (2005). Language of lies in prison: Linguistic classification of prisoners' truthful and deceptive natural language. *Applied Cognitive Psychology, 19*(3), 313-329. doi: 10.1002/Acp.1087
- 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: Lawrence Erlbaum.
- Buller, D. B., Burgoon, J. K., Buslig, A., & Roiger, J. (1996). Testing interpersonal deception theory: The language of interpersonal deception. *Communication Theory, 6*(3), 268-288.
- Burgoon, J. K., Blair, J. P., Qin, T. T., & Nunamaker, J. F. (2003). Detecting deception through linguistic analysis. *Intelligence and Security Informatics, Proceedings, 2665*, 91-101.

- Burgoon, J. K., Buller, D. B., White, C. H., Afifi, W., & Buslig, A. L. S. (1999). The role of conversational involvement in deceptive interpersonal interactions. *Personality and Social Psychology Bulletin*, 25(6), 669-685.
- Burgoon, J. K., Chen, F., & Twitchell, D. P. (2010). Deception and its detection under synchronous and asynchronous computer-mediated communication. *Group Decision and Negotiation*, 19(4), 345-366. doi: 10.1007/s10726-009-9168-8
- Burgoon, J. K., & Nunamaker, J. F. (2004a). Introduction: Computer-aided support of the detection of deception. *Group Decision and Negotiation*, 13(2), 107-110.
- Burgoon, J. K., & Nunamaker, J. F. (2004b). Toward computer-aided support for the detection of deception. *Group Decision and Negotiation*, 13(1), 1-4.
- Burgoon, J. K., & Qin, T. T. (2006). The dynamic nature of deceptive verbal communication. *Journal of Language and Social Psychology*, 25(1), 76-96. doi: 10.1177/026192x05284482
- Cohen, J. (1988). A power primer. *Psychological Bulletin*, 112(1), 155-160.
- Daft, R. L., & Lengel, R. H. (1984). Information richness - a new approach to managerial behavior and organization design. *Research in Organizational Behavior*, 6, 191-233.
- Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science*, 32(5), 554-571.
- DePaulo, B. M., Kirkendol, S. E., Kashy, D. A., 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(3), 177-202.

- DePaulo, B. M., Lanier, K., & Davis, T. (1983). Detecting the deceit of the motivated liar. *Journal of Personality and Social Psychology, 45*, 1096-1103.
- 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. doi: 10.1037//0033-2909.129.1.74
- deTurck, M. A., Feeley, T. H. & Roman, L. A. . (1997). Vocal and visual cue training in behavioral lie detection. *Communication Research Reports, 14*(3), 249-259.
- deTurck, M. A., Harsztrak, J. J., Bodhorn, D. J., & Texter, L. A. (1990). The effects of training social perceivers to detect deception from behavioral cues. *Communication Quarterly, 38*(2), 189-199.
- Ebesu, A. S., & Miller, M. D. (1994). Verbal and nonverbal behaviors as a function of deception type. *Journal of Language and Social Psychology, 13*(4), 418-442.
- Ekman, P. (2001). *Telling Lies: Clues to Deceit in the Marketplace, Politics and Marriage* (3rd ed.). New York: Norton.
- Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origin, usage and coding. *Semiotica, 1*, 49-98.
- Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 29*(3), 288-298.
- Ekman, P., & Friesen, W. V. (1974). Detecting deception from the body or face. *Journal of Personality and Social Psychology, 29*(3), 288-298.
- Ekman, P., Friesen, W. V., & O'Sullivan, M. (1988). Smiles when lying. *Journal of Personality and Social Psychology, 54*(3), 414-420.

- Ekman, P., Friesen, W. V., & Osullivan, M. (1988). Smiles When Lying. *Journal of Personality and Social Psychology*, 54(3), 414-420.
- Ekman, P., Friesen, W. V., Osullivan, M., Chan, A., Diacoyannitarlatis, I., Heider, K., . . . Tzavaras, A. (1987). Universals and cultural-differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology*, 53(4), 712-717.
- Ekman, P., O'Sullivan, M., & Frank, M. G. (1999). A few can catch a liar. *Psychological Science*, 10(3), 263-266.
- Ennis, E., Vrij, A., & Chance, C. (2008). Individual differences and lying in everyday life. *Journal of Social and Personal Relationships*, 25(1), 105-118. doi: Doi 10.1177/0265407507086808
- Feeley, T. H., & Young, M. J. (1992). Humans as lie detectors: Some more second thoughts. *Communication Quarterly*, 46(2), 109-126.
- Global Deception Research Team. (2006). A world of lies. *Journal of Cross-Cultural Psychology*, 37(1), 60-74. doi: Doi 10.1177/0022022105282295
- Gorham, J. (1988). The relationship between verbal teacher immediacy behaviors and student learning. *Communication Education*, 37(1), 40-53.
- Gudjonsson, G. H., & Sigurdsson, J. F. (2004). The relationship of suggestibility and compliance with self-deception and other-deception. *Psychology, Crime and Law*, 19(4), 447-453.
- Hancock, J. T., Curry, L. E., Goorha, S., & Woodworth, M. (2008). On lying and being lied to: A linguistic analysis of deception in computer-mediated communication. *Discourse Processes*, 45(1), 1-23. doi: 10.1080/01638530701739181

- Hancock, J. T., Woodworth, M. T., & Goorha, S. (2010). See no evil: The effect of communication medium and motivation on deception detection. *Group Decision and Negotiation*, 19(4), 327-343. doi: 10.1007/s10726-009-9169-7
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory--Versions 4a and 54*. Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research.
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114-158). New York, NY: Guilford Press.
- Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, 88(1), 67-85.
- Kashy, D. A., & DePaulo, B. M. (1996). Who lies? *Journal of Personality and Social Psychology*, 70(5), 1037-1051.
- Lee, C. C., Welker, R. B., & Odom, M. D. (2009). Features of computer-mediated, text-based messages that support automatable, linguistics-based indicators for deception detection. *Journal of Information Systems*, 23(1), 19-24.
- Levine, T. R., Kim, R. K., Park, H. S., & Hughes, M. (2006). Deception detection accuracy is a predictable linear function of message veracity base-rate: A formal test of Park and Levine's probability model. *Communication Monographs*, 73(3), 243-260. doi: 10.1080/03637750600873736
- Mccroskey, J. C., Richmond, V. P., Sallinen, A., Fayer, J. M., & Barraclough, R. A. (1995). A cross-cultural and multi-behavioral analysis of the relationship between nonverbal immediacy and teacher evaluation. *Communication Education*, 44(4), 281-291.

- Mehrabian, A. (1966). Attitudes in relation to the forms of communicator-object relationship in spoken communication. *Journal of Personality, 34*, 80-93.
- Montgomery, B. M. (1981). Verbal immediacy as a behavioral indicator of open communication content. *Communication Quarterly, 30*(1), 28-34.
- Neuliep, J. W. (1997). A cross-cultural comparison of teacher immediacy in American and Japanese college classrooms. *Communication Research, 24*(4), 431-451.
- O'Sullivan, P. B., Hunt, S. K., & Lippert, L. R. (2004). Mediated immediacy - A language of affiliation in a technological age. *Journal of Language and Social Psychology, 23*(4), 464-490. doi: 10.1177/0261927x04269588
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic Inquiry and Word Count: LIWC 2001*. Mahwah, NJ: Lawrence Erlbaum.
- Pennebaker, J. W., Newman, M. L., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin, 29*(5), 665-675. doi: 10.1177/0146167203251529
- Porter, S., ten Brinke, L., & Wallace, B. (2011). Secrets and lies: Involuntary leakage in deceptive facial expressions as a function of emotional intensity. *Journal of Nonverbal Behavior, Online First (n/a)(n/a)*.
- Richmond, V. P., Gorham, J., & McCroskey, J. C. (1987). The relationship between selected immediacy behaviors and cognitive learning *Communication Yearbook 10* (pp. 574-590). Beverly Hills, CA: Sage.
- Robinson, R. Y., Richmond, V. P. (1995). Validity of the verbal immediacy scale. *Communication Research Reports, 12*(1), 80-83.
- SAS Institute, I. (1989-2007). JMP, Version 7. Cary, NC.

- Seiter, J. S., Brusckhe, J., & Bai, C. (2002). The acceptability of deception as a function of perceiver's culture, deceiver's intention, and deceiver-deceived relationship. *Western Journal of Communication, 66*, 158-180.
- Siegel, L. J. (1986). *Criminology* (2nd ed.). St. Paul: West Pub. Co.
- Sporer, S. L. (1997). The less travelled road to truth: Verbal cues in deception detection in accounts of fabricated and self-experienced events. *Applied Cognitive Psychology, 11*(5), 373-397.
- Steller, M., & Kohnken, G. (1989). Criteria-based content analysis. In D. C. Raskin (Ed.), *Psychological Methods in Criminal Investigation and Evidence*. New York: Springer-Verlag.
- Twitchell, D. P., Forsgren, N., Wiers, K., Burgoon, J. K., & Nunamaker, J. F. (2005). Detecting deception in synchronous computer-mediated communication using speech act profiling. *Intelligence and Security Informatics, Proceedings, 3495*, 471-478.
- Undeutsch, U. (1967). Beurteilung der Glaubhaftigkeit von Aussagen [Assessing the credibility of statements]. In U. Undeutsch (Ed.), *Handbuch der Psychologie. Bd. 11: Forensische Psychologie* (pp. 26-181). Goettingen: Hogrefe.
- Voutilainen, A. (2000). Helsinki taggers and parsers for English. In J. M. Kirk (Ed.), *Corpora Galore: Analysis and techniques in describing English*. Amstersam & Atlanta: Rodopi.
- Vrij, A., Akehurst, L., Soukara, S., & Bull, R. (2004). Detecting deceit via analyses of verbal and nonverbal behavior in children and adults. *Human Communication Research, 30*(1), 8-41.
- 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.

- Walther, J. B. (1992). Interpersonal effects in computer-mediated interaction: A relational perspective. *Communication Research* 19, 37.
- Walther, J. B., Anderson, J. F., & Park, D. W. (1994). Interpersonal effects in computer-mediated interaction - A metaanalysis of social and antisocial communication. *Communication Research*, 21(4), 460-487.
- Walther, J. B., Loh, T., & Granka, L. (2005). Let me count the ways - The interchange of verbal and nonverbal cues in computer-mediated and face-to-face affinity. *Journal of Language and Social Psychology*, 24(1), 36-65. doi: Doi 10.1177/0261927x04273036
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063-1070.
- Weaver III, J. (2005). Mapping the links between personality and communicator style. *Individual Differences Research*, 3(1), 59-70.
- World Bank. (2011). World Internet Usage. *World Development Indicators*. Retrieved from <http://data.worldbank.org/indicator>
- Zhang, Q., & Oetzel, J. G. (2006). Constructing and validating a teacher immediacy scale: A Chinese perspective. *Communication Education*, 55(2), 218-241.
- Zhang, Q., Oetzel, J. G., Gao, X., Wilcox, R. G., & Takai, J. (2007). Teacher immediacy scales: Testing for validity across cultures. *Communication Education*, 56(2), 228-248.
- Zhou, L. (2005). An emperical investigation of deception behavior in instant messaging. *IEEE transactions on professional communication*, 48(2), 147-160.

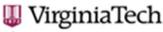
- Zhou, L., Burgoon, J. K., Nunamaker, J. F., & Twitchell, D. (2004a). Automating linguistics-based cues for detecting deception in text-based asynchronous computer-mediated communication. *Group Decision and Negotiation*, 13(1), 81-106.
- Zhou, L., Burgoon, J. K., & Twitchell, D. P. (2003). *A longitudinal analysis of language behavior of deception in e-mail*. Paper presented at the 1st NSF/NIJ conference on intelligence and security informatics, Tucson, AZ, USA.
- Zhou, L., Burgoon, J. K., Twitchell, D. P., Qin, T. T., & 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., Burgoon, J. K., Zhang, D. S., & Nunamaker, J. F. (2004c). Language dominance in interpersonal deception in computer-mediated communication. *Computers in Human Behavior*, 20(3), 381-402. doi: 10.1016/S0747-5632(03)00051-7
- Zhou, L., & Sung, Y. W. (2008). *Cues to deception in online Chinese groups*. Paper presented at the Proceedings of the 41st Hawaii International Conference on System Sciences (HICSS 2008), Waikoloa, Big Island, Hawaii.
- Zhou, L., & Zhang, D. S. (2008). Following linguistic footprints: Automatic deception detection in online communication. *Communications of the ACM*, 51(9), 119-122. doi: 10.1145/1378727.1389972
- Zuckerman, M., Defrank, R. S., Hall, J. A., Larrance, D. T., & Rosenthal, R. (1979). Facial and vocal cues of deception and honesty. *Journal of Experimental Social Psychology*, 15(4), 378-396.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 14, pp. 1-59). New York: Academic Press.

# Appendices

## Appendix A: Institutional Board Approval Letters

### Appendix A.1: Approval Letter



Office of Research Compliance  
Institutional Review Board  
2000 Kraft Drive, Suite 2000 (0497)  
Blacksburg, Virginia 24060  
540/231-4606 Fax 540/231-0959  
e-mail: irb@vt.edu  
Website: www.vt.edu/irb

**MEMORANDUM**  
**DATE:** November 28, 2011  
**TO:** Douglas Cannon, Caroline Connell  
**FROM:** Virginia Tech Institutional Review Board (FWA00000572, expires May 31, 2014)  
**PROTOCOL TITLE:** Persuasive Writing  
**IRB NUMBER:** 11-974

*Director Approval 02/08/11, the Virginia Tech IRB Chair, Dr. David S. Moore, approved the new protocol for the above-mentioned research protocol.*

*This approval provision pertains to begin the human subject activities outlined in the IRB-approved protocol and supporting documents.*

*Please to revoke from the approved protocol under supporting documents must be submitted to the IRB as an amendment request and approved by the IRB prior to the implementation of any changes, regardless of how minor, except where necessary to continue approved human subjects to the protocol. Report promptly to the IRB any injuries or other unexpected or adverse events involving data or items to which research subjects or others.*

*All investigators listed above are required to comply with the regulatory requirements outlined at [www.fda.gov/oc/ohrt/IRBguide.htm](http://www.fda.gov/oc/ohrt/IRBguide.htm) before review of the continuation of your protocol.*

**PROTOCOL INFORMATION:**  
 Approved on: **Expedited, under 45 CFR 46.110 category(ies) 7**  
 Protocol Approval Date: **11/22/2011**  
 Protocol Expiration Date: **11/22/2012**  
 Continuation Review Due Date: **11/22/2012**

*\*Write a Checklist Review application to due to the IRB office if events related to the protocol occur under this protocol, including data analysis, are to continue beyond the Protocol Expiration Date.*

**FEDERALLY FUNDED RESEARCH REQUIREMENTS:**  
 Per federal regulations, 45 CFR 46.103(i), the IRB is required to compare all federally funded grant proposals / work statements to the IRB protocol(s) which cover the human research activities included in the proposal / work statement before funds are released. Note that this requirement does not apply to Exempt and Interim IRB protocols, or grants for which VT is not the primary awardee.

*The table on the following page indicates whether grant proposals are related to this IRB protocol, and which of the listed proposals, if any, have been compared to this IRB protocol, if required.*

IRB Number 11-974      page 2 of 2      Virginia Tech Institutional Review Board

Date*	OSP Number	Sponsor	Grant Comparison Conducted?

*\*Date this proposal number was compared, assessed as not requiring comparison, or comparison information was received.*

*If this IRB protocol is to cover any other grant proposals, please contact the IRB office (540/231-4606) immediately.*

on file  
 Department of Communication & Technology

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY  
*An equal opportunity, affirmative action institution*

*Invent the Future*

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY  
 An equal opportunity, affirmative action institution



## Appendix B: Experimental Measures

### Persuasive Writing Study: Questionnaire

(Introduction and Instructions)

Please complete the following questionnaire.

The participant putting the most effort into the essay task, as judged by someone else in the study, will be awarded a **PRIZE WORTH \$50!**

Part 1: Basics

(BOTH conditions)

1. Please enter your age \_\_\_\_.
2. Please select your gender:
  - A. Male
  - B. Female
  - C. Don't want to say
3. Please enter your major or majors: \_\_\_\_.
4. Please select the number of years you have studied at the university level:
  - A. less than 1
  - B. 1
  - C. 2
  - D. 3
  - E. 4
  - F. 5
  - G. 5 or more

Part 2: (Personality)

(BOTH Conditions. Each statement is followed by five options: Disagree strongly, Disagree a little, Neither agree nor disagree, Agree a little, Agree strongly)

About You. Enter your answers to the items below. Each begins with the statement:

YOU ARE SOMEONE WHO...

1. Tends to find fault with others
2. Worries a lot
3. Has an assertive personality

4. Is talkative
5. Has few artistic interests
6. Is ingenious, a deep thinker
7. Is outgoing, sociable
8. Is considerate and kind to almost everyone
9. Makes plans and follows through with them
10. Tends to be disorganized
11. Is generally trusting
12. Generates a lot of enthusiasm
13. Gets nervous easily
14. Does a thorough job
15. Is helpful and unselfish with others
16. Is a reliable worker
17. Is sometimes shy, inhibited
18. Likes to cooperate with others
19. Is sometimes rude to others
20. Has an active imagination
21. Is sophisticated in art, music, or literature
22. Values artistic, aesthetic experiences
23. Can be somewhat careless
24. Can be cold and aloof
25. Is original, comes up with new ideas
26. Does things efficiently
27. Tends to be quiet
28. Is curious about many different things
29. Is full of energy
30. Is inventive
31. Can be tense
32. Is emotionally stable, not easily upset
33. Likes to reflect, play with ideas
34. Perseveres until the task is finished
35. Remains calm in tense situations
36. Prefers work that is routine
37. Is reserved
38. Is relaxed, handles stress well.
39. Starts quarrels with others
40. Can be moody
41. Is easily distracted
42. Has a forgiving nature
43. Tends to be lazy

44. Is depressed, blue

Part 3(A): Essay (Desert Survival Problem, TRUTH Condition)

Part 3(B): Essay (Desert Survival Problem, DECEPTION Condition)

Due to a request by legal counsel for Human Synergistics, “Appendix B: Experimental Measures” has been removed.

“The Desert Survival Situation™ is copyrighted and trademarked by Human Synergistics. Please visit [www.humansynergistics.com](http://www.humansynergistics.com) for more information.”

### Appendix C: LIWC 2007 Output Variable Information

Category	Abbrev	Examples	Words in category	Validity (judges)	Alpha: Binary/raw
<b>Linguistic Processes</b>					
Word count	wc				
words/sentence	wps				
Dictionary words	dic				
Words>6 letters	sixltr				
Total function words	funct		464		.97/.40
Total pronouns	pronoun	I, them, itself	116		.91/.38
Personal pronouns	ppron	I, them, her	70		.88/.20
1st pers singular	i	I, me, mine	12	.52	.62/.44
1st pers plural	we	We, us, our	12		.66/.47
2nd person	you	You, your, thou	20		.73/.34
3rd pers singular	shehe	She, her, him	17		.75/.52
3rd pers plural	they	They, their, they'd	10		.50/.36
Impersonal pronouns	ipron	It, it's, those	46		.78/.46
Articles	article	A, an, the	3		.14/.14
Common verbs	verb	Walk, went, see	383		.97/.42
Auxiliary verbs	auxverb	Am, will, have	144		.91/.23
Past tense	past	Went, ran, had	145	.79	.94/.75
Present tense	present	Is, does, hear	169		.91/.74
Future tense	future	Will, gonna	48		.75/.02
Adverbs	adverb	Very, really, quickly	69		.84/.48
Prepositions	prep	To, with, above	60		.88/.35
Conjunctions	conj	And, but, whereas	28		.70/.21
Negations	negate	No, not, never	57		.80/.28
Quantifiers	quant	Few, many, much	89		.88/.12
Numbers	number	Second, thousand	34		.87/.61
Swear words	swear	Damn, piss, fuck	53		.65/.48
<b>Psychological Processes</b>					
Social processes	social	Mate, talk, they,	455		.97/.59

Category	Abbrev	Examples	Words in category	Validity (judges)	Alpha: Binary/raw
		child			
<b>Family</b>	family	Daughter, husband, aunt	64	.87	.81/.65
<b>Friends</b>	friend	Buddy, friend, neighbor	37	.70	.53/.12
<b>Humans</b>	human	Adult, baby, boy	61		.86/.26
<b>Affective processes</b>	affect	Happy, cried, abandon	915		.97/.36
<b>Positive emotion</b>	posemo	Love, nice, sweet	406	.41	.97/.40
<b>Negative emotion</b>	negemo	Hurt, ugly, nasty	499	.31	.97/.61
<b>Anxiety</b>	anx	Worried, fearful, nervous	91	.38	.89/.33
<b>Anger</b>	anger	Hate, kill, annoyed	184	.22	.92/.55
<b>Sadness</b>	sad	Crying, grief, sad	101	.07	.91/.45
<b>Cognitive processes</b>	cogmech	cause, know, ought	730		.97/.37
<b>Insight</b>	insight	think, know, consider	195		.94/.51
<b>Causation</b>	cause	because, effect, hence	108	.44	.88/.26
<b>Discrepancy</b>	discrep	should, would, could	76	.21	.80/.28
<b>Tentative</b>	tentat	maybe, perhaps, guess	155		.87/.13
<b>Certainty</b>	certain	always, never	83		.85/.29
<b>Inhibition</b>	inhib	block, constrain, stop	111		.91/.20
<b>Inclusive</b>	incl	And, with, include	18		.66/.32
<b>Exclusive</b>	excl	But, without, exclude	17		.67/.47
<b>Perceptual processes</b>	percept	Observing, heard, feeling	273		.96/.43
<b>See</b>	see	View, saw, seen	72		.90/.43
<b>Hear</b>	hear	Listen, hearing	51		.89/.37
<b>Feel</b>	feel	Feels, touch	75		.88/.26
<b>Biological processes</b>	bio	Eat, blood, pain	567	.53	.95/.53
<b>Body</b>	body	Cheek, hands, spit	180		.93/.45
<b>Health</b>	health	Clinic, flu, pill	236		.85/.38
<b>Sexual</b>	sexual	Horny, love, incest	96		.69/.34
<b>Ingestion</b>	ingest	Dish, eat, pizza	111		.86/.68
<b>Relativity</b>	relativ	Area, bend, exit, stop	638		.98/.51

Category	Abbrev	Examples	Words in category	Validity (judges)	Alpha: Binary/raw
<b>Motion</b>	motion	Arrive, car, go	168		.96/.41
<b>Space</b>	space	Down, in, thin	220		.96/.44
<b>Time</b>	time	End, until, season	239		.94/.58
<b>Personal Concerns</b>					
<b>Work</b>	work	Job, majors, xerox	327		.91/.69
<b>Achievement</b>	achieve	Earn, hero, win	186		.93/.37
<b>Leisure</b>	leisure	Cook, chat, movie	229		.88/.50
<b>Home</b>	home	Apartment, kitchen, family	93		.81/.57
<b>Money</b>	money	Audit, cash, owe	173		.90/.53
<b>Religion</b>	relig	Altar, church, mosque	159		.91/.53
<b>Death</b>	death	Bury, coffin, kill	62		.86/.40
<b>Spoken categories</b>					
<b>Assent</b>	assent	Agree, OK, yes	30		.59/.41
<b>Nonfluencies</b>	nonflu	Er, hm, umm	8		.28/.23
<b>Fillers</b>	filler	Blah, I mean, you know	9		.63/.18

**Words in category** refers to the number of different dictionary words that make up the variable category. **Validity judges** reflect the simple correlations between judges' ratings of the category with the LIWC variable (from Pennebaker & Francis, 1996). **Alphas** refer to the Cronbach alphas for the internal reliability of the specific words within each category. The binary alphas are computed on the occurrence/non-occurrence of each dictionary word whereas the raw or uncorrected alphas are based on the percentage of use of each of the category words within the texts. All alphas were computed on a sample of 2800 randomly selected text files from our language corpus.

The LIWC dictionary generally arranges categories hierarchically. For example, all pronouns are included in the overarching category of function words. The category of pronouns is the sum of personal and impersonal pronouns. There are some exceptions to the hierarchy rules:

**Common verbs** are not included in the function word category. Similarly, common verbs (as opposed to auxiliary verbs) that are tagged by verb tense are included in the past, present, and future tense categories but not in the overall function word categories.

**Social processes** include a large group of words (originally used in LIWC2001) that denote social processes, including all non-first-person-singular personal pronouns as well as verbs that suggest human interaction (talking, sharing).

**Perceptual processes** include the entire dictionary of the Qualia category (which is a separate dictionary), which includes multiple sensory and perceptual dimensions associated with the five senses.

## **Appendix D: Experiment Announcement, Recruitment and Consent Form, and Debriefing Statement**

### Appendix D.1: Announcement

#### **Persuasive Writing: Online Recruiting Announcement**

This study will investigate persuasive writing techniques. Participants will complete a questionnaire and write short persuasive texts in a computer lab. The participant writing the most persuasive essay will be awarded a \$50 gift card to Best Buy.

One hour of research credit will be allotted to each person participating in the study.

Participation will not take more than one hour. This study has multiple sessions. Sign up for one that fits your schedule, and be sure to arrive on time. The door will be locked after the study has begun. If you're late, you won't be allowed in, won't be granted participation credit, and won't be eligible for the \$50 gift card.

Information on how to obtain study results will be given to each participant after completion of the study.

All participants' privacy will be protected. No individuals will be identified in research reports. Information used in the study will be anonymous. Emotional discomfort, if any, is minor and would not be greater than that experienced in everyday life. Students under the age of 18 are allowed to participate in this study.

### Appendix D.2: Recruitment and Consent Form

#### **Persuasive Writing Study: Recruitment and Consent Form Dr. Douglas Cannon and Ms. Caroline Connell**

##### **I. Purpose of this Research**

This research study is to investigate persuasive language. You are among a maximum of four hundred Virginia Tech students participating.

##### **II. Procedures**

After you consent to and sign up for the study, you will be asked to attend an experiment session. When you arrive, you will be asked to fill out a series of questionnaires and perform a writing task. The entire study is expected to take no more than an hour to complete.

### **III. Risks**

Discomfort or risk, if any, will be minimal and no greater than what can be expected in everyday life.

### **IV. Benefits**

The participant who we judge wrote the most persuasive essay will receive a \$50 gift card to Best Buy. You will receive one hour of research credit for participating in the study. Participants will learn about communication research and have access to study results. Benefits are not guaranteed.

### **V. Extent of Anonymity Confidentiality**

Your privacy will be protected throughout this process, and all the information you submit will be kept confidential. Identifying information will be used to award your benefits, but will not be analyzed as part of the study's data. Once benefits are awarded, names will be deleted from the data. Data will be kept on an encrypted hard drive and/or on a CD stored in 106 Shanks. The Institutional Review Board (IRB) may view this study's collected data for auditing purposes. The IRB oversees the protection of human participants involved in research.

### **VI. Compensation**

If you complete the study, you will be eligible for the \$50 gift card to Best Buy. By participating in this study, you will receive Research Participation credit for the course you arrange through the Department of Communication Research Participation System.

### **VII. Freedom to Withdraw**

You are free to withdraw from a study at any time without penalty. You are free not to answer any questions or respond to experimental situations that you choose without penalty.

### **VIII. Participant's Responsibilities**

I have read this document and voluntarily agree to participate in this study. I will sign up for and attend a study session on time. I understand that if I sign up but do not attend a session, I will not be granted research credit nor will I be eligible for the \$50 gift card to Best Buy.

### **IX. Participant's Permission**

I have read this Recruitment and Consent Form and understand the conditions of this study. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent by signing up for this study.

Should I have any further questions about this study and/or research subjects' rights, I may contact:

Investigator(s) Telephone/e-mail:

Caroline Connell 540.231.8333/connellc@vt.edu

Faculty Advisor Telephone/e-mail:

Dr. Douglas Cannon 540-231-2331/dfcannon@vt.edu

Chair, Virginia Tech Institutional Review Telephone/e-mail and address

David M. Moore 540-231-4991/moored@vt.edu

Board for the Protection of Human Subjects

Office of Research Compliance

2000 Kraft Drive, Suite 2000 (0497)

Blacksburg, VA 24060

Appendix D.3: Debriefing Statement

### **Persuasive Writing Study: Debriefing Statement**

Thank you for participating in this study!

The essay task had two versions. One asked participants to argue for their ranking of the 15 desert survival items; the other asked participants to argue for the opposite ranking of the 15 items. We are looking at the way people use language differently in the two tasks.

You were asked to target your essays to “someone else in the study” and told that “someone else in the study” would judge the prize winner. That “someone else” was the experimenter. This study focuses on peer-to-peer communication, so it was necessary to create the impression that you were writing to a peer.

The \$50 Gift Card to Best Buy was awarded to the participant writing the most words. Those arguing for their ranking, or the opposite ranking, had equal opportunity to win.

Once again, thank you for your participation. If you have any questions, or would like to find out about the results of this study, please feel free to contact Caroline Connell (e-mail: connellc@vt.edu).