An Exploratory Study into Deception Detection in Text-based Computer-Mediated Communication*

Lina Zhou
Department of Information Systems
University of Maryland, Baltimore County
Baltimore, MD 21250, USA
zhoul@umbc.edu

Douglas P. Twitchell
Department of MIS
University of Arizona
Tucson, AZ 85721
dtwitchell@cmi.arizona.edu

Tiantian Qin
Department of MIS
University of Arizona
Tucson, AZ 85721
tqin@cmi.arizona.edu

Judee K. Burgoon
Department of MIS
University of Arizona
Tucson, AZ 85721
judee@cmi.arizona.edu

Jay F. Nunamaker, Jr.
Department of MIS
University of Arizona
Tucson, AZ 85721
jnunamaker@cmi.arizona.edu

Abstract

Deception is an everyday occurrence across all communication media. The expansion of the Internet has significantly increased the amount of textual communication received and stored by individuals and organizations. Inundated with massive amounts of textual information transmitted through Computer-mediated Communication, CMC, people remain largely unsuccessful and inefficient in detecting those messages that may be deceptive. Creating an automated tool that could help people flag the possible deceptive messages in CMC is desirable, but first it is necessary to understand cues used to deceive in textual instances. This study focuses on the identification of deceptive cues deceivers use in a textual CMC environment. 30 dyads (n =14 truthful, n = 16 deceptive) were able to complete the Desert Survival Problem. Findings have demonstrated significant differences between the content within truthful and deceptive messages. Several cues were also found to be significantly more present when deceivers write messages.

1. Introduction

The Internet has allowed for users to distribute and exchange vast amounts of information in a quick manner, thus people are seeing a rapid increase in textual information received on a daily basis. It is imperative for people to be able to not only filter this information but to be able to distinguish which information is misleading or deceptive. Research regarding the ability to identify deception in textual information has been sparse at best. Focusing research efforts to establish a set of cues that can aid in the detection of deception via textual means will not only advance deceptive research but also can benefit the general public, law officials, and military. If cues can be identified for textual deceptive messages, an automated tool based on at least a portion of those cues can be built in order to aid people in deciding if the information they have received is truthful or deceptive. The ability to identify deceptive information in textual forms can decrease an organization's economic costs, labor costs, and reduce the time spent following deceptive leads or information. These reasons become even more pressing in situations where time is of essence (e.g., kidnappings, national security, business deals). This research project aids in expanding the knowledge about deceptive cues over text based systems and proposes future areas for research.

2. Literature Review

Deception is defined as the active transmission of messages and information to create a false conclusion [2]. Messages that are unknowingly sent by a sender are not considered deceptive, as there is no intention to deceive. Most people have experienced deception of one form or another from outright lies and fabrications to little "white" lies [11].

*Portions of this research were supported by funding from the U. S. Air Force Office of Scientific Research under the U. S. Department of Defense University Research Initiative (Grant #F49620-01-1-0394). The views, opinions, and/or findings in this report are those of the authors and should not be construed as an official Department of Defense position, policy, or decision.

The majority of information exchanged on a daily basis normally involves some level of deceit and is done using rich media (e.g., face-to-face, voice). Therefore, research has mainly focused on richly mediated communication channels. However, more and more people are turning to using different modalities to transfer information. As this switch in modalities becomes more of the norm for relaying information, deception within these modalities should increase. Therefore, it is important to understand the benefits and drawbacks of detecting deception over less rich mediums such as e-mail systems.

2.1. Media Richness Theory

Media richness is measured on a continuum against four criteria: feedback (asking questions or making corrections), multiple cues (transmitting voice inflection, body language, numbers, and symbols), language variety (range of meaning that can be encoded in language symbols), and personal focus (transmitting feelings and emotions) [9]. Deceptive research has focused on mainly media rich channels (e.g., FtF interactions, voice) with small amounts of research being conducted in the textual arena therefore the set of cues associated with predicting deception are only those within these highly media rich channels (for a meta-analysis of deceptive literature, see [12]). The highest predictors of deception are nonverbal cues such as pupil dilation, blinking, segmentation of body behavior and pitch [30] but even at best, the ability to correctly identify deception in rich media using these cues is little more than chance (e.g., see [13]). Several possible reasons have been given for the lack of detection accuracy, including truth bias, visual distraction, situational familiarity, and idiosyncratic behaviors that cloud true deception cues [23]. Yet communication media low in richness (e.g., chat, e-mail, computer documents) do not lend themselves to transmitting the most salient non-verbal cues.

In terms of richness, text-based CMC is classified as being between the telephone and written communication in its ability to transmit information [7]. Being that it is one of the least rich media, it does not have the same ability to transmit information, meaning, and emotion as does richer media, such as face-to-face [9, 10]. Due to the nature of e-mail being a less rich medium, deception is claimed to be more difficult to detect over e-mail than deceptive messages transmitted via richer media [16]. This may be attributable to the fact that there is an inability to transfer identifiable nonverbal cues to textual situations. However, if users of text-based systems perceive the channel being used as able to convey richer information than it really does, they may use the system in such a way that begins to mimic the use of more rich systems.

2.2. Channel Expansion Theory

Channel Expansion Theory expands media richness theory to include the experience senders or receivers have with a channel or media, the topic of the communication, the organizational context, and the other parties in the communication [8]. The more experience the senders and the receivers have with each of these domains, the richer they find the media they are using. The heightened perceived richness may have one of two results: those more experienced with text-based CMC will find it more rich and likely transmit more deception cues, or they may have a greater ability to strategically hide possible deception cues. People hold expectations about the discourse of others and assume that the others would satisfy the interaction demands of sufficiently complete, truthful, clear and relevant to the current topics [17]. In other words, deceivers have to make the conversations or messages sufficiently complete, truthful, clear and relevant, to the extent that the deceivers believe their receivers are satisfied with the quantity of the conversations or messages. Therefore, deceptive senders may create patterns of linguistic deceptive cues that convey deception over textual means much the same way as deceptive cues are done in FtF situations.

2.3. Interpersonal Deception Theory

Interpersonal Deception Theory, IDT, attempts to explain deception from an interpersonal conversational perspective and not strictly from any physiological venue [2]. IDT posits that within the context and relationship of the sender and receiver of deception, the deceiver will both engage in strategic modifications of behavior in response to the receiver's suspicions and will display nonstrategic leakage cues or indicators of deception. Tests of this theory have confirmed the existence of brevity and nonimmediacy along with other identifiable cues, which may be useful in detecting deception within any modality [3, 4, 5]. Because the theory does not try to account for just one modality, nor does it focus on only physiological or nonverbal indicators, many of its findings are applicable to lower richly mediated channels.

2.4. Model of Deceptive Communication

A model of deceptive communication proposed by Miller and Stiff [23] suggests that people are motivated by the social context of the communication to deceive. Social context includes 1) familiarity of the parties (e.g., the length and depth of the relationship), 2) context of relationship (e.g., familial, organizational, or friendship), and 3) status of the relationship (e.g., superior-subordinate or equal). The motivation for the deceiver to deceive affects

the contents of the deceptive message and its potential to deceive. The ability to detect the deceptive messages is affected by the receiver's motivation to detect. Both the deceiver's and the receiver's motivations are in turn affected by the overall social context in which the communication is taking place. This model was later on expanded to include the effect of the communication medium on the relationship between the deceiver and the message the deceiver is sending [15, 16]. The new deceptive communication model allows an avenue for the studying of deception in text-based communication.

2.5. Statement Validity Analysis

The above-mentioned theories are concerned with how deception occurs but do not give guidelines on how to detect it. The criminal justice discipline helps resolve these issues and has produced research based on the ability to detect deception. In Statement Validity Analysis (SVA) statements made by suspects in criminal investigations are analyzed using Criteria Based Content Analysis (CBCA), Reality Monitoring (RM), and/or Scientific Content Analysis (SCAN) to detect deception [27].Vrij found that use of CBCA and RM to detect deception was successful at a rate of 79.5% and 64.1% respectively [28].

2.6. Criteria Based Content Analysis (CBCA)

CBCA is based on what is known as the Undeutsch-Hypothesis [28, 25, 26] which states that a statement derived from actual memory will differ in content and quality from a statement derived from fantasy. CBCA uses a set of specific criteria to evaluate this hypothesis (see [27] for list of criteria). Trained investigators rate a criminal statement against each criteria using a three-point scale (i.e. 0 - criteria absent, 1 - criteria present, 2 - criteria strongly present).

2.7. Reality Monitoring (RM)

RM was originally designed for studying memory characteristics. It implies that a truthful memory will differ in quality from remembering an event that has been made up. The former is likely to contain perceptual information (visual details, sounds, smells, tastes, and physical sensations), contextual information, and affective information (details about how someone felt during the event), while the latter is likely to contain cognitive operations (such as thoughts and reasoning) [19]. Considering that deceiver's statements are virtually created from imagined events, RM has been applied in the context of deception detection.

2.8. Scientific Content Analysis (SCAN)

Given the transcript or written statement of a subject, research has shown that SCAN is able to discriminate between adult criminal investigation statements of doubtful validity and those that are probably accurate [20]. The presence of some criteria in SCAN [24] suggests deception, such as lack of memory and missing links; the absence of some indicates deception, such as connection, spontaneous corrections, pronouns, first person singular, past tense, denial of allegations, unnecessary links, changes in language; and the interpretation of others depends on where they occur, such as emotion and time. Due to such complexity in assessing a statement, SCAN should be used cautiously when multiple issues may be involved [20].

2.9. Verbal Immediacy

Verbal immediacy (VI) was originally proposed for inferring people's attitude or affect [22]. Since deception is most frequently associated with a communicator's negative affect, non-immediacy comes into play in detecting deception. Non-immediacy is referred to as any indication of separation, non-identity, attenuation of directness, or change in intensity of interaction among the communicator and her referents. The variations in immediacy include verbal forms, such as pronouns, tense, or words. The basic principle of assessing verbal immediacy is via a literal interpretation of the words, rather than on their connotative meanings [29]. For example, while "you and I selected" may be equivalent to "we selected" in meaning, the former is considered more non-immediate than the latter. Non-immediacy is classified into three major categories: spatio-temporal, denotative specificity, and agent-action-object categories, each of which is further broken down into many sub-categories [29]. The detailed criteria for scoring non-immediacy are also provided. The greater the number of non-immediacy scores assigned to a communication unit, the greater the probability that it is part of a communication about a negative experience [22]. Therefore, VI is easy to operate compared with other criteria or theories.

The VI theory has been applied to conversation analysis and coded on a scale with positive score signifying approach and negative score avoidance [1, 14]. The avoidance is indicated by some non-immediacy sub-categories, such as spatial, temporal, passive voice, and modified, and other expressions such as volitional words, politeness, and automatic phrasing.

2.10. Cues

To find cues useful in deception detection, we first extracted some promising cues from existing criteria and cre-

ated some new cues based on our observations of the experimental deceptive messages and knowledge of linguistics. We then merged them into a candidate cue list for testing in a text-based CMC simulation study.

Based on the findings from the reviewed prior research, deceptive messages are expected to have fewer number of words, sentences, self references, affect, temporal, spatial, and perceptual information; more negative statements, uncertainties, passive voice, cognitive information and higher level of language intensity than truthful messages. The number of words and sentences are related to the quantity aspect of language; self references, temporal, spatial, affective, and perceptual information to the specificity of language; and passive voice and uncertainties to the immediacy of language. Encouraged by the research on stylistic analysis as a predictor of competence-trust of newspaper [6], we selected two of the other stylistic indices that have not been covered in the above list as new dimensions of deceptive language: complexity, and diversity. Complexity could be indicated by the ratio of syllables to words or characters to words, and diversity could be defined as the ratio of total number of different words divided by total number of words. In addition, emotiveness [6], which is the ratio of adjectives plus adverbs to nouns plus verbs modifier, is an indication of expressiveness of language in a message. In text-based CMC, typos are both unavoidable and correctable if wanted. Thus, typos in a message may reflect the informality of language in the communication, which might be another useful aspect to view deceptive messages.

In the study, deceptive senders were instructed to mislead their receivers to making a wrong decision. In order to make the efforts successful, they had to provide sufficient evidence to convince the partner that their opinions were correct. Therefore, the quantity and complexity of language in deceptive messages are expected to be higher than that in truthful messages. Deceivers have cognitive anxiety from the possibility of being detected; thereby they may unintentionally adopt a higher degree of non-immediacy and informality in their messages than truth-tellers. To enhance the impression of their opinions, deceivers are likely to display higher expressiveness of language, higher degree of repetition, and lower degree of diversification of wording than truth-tellers. Deceivers may display more affective information to influence the attitude of their partners; however, they may be less likely to produce perceptual information that is originated from real experience, and less likely to include specific information in messages than truth-tellers. Thus, diversity and specificity in deceptive messages are expected to be lower than that in truthful messages. If we label the initiator of a communication as sender and the other party as receiver, and assume that only senders may deceive, These statements can be narrowed down into Hypotheses 1 and Hypotheses 2.

HYPOTHESES 1. Deceptive senders display higher (a) quantity, (b) complexity, (c) non-immediacy, (d) expressiveness, (e) informality, and (f) affect; and less (g) diversity, and (h) specificity of language in their messages than truthful senders.

HYPOTHESES 2. Deceptive senders display higher (a) quantity, (b) complexity, (c) non-immediacy, (d) expressiveness, (e) informality, and (f) affect; and less (g) diversity, and (h) specificity of language in their messages than the receivers.

Based on prior studies, the linguistic features of messages in text-based CMC, and the possibility of automation, we selected 27 linguistic based cues as dependent variables. The cues are grouped into eight linguistic constructs: quantity, complexity, non-immediacy, diversity, affect, specificity, expressiveness, and informality. These linguistic constructs are summarized in Table 1.

Quantity:

1. Word [a]: a written character or combination of characters representing a spoken word.
2. Verb [a]: a word that characteristically is the grammatical center of a predicate and expresses an act, occurrence, or mode of being.
3. Modifier [b]: describes a word or make the meaning of the word more specific. There are two parts of speech that are modifiers- adjectives and adverbs. Adjectives modify nouns and pronouns. Adverbs modify verbs, adjectives, and other adverbs.
4. Noun phrase [a]: a phrase formed by a noun and all its modifiers and determiners.
5. Sentence [a]: a word, clause, or phrase or a group of clauses or phrases forming a syntactic unit which expresses an assertion, a question, a command, a wish, an exclamation, or the performance of an action, which usually begins with a capital letter and concludes with appropriate end punctuation.

Complexity:

6. Average number of clauses : $\frac{\text{total \# of clauses}}{\text{total \# of sentences}}$
7. Average sentence length [6]: $\frac{\text{total \# of words}}{\text{total \# of sentences}}$
8. Average word length: $\frac{\text{total \# of characters}}{\text{total \# of words}}$
9. Average length of noun phrase: $\frac{\text{total \# of words in noun phrases}}{\text{total \# of noun phrases}}$
10. Pausality [6]: $\frac{\text{total \# of punctuation marks}}{\text{total \# of sentences}}$

Non-immediacy [c]:

11. Passive voice: a form of the verb used when the subject is being acted upon rather than doing something.
12. Modal verb [a]: an auxiliary verb that is characteristically used with a verb of predication and expresses a modal modification.

13. Objectification [a]: an expression given to (as an abstract notion, feeling, or ideal) in a form that can be experienced by others. Externalize one's attitude.
14. Uncertainty [a]: a word that indicates lack of sureness about someone or something.
15. Generalizing term: refers to a person (or object) as a class of persons or objects that includes the person (or object) [29].
16. Self reference: singular first personal pronoun.
17. Group reference: first personal plural pronoun.
18. Other reference: third personal pronoun.

Expressiveness:

19. Emotiveness [6]: $\frac{\text{total \# of adjectives + total \# of adverbs}}{\text{total \# of nouns + total \# of verbs}}$

Diversity:

20. Lexical diversity [18]: $\frac{\text{total \# of different words}}{\text{total \# of words}}$, which is the percentage of unique words in all words.
21. Content word diversity: $\frac{\text{total \# of different content words}}{\text{total \# of content words}}$, where content word primarily expresses lexical meaning.
22. Redundancy: $\frac{\text{total \# of function words}}{\text{total \# of sentences}}$, where function word is a word expressing primarily grammatical relationship.

Informality:

23. Typo ratio: $\frac{\text{total \# of misspelled words}}{\text{total \# of words}}$

Specificity [c]:

24. Spatio-temporal information: information about locations or the spatial arrangement of people and/or objects, or information about when the even happened or explicitly describes a sequence of events.
25. Perceptual information: indicates sensorial experiences such as sounds, smells, physical sensations and visual details [27].
26. Positive affect [a]: conscious subjective aspect of a positive emotion apart from bodily changes.
27. Negative affect [a]: conscious subjective aspect of a negative emotion apart from bodily changes.

[a] Source of definition: http://www.webster.com

[b] Source of definition: http://englishplus.com/grammar/glossary.htm

[c] Individual measures in the construct are for per message unit, i.e. frequency counts divided by the total # of words.

Table 1: Summary of Selected Cues and Measures

3 Methodology

3.1. Participants

Sixty undergraduates, 30 dyads (deceptive dyads n = 16, truth dyads n = 14), participated and completed this study in return for homework credit in an advanced Management Information Systems course at a large southwestern university. The sample consisted of female (n = 34) and male (n = 24), undergraduate students that were predominately native English speakers (n = 43).

3.2. Procedures

Data were gathered over four consecutive days using a 2 (deceptive vs. truth conditions) x 2 (dyad role: sender or receiver) x 3 (time periods) experimental design administered entirely via the Internet. Participants were informed that they would be participating with another random person from class to participate in a Desert Survival Problem (see [21] for a complete explanation) decision-making study. The problem focused on being stranded in the desert and the primary goal for participants was to achieve an agreeable ranking of the items they had in order of their usefulness to survival. Additionally, participants were told that the researchers wanted to better understand how people make decisions over e-mail. All participants consented to participate in the four-day study, the tasks for each day could be completed in less than thirty minutes, and upon consenting they completed a series of demographic questions.

On day 1, participants were randomly assigned to one of two conditions, deceptive or truth. In the deceptive condition, one participant (deceptive participant) in each dyad (n = 15) was randomly selected and explicitly instructed to deceive their partner (naïve participant) about how they ranked their items. All participants then viewed a scenario and expert information about surviving in a desert atmosphere. Upon reading the information, participants ranked items according to how useful they thought they would be for survival. The deceptive partner and one truthful partner from each dyad were asked to write their naïve partner an e-mail informing them of how they ranked their items and why. After they were done ranking the items the deceptive participants were administered a deceptive questionnaire asking them to check-off which items they attempted to deceive their partner on. All participants were thanked and instructed to logon the next day either in the morning or afternoon to continue the experiment. The system also sent e-mail reminders at the end of each task.

The naïve partners were asked to logon in the morning on Day 2 to read the e-mail that was sent to them initially by their partners. Upon reading the e-mail, they were given the opportunity to re-rank their items and were then asked to write a response to their partners explaining how they ranked their items and why. Because the experiment was conducted over several days, participants could view any prior information and had the ability to view how they ranked their items throughout the course of the experiment. After writing their e-mail, the naïve participants took a survey about their general interactions with their partner. The deceptive and initial truth partners logged in during the afternoon and were given a random scenario that rendered one of the 12 items useless. Additionally, the deceptive partner could see how their partner ranked the items. Everyone then read and replied to their partner's e-mail. All partici-

pants were asked to take the general interaction survey and the deceptive condition again received an additional survey asking for them to identify deceptive items.

In the morning on day three the naïve participants were given matching scenarios to their partners that rendered an item useless. They were asked to read the replies and also given the opportunity to re-rank their items. Participants then wrote e-mails to their partners. In the afternoon the deceptive and initial truth partners received another random scenario rendering another item useless. They were able to see the rankings of their partner and if those rankings changed plus the deceptive participants were able to see the correct expert's ranking of the items. Both conditions were asked to read and reply to their partner's e-mail and to re-rank their items. Deceptive participants took the deceptive questionnaire.

By day four, all participants in the experiment could logon at any time. The naïve participants again received the same scenarios as their partners and were told to read the e-mail, given the ability to re-rank their items and reply to the e-mail. They were then administered a questionnaire as to how much they believed and trusted their partner. All participants in both conditions were then given the general interaction survey to complete. Upon completion of the survey participants were debriefed about the nature of the study and thanked for their participation.

3.3. Programming and Data Capture

An email messaging system written in Java automatically captured all of the textual data for the desert survival problem. The Resin application server interacted with the messaging system and JSP to send the information to the system's MS SQL 2000 database from which we retrieved the data. The messages were then divided into deceptive and truthful for each of the four days and tagged for part of speech using the commercial natural language processor Conexor iSkim (for more information, see http://www.conexoroy.com). Lastly, simple parsing program tallied the parts of speech of interest and performed the calculations.

4. Analysis/Results

In order to test Hypotheses 1 and 2, a series of simple effect analyses were conducted based on a multivariate analysis. These analyses showed that messages from deceptive senders were significantly different from those from truthful senders on quantity, Wilk's $\lambda = 0.578$; $F(5, 24) = 3.508$, $p = 0.016$, partial $\eta^2 = 42.2\%$, and diversity of language, Wilk's $\lambda = 0.707$; $F(3, 26) = 3.585$, $p = 0.027$, partial $\eta^2 = 29.3\%$. Seven significant univariate differences in the quantity and diversity constructs emerged between deceptive and truthful conditions. Compared with truthful senders, deceptive senders used more words, $F(1,28) = 6.877$, $p = 0.014$, verbs, $F(1,28) = 11.446$, $p = 0.002$, modifiers, $F(1,28) = 8.547$, $p = 0.007$, and noun phrases, $F(1,28) = 4.644$, $p = 0.040$, and displayed less lexical diversity, $F(1, 28) = 9.322$, $p = 0.005$, and content diversity, $F(1, 28) = 8.116$, $p = .008$, than truthful senders. The simple effects of condition on informality, $F(1, 28) = 3.892$, $p = 0.058$, partial $\eta^2 = 12.2\%$, and affect, $F(2, 27) = 2.852$, $p = 0.075$, partial $\eta^2 = 17.4\%$, were only weakly supported. The follow-up univariate analyses on affect showed that deceptive senders produced more positive affect, $F(1, 28) = 5.272$, $p = 0.029$, than truthful senders. As what was predicted, typo ratios in messages from deceptive senders tended to be higher than those from truthful senders.

No significant simple effects of condition was found on complexity, $F(5, 24) = 1.299$, $p = 0.297$, non-immediacy, $F(8, 21) = 1.515$, $p = 0.211$, specificity, $F(2, 27) = 0.599$, $p = 0.577$, and expressiveness, $F(1, 28) = 0.430$, $p = .517$. However, the univariate analyses revealed that pausality in complexity from truthful senders was significantly more than that from deceptive senders, $F(1, 28) = 4.63$, $p = 0.04$, and conversely, deceptive senders produced slightly more group references, $F(1, 28) = 4.151$, $p = 0.051$, and modal verbs in non-immediacy, $F(1, 28) = 3.881$, $p = 0.059$, than truthful senders. These findings are summarized in Table 2

Dependent Construct	Condition	Significant Cues
Quantity	3.508(0.016)*	1,2,3* 4**
Complexity	1.299(0.297)	10*
Non-immediacy	1.515(0.211)	
Expressiveness	0.430(0.517)	
Informality	3.892(0.058)	
Affect	2.852(0.075)	
Diversity	3.585(0.027)*	20,21**
Specificity	0.599(0.577)	

* Significant at .05 level

** Significant at .01 level

Table 2. F-(and p-values) for simple effects and univariate tests of Hypothesis 1 on dependent constructs and cues

As predicted in Hypotheses 2, the simple effects of dyad were supported on quantity, Wilk's $\lambda = 0.323$; $F(5, 11) = 4.604$, $p = 0.016$, partial $\eta^2 = 67.7\%$; affect, Wilk's $\lambda = 0.487$; $F(2, 14) = 7.375$, $p = .006$, partial $\eta^2 = 51.3\%$; non-immediacy, Wilk's $\lambda = 0.209$; $F(8, 8) = 3.795$, $p = 0.038$, partial $\eta^2 = 79.1\%$; and expressiveness, Wilk's $\lambda = 0.719$; $F(1, 15) = 75.876$, $p = 0.028$, partial $\eta^2 = 28.1\%$, between senders and receivers under the deceptive condition. The

follow-up univariate analyses showed that under the deceptive condition senders mentioned more verbs, F(1,15) = 5.919, p = 0.028, modifiers, F(1,15) = 8.934, p = 0.009, and modal verbs, F(1, 15) = 44.041, p = 0.0, and displayed more negative affect, F(1, 15) = 7.355, p = .016, and emotiveness, F(1, 15) = 5.876, p = 0.028, than receivers. They also showed a trend in deceptive condition that messages from senders included fewer self references, F(1, 15) = 3.585, p = 0.078, and more group references, F(1, 15) = 3.675, p = 0.074, words, F(1, 15) = 3.278, p = 0.09, and sentences, F(1, 15) = 3.567, p = 0.078, than those from the receivers. The simple effect of dyad was weakly supported on diversity, Wilk's λ = .611; F(3, 13) = 2.755, p = 0.085, partial η^2 = 38.9%. The follow-up univariate analyses revealed that under deceptive condition senders displayed lower lexical diversity, F(1, 15) = 6.622, p = 0.021, and content diversity, F(1, 15) = 8.459, p = 0.011.

No significant simple effects of dyad for deceptive condition was found on specificity, F(2, 14) = 0.194, p = 0.825, complexity, F(5, 11) = 1.657, p = 0.225, and informality of language, F(1, 15) = 0.001, p = .981. However, the univariate analysis showed a trend that senders were lower on pausality in complexity than receivers, F(1, 15) = 3.778, p = 0.071. Findings for Hypotheses 2 are listed in Table 3

Dependent Construct	Dyad	Significant Cues
Quantity	4.604(0.016)*	2** 3*
Complexity	1.657(0.225)	
Non-immediacy	3.795(0.038)*	12*
Expressiveness	75.876(0.028)*	19*
Informality	0.001(0.981)	
Affect	7.375(0.006)**	
Diversity	2.755(.085)	
Specificity	0.194(0.825)	27*

* Significant at .05 level

** Significant at .01 level

Table 3. F-(and p-values) for simple effects and univariate tests of Hypothesis 1 on dependent constructs and cues

In addition to the 27 linguistic cues studied above, we performed an exploratory, post-hoc analysis on the messages using Word Perfect® 10 Grammatik's grammatical parser. This tool measures a number of grammatical mistakes and syntactic cues, some of which are listed in Table 4.

Higher levels of the grammatical cue Offensive usage contained within Grammatik may demonstrate that the more offensive language used, the less plausible and credible the messages are; thereby creating a more negative tone to the message. Incomplete sentence and Ellipsis in Grammatik are grammatical cues that represents the existence of unfinished, ignored, or incomplete sentences. Wordy refers to the specification of superfluous usages of words. For example, Wordy detects the unnecessary use of "or not" in the phrase "whether or not." Sentence variety specifies the number of different styles of sentences: simple, compound, complex, and compound-complex. Greater sentence variety implies higher frequency in the change of language and less grammatical consistency. Jargon, on the other hand, could be a good indicator of semantic clarity as its usage introduces sophisticated vocabularies that can cause confusion. Passive voice and Run-on sentences are grammatical indicators that also allow for the testability of semantic ambiguity as they make the grammatical structure unclear. Second person address (you) and Possessive form help to depict the speakers' attitudes and involvement based on their comments. The results of the post-hoc analysis of these cues appear in Table 5.

Cues	Description
Offensive usage	Impolite language, words usages in the message
Incomplete sentences	Incomplete sentence structure such as missing subjects, objects... etc.
Ellipsis	Ellipsis usage
Wordy	Wordy relative clauses or vague modifiers (such as "fairly" or "pretty"), redundant adverbs, too many negatives, the unnecessary use of words... etc.
Sentence variety	The number of different sentence style, such as simple, compound, complex... etc.
Jargon usage	A hybrid language or dialect simplified in vocabulary and grammar and used for communication between peoples of different speech, or technical terminology idiom
Run-on sentences	The usage of sentences with at least two parts, either one of which can stand by itself (in other words, two independent clauses).
Second person	The usage of "you"
Possessive form	The usage of possessive form

Table 4. Description of possible cues analyzed using WordPerfect® 10 Grammatik

The results showed some cues were significant: ellipsis (p=0.056), wordy (p=0.011), passive voice (p>0.000), second person address (p=0.002) and possessive form (p=0.030), and the rest were insignificant: offensive usage (p=0.332), incomplete sentences (p=0.352), sentence variety (p=0.619), jargon usage (p=0.579), and run-on sentences (p=0.422).

	Truthful		Deceptive		
Cues	Mean	Stdev.	Mean	Stdev.	p
Offensive usage	0.18	39	0.06	0.39	0.33
Incomplete sentences	3.88	4.03	2.29	4.55	0.35
Ellipsis	1.59	2.40	0.29	0.59	0.06
Wordy	0.35	0.79	2.35	2.62	0.01
Sentence variety	1.41	1.12	1.71	1.83	0.62
Jargon usage	0.12	0.33	0.06	0.24	0.58
Run-on sentences	0.29	0.77	0.12	0.33	0.42
Second person	1.06	1.34	5.41	4.60	.001
Possessive form	0.41	0.71	0	0	0.03

Table 5. Results of analysis using WordPerfect® 10 Grammatik

5. Discussion

In general, the test results support Hypotheses 1a and 1g that messages from deceptive senders were significantly higher on quantity and lower on diversity than those from truthful senders. Hypotheses 1e and 1f were weakly supported, which showed a trend that the informality and affect of language from deceptive senders were higher than that from truthful senders. Hypotheses 1b and 1c, which stated that senders from the two conditions differed on complexity and non-immediacy of language, were only partially supported in the univariate analyses. It found that deceptive senders displayed less pausality, and more group references and modal verbs than truthful senders. However, the directions on pausality and group reference were opposite to the prediction. Hypotheses 1d and 1h were not supported, for there was no significant difference on expressiveness and specificity between senders under the two conditions. Overall, Hypotheses 2 was well supported. In particular, the results supported Hypotheses 2a, 2c, 2d, and 2f that senders' language was higher on quantity, non-immediacy, expressiveness and affect than receivers' under the deceptive condition. Hypotheses 2g which stated that deceptive senders displayed lower diversity of language than their receivers was weakly supported. Hypothesis 2b was only partially weakly supported in the univariate analyses in that senders tended to show more pausality than receivers under the deceptive condition. However, Hypothesis 2f and 2h were not supported; there was no significant difference on informality and specificity between senders and receivers under the deceptive condition.

While deceptive subjects in this study displayed less lexical diversity and self references, and more modal verbs, as shown in some of the other studies, the deceptive condition under text-based CMC led to more rather than fewer words, group references, and affective information, and less rather than more pausality and generalizing term in messages. Contrary to the proposition in RM, we found a trend that deceivers showed more affective information than truth-tellers. The lack of evidence in support of spatio-temporal information was against the prediction in CBCA, RM, and VI. The ratio of generalizing terms produced in deceptive condition was surprisingly higher than that in the truthful condition. However, The context and task in this study were unusual compared with the previous studies; therefore, the results cannot be accepted as evidence against other criteria and constructs. The opposite findings in deceptive messages in these cases were likely to illustrate the unique nature of text-based CMC.

Additionally, most previous deception research suggests that deceptive messages should be shorter than truthful messages because deceivers do not have the details to put into a message that a truth teller does. Most of the literature, however, focuses on statements of fact or recollections such as criminal statement analysis. An informal review of some of the messages produced by the Desert Survival study showed that it is possible the subjects in this study were trying to boost their credibility to make their rankings believable. Furthermore, it seemed the deceivers were giving more elaborate reasons for their rankings while the truth teller, not needing to prove anything, ranked the items with short, common sense reasons or no reasons at all.

This discrepancy between previous research and our findings suggests that deceptive cues may tend to differ based on the intent of the deceiver. For example, a deceiver who is covering up an occurrence in the past will likely give less detail in a statement about the occurrence than someone who is simply stating what happened. The lack of detail exists because the deceiver simply does not have the detail of reality to give when fabricating the message. Because of the lack of detail, the length of the statement should be shorter than a truth teller. On the other hand, if the deceiver's intent is to establish credibility to enhance the deception's success, then he or she will likely attempt to support that credibility with persuasive text. In this case, the deceiver's message length is likely to be greater than the truth teller's message length.

The considerable difference between the deceptive and truthful conditions in the wordiness of sentences supports the idea that deceivers are likely to put in superfluous words or meaningless expressions in case they do not have much to say but still want to have the impression of "completeness." Particularly in an asynchronous experiment, where subjects are given a length of time to complete each e-mail message, deceivers have enough time to manipulate their messages by adding more but repetitive or meaningless (wordy) expressions to make their e-mail look informative. As a result, deceivers used less ellipsis in sentences than truth tellers. The finding is consistent with our other findings that deceivers in

the Desert Survival Problem provide longer messages. The fact that deceivers used more "you" pronoun and less possessive form implied they tend toward communicating with less immediacy.

5.1. Future Areas

Cross-validation studies are crucial with this type of applied research. We plan to test the effective cues found in this study in messages created with other types of tasks under text-based CMC. As the first step, we can use the remaining hundreds of messages from uncompleted subjects in this experiment to validate the linguistic based cues. Moreover, linguistic cues as a function of deception strategy or speech act are worthy to be explored. The next step is to create computer model for deception detection and derive weights for various cues.

Attempting deception detection in text-based CMC does have the advantage that the text content of the communication is readily processable by current computer technology. In contrast, processing either audio or video requires unproven speech recognition software, which simply produces text (transcription) already obtainable with text-based CMC. Furthermore, those attempting to detect deception in text-based CMC may do so while spatially separated from a potential deceiver, which allows the detector to employ automated tools that might otherwise arouse suspicion in a face-to-face environment.

To take advantage of the maturity of text processing technology, we propose building a computer-based automated tool that will objectively measure text-based deception cues and give the users of the tool a probability-based score to aid their judgment of the deceptive intent of a message. It is unlikely that the automated tool would be very successful in detecting deception on its own; rather it would be a useful tool for supporting trained human detectors in analyzing messages in text-based CMC. The beginning steps in building such an automated tool are first, finding those cues that are transmitted in text-based CMC, and second, identifying those cues that are both possible and simple to program as part of a software tool. The method we used in this study focuses on the first step.

6. Conclusion

Deception may always be a part of every day life, but individuals and organizations need not be powerless to detect its use against them. Researchers in the communication and psychology fields have studied deception for years and have made some progress toward useable deception detection. To achieve the end of building such an automated tool, we are taking a stepwise approach. This study reflects our partial efforts in the first step, that is to find effective cues to deception that are transmitted in text-based CMC. As more and more promising cues are accumulated, we will be able to develop a detection tool based on the cues in the next step.

References

[1] C. P. Borchgrevink. Verbal immediacy, verbal immediacy coding scheme (unpublished).

[2] D. B. Buller and J. K. Burgoon. Interpersonal deception theory. *Communication Theory*, 6(3):203–242, 1996.

[3] D. B. Buller, J. K. Burgoon, A. L. S. Buslig, and J. F. Roiger. Interpersonal deception: Viii. further analysis of nonverbal and verbal correlates of equivocation from the bavelas et al (1990) research. *Journal of Language & Social Psychology Special Issue: Interpersonal deception*, 13(4):396–417, 1994.

[4] J. K. Burgoon and D. B. Buller. Interpersonal deception: Iii. effects of deceit on perceived communication and nonverbal behavior dynamics. *Journal of Nonverbal Behavior*, 18(2):155–184, 1994.

[5] J. K. Burgoon, D. B. Buller, C. H. White, W. Afifi, and A. L. S. Buslig. The role of conversational involvement in deceptive interpersonal interactions. *Personality & Social Psychology Bulletin*, 25(6):669–685, 1999.

[6] J. K. Burgoon, M. Burgoon, and M. Wilkinson. Writing style as predictor of newspaper readership, satisfaction and image. *Journalism Quarterly*, 58(2):225–231, 1981.

[7] J. R. Carlson, J. F. George, J. K. Burgoon, M. Adkins, and C. H. White. Deception in computer-mediated communiction. *Academy of Management Journal (Under Review)*, 2001.

[8] J. R. Carlson and R. W. Zmud. Channel expansion theory and the experiential nature of media richness perceptions. *Academy of Management Journal*, 42(2):153–170, 1999.

[9] R. Daft and R. Lengel. Organizational information, media richness and structural design. *Management Science*, 32:554–571, 1986.

[10] R. L. Daft, R. H. Lengel, and L. K. Trevino. Message equivocality, media selection, and manager performance: Implications for information systems. *MIS Quarterly*, 11(3):355–366, 1987.

[11] B. M. DePaulo, D. A. Kashy, S. E. Kirkendol, M. M. Wyer, and J. A. Epstein. Lying in everyday life. *Journal of personality and social psychology*, (70):979–995, 1996.

[12] B. M. DePaulo, J. J. Lindsay, B. E. Malone, L. Muhlenbruck, K. Charlton, and H. Cooper. Cues to deception. *Psychology Bulletin (under review)*, 2000.

[13] B. M. DePaulo, J. I. Stone, and G. D. Lassiter. Deceiving and detecting deceit. In B. R. Schlenker, editor, *The Self and Social Life*, pages 323–370. McGraw-Hill, New York, 1985.

[14] W. A. Donohue. Verbal intensity. In *Communication, Marital Dispute, and Divorce Mediation*. Lawrence Erlbaum Associates, Hillsdale, NJ, 1991.

[15] J. F. George and J. R. Carlson. Electronic lies: Lying to others and detecting lies using electronic media. In *Proceedings of the Fifth Americas Conference on Information Systems*, pages 612–614, Milwaukee, WI, 1999.

[16] J. F. George and J. R. Carlson. Group support systems and deceptive communication. In *Hawaii International Conference on System Sciences*, Maui, Hawaii, 1999.

[17] H. P. Grice. Logic and conversation. In J. L. Cole, P.andMorgan, editor, *Synax and Semantics Speech Acts*, volume 3, pages 41–58. 1975.

[18] H. Hollien. *The Acoustics of Crime: The New Science of Forensic Phonetics*. Plenum, New York, 1990.

[19] G. Köhnken, E. Schimossek, E. Aschermann, and E. Höfer. The cognitive interview and the assessment of the credibility of adults' statements. *Journal of Applied Psychology*, 80:671–684, 1995.

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

[21] J. Lafferty and P. Eady. *The Desert Survival Problem*. Experimental Learning Methods, Plymouth, Michigan, 1974.

[22] A. Mehrabian and M. Wiener. Non-immediacy between communicator and object of communication in a verbal message: Application to the inference of attitudes. *Journal of Consulting Psychology*, 30:420–425, 1966.

[23] G. R. Miller and J. B. Stiff. *Deceptive Communication*. Sage Publications, Thousand Oaks, CA, 1993.

[24] S. Sporer. The less travelled road to truth: Verbal cues in deception detection in accounts of fabricated and self-experienced events. *Applied Cognitive Psychology*, 11:373–397, 1997.

[25] M. Steller and G. Köhnken. Criteria-based content analysis. In D. C. Raskin, editor, *Psychological methods in criminal investigation and evidence*, pages 26–181. Springer-Verlag, New York, 1989.

[26] U. Undeutsch. The development of statement reality analysis. In U. Undeutsch, editor, *Credibility Assessment*, pages 101–121. Kluwer, Dordrecht, The Netherlands, 1989.

[27] A. Vrij. *Detecing Lies and Deceit: the psychology of lying and implications for professional practice*. John Wiley & Sons, Chichester, 2000.

[28] A. Vrij, K. Edward, K. P. Roberts, and R. Bull. Detecting deceit via analysis of verbal and nonverbal behavior. *Journal of Nonverbal Behavior*, 24(4):239–263, 2000.

[29] M. Wiener and A. Mehrabian. *Language Within Language*. Appleton-Cetury-Crofts, New York, 1968.

[30] M. Zuckerman, B. M. DePaulo, and R. Rosenthal. Verbal and nonverbal communication of deception. In L. Berkowitz, editor, *Advances in Experimental Social Psychology*, pages 1–59. Academic, New York, 1981.