Ethical Dilemma: Deception Dynamics in Computer-Mediated Group Communication

Shuyuan Mary Ho
School of Information, Florida State University, 142 Collegiate Loop, Tallahassee, FL 32306-2100. E-mail: smho@fsu.edu

Jeffrey T. Hancock
Department of Communication, Stanford University, Building 120, 450 Serra Mall, Stanford, CA 94305-2050. E-mail: hancockj@stanford.edu

Cheryl Booth
School of Information, Florida State University, 142 Collegiate Loop, Tallahassee, FL 32306-2100. E-mail: clb14h@my.fsu.edu

Words symbolically represent communicative and behavioral intent, and can provide clues to a communicator's future actions in online communication. This paper describes a sociotechnical study conducted from 2008 through 2015 to identify deceptive communicative intent within group context as manifested in language-action cues. Specifically, this study used an online team-based game that simulates real-world deceptive insider scenarios to examine several dimensions of group communication. First, we studied how language-action cues differ between groups with and groups without a compromised actor. We also examine how these cues differ within groups in terms of the group members' individual and collective interactions with the compromised actor. Finally, we look at how the cues of compromised actors differ from those of noncompromised actors, and how communication behavior changes after an actor is presented with an ethical dilemma. The results of the study further our understanding of language-action cues as indicators for unmasking a potential deceptive insider.

Introduction

"Actions speak louder than words." This well-known phrase implies that actions and words can be seen as distinct, measureable, and quantifiable, and highlights our general tendency to view words as being separate from actions. However, Austin (1962) suggested a different view of words and actions. Words—whether written or spoken—can be interpreted as symbolic action that reflects behavioral intent. Studies have demonstrated that social actors' intent can be attributed through interactions as well as by words exchanged, in expressions of information behavior (Ho & Benbasat, 2014; Ho, Fu, et al., 2015; Ho, Hancock, Booth, Burmester, et al., 2016; Ho, Hancock, Booth, Liu, et al., 2016; Ho, Hancock, et al., 2015).

The ability to detect deception using linguistic cues is particularly important in the discipline of information studies (Ho & Hollister, 2013; Rubin, 2010; Rubin & Lukoianova, 2015; Whitty, Buchanan, Joinson, & Meredith, 2012). The growing prevalence of computer-mediated communication (CMC) has significantly changed the way information is communicated. Many of the CMC tools used today (e.g., e-mail, instant-messaging/chat, blogs) are text-based and thus lack the physical communication cues (e.g., body language, facial expression) that can help discern deception. This has made it a relatively easy for malevolent individuals to successfully misrepresent not only the content of their messages, but their actual identity. Several forms of deceptive online communication have become especially prevalent, including phishing (particularly "spear phishing") attacks (Fuller, Marett, & Twitchell, 2012; Wright & Marett, 2010), "griefing" (i.e., a deception strategy to trick other players) in online games (Rubin & Camm, 2013), and CMC-based social engineering (Mitnick, Simon, & Wozniak, 2002). Computer-mediated deception has also become prevalent in online employment recruitment (Allen, Mahto, & Otondo, 2007), virtual professional communities (Joinson & Dietz-Uhler, 2002), and social media relationships

Received April 28, 2016; revised September 8, 2016; accepted December 15, 2016

 DOI: 10.1002/asi.23849

(Hancock et al., 2009; Toma & Hancock, 2010, 2012). As geographically dispersed individuals, teams, and organizations increasingly rely on CMC for collaboration, threats from deceptive communications are also on the increase.

Unfortunately, detecting and interpreting behavioral intent through language is especially challenging—especially in the case of text-based CMC—where language cues are the only means by which communicative intent can be ascertained (Hancock, Woodworth, & Goorha, 2010). If identifying deception in interpersonal communication is difficult; it is even more so in the context of group communication. Yet, being able to identify deception and determine whether communication can be trusted is critical to successful group collaboration (Jones & Marsh, 1997).

While an actor's communicative intent can be observed in and ascertained through language as symbolic action, it is challenging to benchmark consistency in words as compared to actions, especially considering this can fluctuate based on changes in situation or circumstance. In this study, we seek to understand the challenges in determining an actor's deceptive intentions through the words exchanged within a collaborative online group. Our overarching research question is thus framed as: *Can we detect shifts in a deceptive actor's intent when facing an ethical dilemma in a collaborative group context through language-action cues?* Here, "language-action cues" refers to linguistic styles, or patterns in an actor's written expression and manifested as a subtle signal to other actors (Ho, Hancock, Booth, Burmester, et al., 2016; Ho, Liu, Booth, & Hariharan, 2016).

To answer this question we developed an interactive online game in which an ethical dilemma is presented to focal actors in one or more randomly assigned virtual groups of collaborative problem-solvers, thereby introducing a motivated deceptive actor as a factor in the groups' interactions (Ho, Fu et al., 2015; Ho, Hancock, Booth, Burmester, et al., 2016; Ho & Warkentin, 2017). The study was conducted across different technology platforms, and data were collected both in 2008 and in 2014. The conversations of each group were archived, processed, and analyzed to study the language-action cues of actors facing an ethical dilemma, and how deceptive intent can be uncovered within group-based social interaction. This paper first discusses the viability of using language in detecting deceptive intent expressed as an ethical dilemma in CMC. Then, research hypotheses are posited regarding language-action cues during conflict of interest scenarios. The research design and data collection employed in the study are discussed next, followed by the results. Finally, the paper discusses implications and limitations of the study, concluding with some remarks on future work.

Language and Ethical Dilemma in Computer-Mediated Deception

Whether in computer-mediated or face-to-face (F2F) environments, there are several fundamental aspects of deception that should be understood. First, there is a distinction between being deceptive and simply being mistaken. As Buller and Burgoon (1996) stated, deception involves the intentional distortion of the informational content of a message in order to convey something other than the truth. Because of the volitional and intentional nature of this action, "deception" as used herein excludes accidentally incorrect or incomplete information (i.e., mistakes or misstatements of fact). Second, deception can be categorized as being either spontaneous (on-the-fly) or planned (Whitty et al., 2012). Third, as Whitty et al. (2012) pointed out, a deception may be for the benefit of the deceptive actor (i.e., self-serving) or the recipient and/or other third parties (i.e., for altruistic reasons). Finally, deception can be insignificant (e.g., "everyday" lies, or little white lies) (DePaulo & Kashy, 1998; DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996), or serious (i.e., having significant consequences) (DePaulo, Ansfield, Kirkendol, & Boden, 2004), or may fall somewhere in between.

Deceptive Intent in Interpersonal CMC

Cues available to detect deception differ across environments. Cues available in one environment may be lacking in another. For example, the deceptive communication cues that exist in F2F communication (e.g., body language, facial expression, hand gestures, etc.) are not present in a text-based CMC environment (Hancock et al., 2009; Hirschberg, 1993, 2002; Hirschberg & Litman, 1993). Pennebaker and King (1999) suggested that linguistic styles and use of certain linguistic cues (i.e., self-references, negations, inclusivity, insight, causality) can provide insight into an actor's psychological state, revealing deceptive intent. Pennebaker, Mehl, and Niederhoffer (2003) further suggested that words reveal the inner characteristics of an individual (i.e., words associated with cognitive and affective processes), and convey cognitive thinking and emotions.

The proposition that language-action cues can be used to detect deception has been supported in a number of studies. For example, in studies examining deception in online dating profiles, users who were highly deceptive in their profiles included fewer self-references, more negative words and negations, more motion words, but fewer words overall when compared to less deceptive profiles (Toma & Hancock, 2010, 2012). Zhou, Burgoon, Twitchell, Qin, and Nunamaker (2004) found that deceptive actors tend to use more modal verbs and fewer self-references. Likewise, in synchronous text-based CMC environments, Hancock, Curry, Goorha, and Woodworth (2008) found an observable difference in these cues between a nonmotivated deceptive actor and an actor who has been motivated to deceive, and that motivated deceptive actors used more words overall, more sense-based words, more other-oriented pronouns, and fewer self-oriented pronouns. In particular, motivated deceptive actors used more negations than truthful actors, with no difference in the frequency of causal words between truthful and false statements.

DOI: 10.1002/asi

In addition to specific words chosen by a deceptive actor, the overall descriptiveness of his/her communication and even the overall number of words used may also reveal deceptive intent. For example, studies suggest that a deceptive actor may try to gain credibility to enhance the success of deception by being increasingly wordy with peripheral expressions (Hancock et al., 2008; Zhou, Twitchell, Qin, Burgoon, & Nunamaker, 2003). Indeed, although excessive descriptiveness may undermine credibility by revealing inconsistencies, sufficient description in a deceptive statement is necessary to provide authenticity (Pennebaker Conglomerates, 2007). Zhou et al. (2004) also found that deceptive actors use proportionately more imagery words (i.e., sensory, spatial, and temporal expressions) than truth-tellers. Because deceptive actors are unable to rely on experience or memory to deceive, they tend to use more sensory expressions (e.g., sounds, smells, physical sensations, and visual details), spatial (e.g., locations of people or objects), and temporal (e.g., time when the event happened) words. In contrast to F2F interactions, where deceptive actors have been found to be more laconic (DePaulo et al., 2003), Zhou et al. (2004) found that deceptive actors tend to be more active, wordy, taking shorter pauses between messages than truth tellers. In a two-person online game that creates spontaneous interpersonal deception scenarios, Ho, Hancock, Booth, and Liu (2016) found that deceivers tend to use words associated with affective and cognitive processes, while tending to take longer to respond than truth-tellers.

Ethical Dilemma Expressed in Group Communication

All the above studies were designed and conducted in the context of interpersonal communication. Although these studies do not specifically address deception in group dynamics, they nonetheless inform our research assumptions. As the base of all human relationships, trust influences relationships at the interpersonal level (Rotter, 1967, 1980), but also within, between, and among groups (Hosmer, 1995). The context of trust further extends to sociopolitical relations, both nationally (citizens' trust in government) (Hoffman, 2002), and internationally (one government's trust of another) (Welch, Hinnant, & Moon, 2005). Hosmer (1995) defined trust as follows:

> *Trust is the expectation by one person, group or firm of ethically justifiable behavior—that is, morally correct decisions and actions based upon ethical principles of analysis—on the part of the other person, group or firm in a joint endeavor or economic exchange*. (p. 399)

Hosmer's definition emphasizes that *trust* incorporates ethics and ethical behavior. In a collaborative group environment, ethics can be observed and identified through actors' communicative linguistic cues and behavior. For example, Greitzer, Kangas, Noonan, Brown, and Ferryman (2013) attempted to create a behavioral/psychological model to identify deceptive insiders. The focus involves categorizing and modeling behaviors and psycholinguistic cues (Brown, Greitzer, & Watkins, 2013; Brown, Watkins, & Greitzer, 2013; Taylor et al., 2013). Brown, Watkins et al. (2013) focused on linguistic cues, "translating" them into behavioral categories corresponding to behaviors associated with deceptive insiders. Taylor et al. (2013) conducted an experiment where 25% of the participants were incentivized to "act" as deceptive insiders. It was discovered that insiders who passed information—without authorization—to a provocateur seemed to be more self-focused, and used more linguistic features associated with negative emotion and cognitive processes than nondeceptive actors. Ho and Warkentin (2017) set up experiments to study how a group member, when facing an ethical dilemma, would react to group members, to simulate insider threat scenarios in virtual groups.

Ho, Hancock, Booth, Burmester, et al. (2016) further studied the differences in language-action cues between deceptive actors and nondeceptive actors in group interaction, and how the language-action cues reflected in the groups with a deceptive actor differed from those without a deceptive actor. The study was conducted using an online game, which simulated real-world interaction scenarios, and was designed such that the deceptive actor's actions were allowed to naturally evolve. From this, Ho, Hancock, Booth, Burmester, et al. (2016) concluded not only that differences in language-action cues can be identified between deceptive actors and nondeceptive actors as well as between groups with a deceptive actor and groups without one, but also suggested that differences can be observed between the language-action cues of actively deceptive actors and those of passively deceptive actors. That is, based on McCornack's (1992) classification scheme of information manipulation; in some groups, the deceptive insider would actively sabotage her/his team, while in other groups, the deceptive actor would undermine the group passively, such as by simply not actively participating. The results of Ho, Hancock, Booth, Burmester, et al. (2016) suggested that the cues to a deceptive insider's behavior tend to be subtle, or even unnoticeable. The results also support the proposition that in these scenarios, a statistically significant difference exists in the use of certain language-action cues.

Research Model

We assume that when an actor faces an ethical dilemma, his/her language can provide indications of deceptive intent within group communication—both intra- and extragroup communications. Using group level analysis, we explore differences in those cues when used by a deceptive focal actor versus a nondeceptive actor (H1–H2). Additionally, using individual level analysis, we compare the language-action cues of focal actors who had been motivated to deceive with those of nondeceptive focal actors, as well as those cues from independent communications with a designated external authority having an interest in the group's outcomes

DOI: 10.1002/asi

(H3–H4). For the purposes of our discussion, we will refer to this authority as an *overseer*.

Deceptive Actors Versus Others

Zuckerman, DePaulo, and Rosenthal (1981) suggested that a deceptive actor uses more cognitive efforts in constructing deceptive messages. Hancock et al. (2008) further informed our understanding of how deceptive intent might vary during interpersonal synchronous CMC. Motivated deceivers tend to use more words of negation (e.g., "no," "not," "never") to avoid being caught in a contradiction. They tend to shape their conversation by how their partners behave (Hancock et al., 2008). Toma and Hancock (2010) also reported that deceivers used more words of negation in online dating profiles. In a group, a deceptive actor may alter the style of communication to sound more persuasive, using more negation words so as to sound more in control. Thus, we hypothesize that:

> H1: *Deceptive actors will use more negation words than other (nondeceptive) actors during synchronous CMC group interaction.*

Human Sensors in Group Dynamics

Ho, Hancock, Booth, Burmester, et al. (2016) further suggested that, compared to groups without a deceptive actor, groups with a deceptive actor may tend to use i) more words of inclusion (e.g., "and," "with," and "include") to support a sense of belonging; ii) more words of exclusivity ("but," "without," "exclude") to create a sense of discrete uniqueness; iii) more words of certainty ("always," "never") to bolster group confidence; and iv) more words of achievement (e.g., "win," "earn," and "hero") to create a sense of common purpose. Thus, broadly speaking, we may assume an increased use of words around cognitive processes in groups with a deceptive actor, and further assume that groups with a deceptive actor will use more cognitive process-oriented words than will groups without a deceptive actor. We thus hypothesize that:

> H2(a): *Language-action cues associated with cognitive processes will be more prevalent in groups with a deceptive actor than in groups without a deceptive actor.*

While Hancock et al. (2008) suggested that motivated liars tend to avoid using causal terms but more sense-based words (e.g., touching, feeling, etc.), there is little research about how groups react in situations where an actor intends to deceive. Ho, Hancock, Booth, Burmester, et al. (2016) suggested the effect of the introduction of a deceptive actor into a group, and led us to further speculate that group interaction will be altered as a result. In such a scenario, the group's language-action cues may demonstrate a significant reduction in group members' willingness to cooperate with the deceptive actor. In particular, one may observe an increased usage of language-action cues relating to affective processes, as well as in words connoting negation. To test this assumption, we hypothesize as follows:

> H2(b): *The use of language-action cues associated with negation and affective process in a group's interactions will increase after an incentive for deception has been accepted by an actor within the group.*

Nonobservable Hidden Agenda

One hallmark of deceptive behavior is the attempted concealment (whether or not successful) to avoid detection. We would expect the deceptive actor to attempt to sound and/or act much the same as if s/he were telling the truth. Since this makes it difficult (or impossible) to identify or detect deceptive cues—including language-action cues—it seems unlikely that there would be any observable differences between the language-action cues used by deceptive actors and those used by nondeceptive actors. But this analysis does not account for *leakage*, which Ekman and Friesen (1969) described as unconscious/subconscious behaviors and cues on the part of the deceptive actor that undermine deception. The increased cognitive load associated with the ethical dilemma of deceiving others in a way that is intentionally harmful may make it difficult for the deceptive actor to maintain the deception. For this reason, there may be observable differences between the language-action cues of deceptive actors and nondeceptive actors. We assume that a deceptive actor will communicate and interact with his/her group differently when compared to a nondeceptive actor. Accordingly, we frame our hypothesis in the affirmative:

> H3(a): *The language-action cues of a deceptive actor's communications with her/his group will differ from those of a nondeceptive actor.*

This same fundamental goal of concealment also makes it seem unlikely that there would be any observable changes in language-action cues used by an actor before an incentive for deception has been introduced and after the deceptive actor has accepted it. That is, we would expect that the deceptive actor would attempt to conceal deceptive intent by attempting to behave the same before and after this event (cf. Ekman and Friesen, 1969). However, even if not probable, it is nonetheless possible that differences may be observed. In other words, a deceptive actor may or may not change the way in which s/he communicates with her/his group after being offered an incentive to deceive. Thus, we again frame our hypothesis in the affirmative:

> H3(b): *The language-action cues used by a deceptive actor in communicating with his/her group will be different before an incentive for deception has been introduced, when compared to after an incentive has been accepted.*

DOI: 10.1002/asi

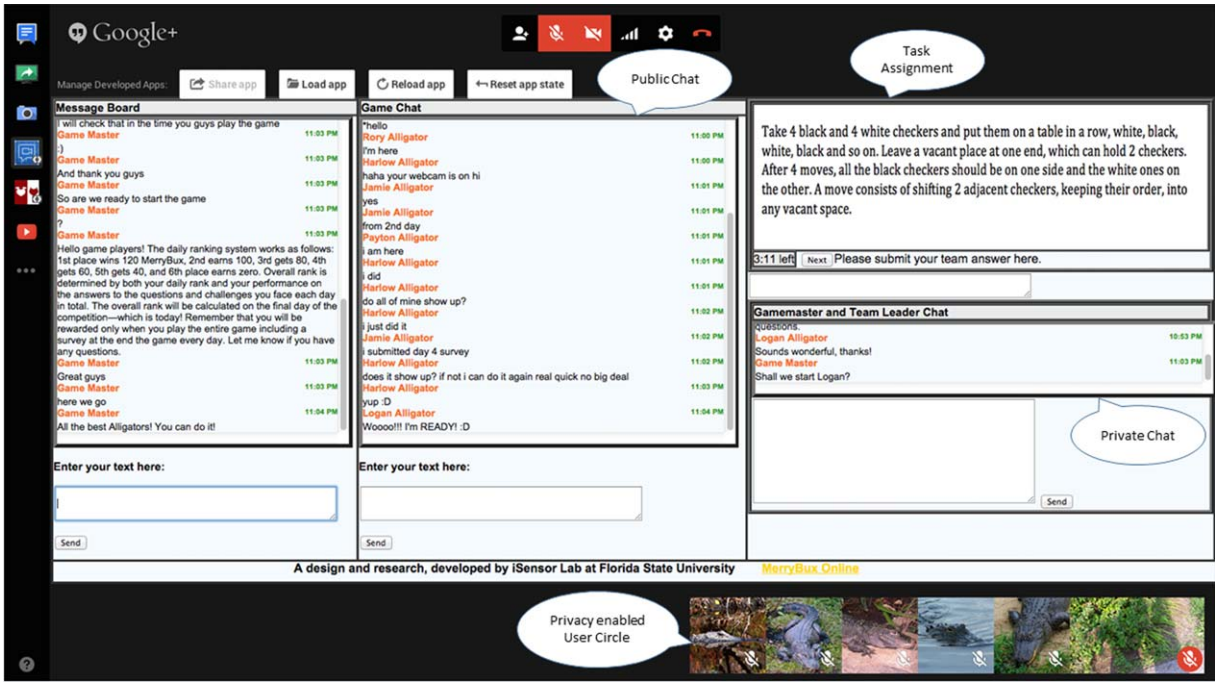

FIG. 1. Illustration of an online game designed and developed for data collection. [Color figure can be viewed at wileyonlinelibrary.com]

Motivated Deception and Ethical Dilemmas

As discussed, deception is generally motivated either by the deceptive actor's desire to benefit him/herself, or by her/his desire to benefit others (Whitty et al., 2012). These benefits can be tangible or intangible. For example, a deceptive actor may withhold information—such as a painful truth—from a loved one. This is motivated by a desire for that person not to be hurt (i.e., intangible benefit to others). In this case, we may also consider that there is an intangible benefit to the deceptive actor in that s/he may avoid a painful or uncomfortable conversation. This example illustrates the differences between benefit to "self" and benefit to "others," which are not mutually exclusive. On the other hand, a deceptive actor may withhold information about a shared winning lottery ticket from other owner/s of the ticket (i.e., a tangible benefit to the self). This example illustrates how benefit to self can result in sacrificing others' interests. Indeed, not only does the deception in this case not benefit the "others," it can be seen as harmful. This example also illustrates the type of ethical dilemma a group member may face in choosing whether or not to deceive.

Just as deceptive intent *per se* may be revealed by the use of certain language-action cues, the group member's decision-making process in resolving this ethical dilemma (i.e., whether to deceive or not) can also be observed using certain language-action cues. Other research investigating this question, such as Hancock et al. (2008), have focused only on deceptive actors (rather than comparing deceptive and nondeceptive actors) and the differences in language-action cues between a motivated and an unmotivated deceptive actor. In contrast, our emphasis is on examining differences in language-action cues between a nondeceptive and a deceptive actor within an interacting group. Moreover, we examine this question in a dynamic environment—comparing language-action cues used by a focal actor both before and after s/he has been offered an incentive to deceive. Specifically, we look at the private "shadow" communications between the selected focal actors in each group and the *overseer*. We thus hypothesize that:

H4(a): *A motivated deceptive actor will use more language-action cues associated with negation and cognitive process in the course of communicating with an overseer than a nondeceptive actor will.*

Furthermore, when an ethical dilemma is introduced to a deceptive actor, it may be possible to observe differences in language-action cues used by him/her when "baited" to deceive by the *overseer*. In other words, the deceptive actor is conscious of the bait, and may, accordingly, express and demonstrate more inclusivity (cognitive load) to the bait-giver as a "partner in crime." Thus, we hypothesize that:

H4(b): *A deceptive actor will use more language-action cues associated with affective process and cognitive process in communicating with an overseer after having accepted an incentive for deception than before the incentive is introduced.*

Research Design

To investigate our overarching research question and the hypotheses set forth above, we developed an online multiplayer/team gaming environment to simulate group-interaction scenarios (Figure 1). The game uses a cue-lean, synchronous

DOI: 10.1002/asi

TABLE 1. Control/treatment group designations for team assignments.

Experiment design	Control/no bait					Treatment/bait				
# of day(s)	1	2	3	4	5	1	2	3	4	5
Negative feedback			Team A					Team C		
Positive feedback			Team B					Team D		
Neutral feedback			Team E					Team F		

text-based CMC environment to capture messages among a team in the course of each game session, as well as between the designated focal actor in each team and an external overseer. Using these chat messages, we can compare and contrast group dynamics, behaviors of teams with and without a deceptive actor, as well as before and after an incentive to deceive ("bait") is accepted by a team member. In addition, we can compare and contrast the behaviors of a deceptive actor before and after an incentive to deceive has been accepted, as well as the behaviors of an incentivized deceptive actor and a nondeceptive actor.

Experiment Design and Overview

Virtual teams were established of approximately three or four members, plus one randomly designated focal actor (called the "Team Leader"). These teams were randomly designated as being either a control group (no "bait" offered) or a treatment group (Team Leader is offered "bait") (Table 1). Players were randomly assigned to teams, and real names were replaced with pseudonyms to protect privacy. Each team's stated objective was to work collaboratively to complete a series of assigned tasks within a specified timeframe. Specifically, each team was presented with a set of seven logic puzzles as task assignments—a new/different set for each game session—each puzzle must be solved within 6 minutes. Although the same set of puzzles was used for all teams on the same day, the specific order of the puzzle-questions presented to any given team was randomized, and the difficulty of the puzzles was also controlled (i.e., the level of difficulty was the same for every team on a particular day in a 5-day competition).

Play consisted of daily game sessions, each of which lasted ~40 minutes, over 5 consecutive days. The game was competitive between the teams, so each team was effectively competing against the others to correctly complete the most puzzles within the allotted time.

In addition to the players/participants, an external party (i.e., a research experimenter) is appointed to function as the *overseer* (called the "Game Master"). Although the Game Master has visibility to all team chats/communications, s/he interacts only with the Team Leaders, who communicate using a private chat function not available to team members. Each Team Leader is responsible for reporting to the Game Master on his/her team's progress. Specifically, Team Leaders submit answers on behalf of the team.

Team Performance-Based Award System Creates Opportunity for Ethical Dilemma

Team performance in this competitive game is rewarded through use of a micropayment online banking system called MerryBux. That is, the team achieving the highest overall performance score (i.e., the team who correctly solved the most puzzles in the required time) at the end Day Five wins the most MerryBux. Likewise, the team with the next-highest score receives a lesser amount of MerryBux, and so on, with the last place team receiving nothing. In order to create the "bait" introduced to the focal actor, the "value" of the MerryBux ostensibly corresponds to the value of an Amazon gift card that each player receives at the end of the game. Team winnings are to be distributed among the team members in proportions determined by the Team Leader.

Ethical Dilemma Created Through Bait Treatment

In addition to the winnings based on the team performance from the game competition, a further financial incentive (i.e., 200 additional MerryBux as "bait") is offered to Team Leaders by the Game Master in the treatment groups at the end of Day Two (Table 1). This incentive is not tied to team performance *per se*, but is offered with two conditions attached: i) the Team Leader must keep this a secret from her/his team; and ii) if his/her team wins (i.e., finishes in first place), the Team Leader must split the incentive equally with all team members. However, if the team does not finish in first place, the Team Leader has no obligation to share the incentive at all.

With the introduction of this extra financial incentive, Team Leaders in the treatment groups are thus presented with an ethical dilemma: whether i) to collaborate with their teammates to achieve the best outcome and, if they win first place, share the incentive with them, or ii) to undermine the team's collaborative efforts—whether actively (e.g., by failing to turn in puzzles on time) or passively (e.g., by not correcting the answer when s/he knows it to be wrong)—accept less in team winnings, and keep them for him/herself (Ho, Hancock, Booth, Burmester, et al., 2016; Ho & Warkentin, 2017).

Data Collection and Analysis

Our objective in this study is not to identify deceptive intent *per se*, but to identify deceptive intent based on words used within groups' online interaction. More specifically, we not only analyze the focal deceiver's words, but also the

DOI: 10.1002/asi

words from interacting group members. Our data collection and analysis were aimed at examining the collective reaction of a team when a focal actor is presented with an ethical dilemma involving an act of deception. Rather than a post-hoc analysis of the participants' perceptions (i.e., *after* the deceptive act), we collect and objectively analyze conversations before, during, and after the deceptive act. In addition, our data collection and analysis also targets the reactions/changes within the focal actor's conversations before, during, and after the deceptive act.

Data Collection

Two separate research experiments were conducted using the research design described above, but each of the two studies was carried out at a different research institution. Experimental data for Study 1 were collected from the Blackboard system in 2008, with research protocol #07-276 approved by the Syracuse University Institute Review Board. Experimental data for Study 2 were collected from the Google+ Hangout platform, with an interface designed at Florida State University (FSU) in 2014 with research protocols #2013.19010 and #2014.12194 approved by the FSU Human Subject Committee. Participants in both studies were geographically dispersed across the United States. Participants were primarily students in undergraduate and graduate programs, some of which were also full-time employed.

In Study 1, participants ranged in age between 20 and 65 years of age. Twenty-six participants were recruited in Study 1, of which 64% ($n = 16$) were men and 36% ($n = 10$) were women. Twenty-seven participants were recruited in Study 2, of which 63% ($n = 17$) were men and 37% ($n = 10$) were women. In Study 2, participants ranged in age between 18 and 65 years of age. In both studies, data were collected over 5 consecutive days.

TABLE 2. Coding schema extracted from LIWC categories.

LIWC categories	Coding schema	Examples
Affective process	affect	happy, cried, abandon
Positive emotion	posemo	love, nice, sweet
Negative emotion	negemo	hurt, ugly, nasty
Anxiety	anx	worried, fearful, nervous
Anger	anger	hate, kill, annoyed
Sadness	sad	crying, grief, sad
Cognitive process	cogmech	cause, know, ought
Insight	insight	think, know, consider
Causation	cause	because, effect, hence
Discrepancy	discrep	should, would, could
Certainty	certain	always, never
Inclusive	incl	and, with, include
Exclusive	excl	but, without, exclude
Achievement	achieve	earn, hero, win
Negation	negate	no, not, never

Data Process and Analysis

All raw data sets were cleaned and preprocessed. During the data-cleaning phase, one subject's data from Study 1 was excluded from analysis because the subject did not complete the entire study. Thus, only 25 subjects' data were analyzed in Study 1, and all data from 27 participants were analyzed in Study 2. We separated all actors' conversations and kept them separate from Day 1 through Day 5.

The data collected in both studies were analyzed using the Linguistic Inquiry and Word Count (LIWC) software program (Pennebaker Conglomerates, 2007). LIWC employs customizable dictionaries to determine psychological state based on the types and categories of words used. LIWC analyzes text files to generate incremental word counts in each category (such as words reflecting positive or negative emotions, negations, self-references, and causal words) for a psychological picture of the speaker (affective processes, cognitive processes, social processes, etc.) based on the rate of occurrence of particular word(s). Several LIWC categories also have subcategories, offering more specific insight into word usage and the speaker. Consequently, most words can be coded into multiple LIWC categories (Pennebaker et al., 2003). For example, positive and negative emotion such as "love," "hate," "happy," and "sad," are further subcategorized as both a specific type of emotion (e.g., "hate" is subcategorized into "anger," etc.) as well as words relating to the main category, affective process (Table 2). Likewise, the category "cognitive processes" is also broken down into several subcategories (e.g., insight, causation, and certainty) (Table 2). Words and cues in these subcategories represent a specific instance of the overarching category (i.e., affective process or cognitive process, respectively).

Basic linguistic and psychological categories were extracted using LIWC (Table 2). Overall, the cleaned data set for Study 1 consists of a total word count of 20,452 words in 9,682 total lines of chat. The cleaned data set for Study 2 consists of a total word count of 13,086 words in 9,477 total lines of chat.

TABLE 3. Number of data sets after processing.

	Control/no bait		Treatment/bait	
Experiment design	Prebait	After bait	Prebait	After bait
Group interaction	Team A (Days 1-2)	Team A (Days 3-5)	Team C (Days 1-2)	Team C (Days 3-5)
Private (shadow)	Actor A (Days 1-2)	Actor A (Days 3-5)	Actor C (Days 1-2)	Actor C (Days 3-5)
Group interaction	Team B (Days 1-2)	Team B (Days 3-5)	Team D (Days 1-2)	Team D (Days 3-5)
Private (shadow)	Actor B (Days 1-2)	Actor B (Days 3-5)	Actor D (Days 1-2)	Actor D (Days 3-5)
Group interaction	Team E (Days 1-2)	Team E (Days 3-5)	Team F (Days 1-2)	Team F (Days 3-5)

DOI: 10.1002/asi

TABLE 4. Independent samples test Study 1 & 2.

	Independent samples test				
	t-test for equality of means				
	t	df	Sig. (2-tailed)	Mean difference	Std. error eifference
Negate (Study1)	2.258	10	.048	1.24600	.55189
	3.057	2.137	.085	1.24600	.40760
Negate (Study2)	2.506	12	.028	.69515	.27744
	2.572	3.306	.075	.69515	.27025

TABLE 5. Independent samples test Study 1.

	Independent samples test				
	t-test for equality of means				
	t	df	Sig. (2-tailed)	Mean difference	Std. error difference
Cogmech	−1.488	22	.151	−1.71167	1.15002
	−1.488	21.036	.151	−1.71167	1.15002
Insight	−2.993	22	.007	−1.25583	.41957
	−2.993	18.948	.007	−1.25583	.41957
Certain	2.674	22	.014	.68417	.25586
	2.674	21.491	.014	.68417	.25586
Incl	-.140	22	.890	-.05500	.39372
	-.140	19.004	.890	-.05500	.39372
Excl	-.458	22	.652	-.24250	.52975
	-.458	19.983	.652	-.24250	.52975
Achieve	2.092	22	.048	1.48667	.71078
	2.092	19.867	.050	1.48667	.71078

For each study, we derived two sets of clean data: i) the group interaction represents communications between the Team Leader and the team players while solving puzzles (group-based data), and ii) the private conversations between each Team Leader and the Game Master (individual-based data). Both data sets were further subdivided by timeframe for analysis: Days 1–2 (prebait) and Days 3–5 (postbait). In order to normalize the data for the group-based analysis (H1–H2), word count was converted to percentages. Since the data for the individual-based analysis (H3–H4) were already normalized, the word-count was retained. Table 3 depicts the breakdown of the data sets derived to test our hypotheses. We used IBM SPSSv22 (Armonk, NY) to run the paired sample and independent sample *t*-tests on the word count data set for the statistical significance of the data.

Results and Discussion

A variety of LIWC categories (as illustrated in Table 2) were selected, along with specific word counts, to analyze the textual data collected during interaction.

H1: Supported

Hypothesis 1 posits that under the influence of financial incentives (bait), deceptive actors will use more negation words than nondeceptive actors during synchronous CMC group interaction. We ran independent samples *t*-tests to compare words from the deceptive actor with words from nondeceptive actors in the same team, and observed the differences in use of negation words between deceptive and nondeceptive actors. With equal variance assumed, the relevant results are from the pooled variance estimator (Table 4). These results show that the use of negation words between deceptive and nondeceptive actors is statistically significant. Specifically, in Study 1, ($t(10) = 2.258$ $p = .048$) and in Study 2, ($t(12) = 2.506$ $p = .028$) (Table 4). We therefore submit that Hypothesis 1 is supported.

H2(a): Supported

Hypothesis 2(a) posits that the language-action cues associated with cognitive processes will be more prevalent in groups with a deceptive actor than in groups without one. We ran an independent samples *t*-test, assuming equal variance for each of the variables being compared, to observe differences in cognitive-process-related cues between these two categories of groups.

In Study 1, our results indicate that the words associated with insight and certainty (as part of cognitive processes) in the groups with a deceptive actor was statistically significant (Table 5) when compared to groups without a deceptive actor. Likewise, in Study 2 our results suggest that groups containing a deceptive actor tend to use more words associated with inclusivity and exclusivity (as part of cognitive process) than groups without a deceptive actor. The result overall was found statistically significant (Table 6). Accordingly, Hypothesis 2(a) is supported.

DOI: 10.1002/asi

TABLE 6. Independent samples test Study 2.

	Independent samples test				
	t-test for equality of means				
	t	df	Sig. (2-tailed)	Mean difference	Std. error difference
Cogmech	2.990	25	.006	3.41264	1.14153
	2.994	24.970	.006	3.41264	1.13967
Insight	1.694	25	.103	.80984	.47800
	1.715	23.616	.099	.80984	.47223
Certain	−1.498	25	.147	-.42791	.28563
	−1.505	24.947	.145	-.42791	.28428
Incl	3.614	25	.001	1.27522	.35285
	3.594	23.816	.001	1.27522	.35485
Excl	2.535	25	.018	1.00236	.39540
	2.500	21.030	.021	1.00236	.40096
Achieve	−2.119	25	.044	-.60665	.28634
	−2.143	23.842	.043	-.60665	.28311

TABLE 7. Paired samples *t*-test Study 1.

		Paired samples test					
		Paired differences					
		Mean	Std. deviation	Std. error mean	*t*	df	Sig. (2-tailed)
Pair 1	Affect - Affect_2	.40417	1.74455	.50361	.803	11	.439
Pair 2	Posemo - Posemo_2	−.14417	2.12648	.61386	−.235	11	.819
Pair 3	Negate - Negate_2	−1.13083	.97212	.28063	−4.030	11	.002

It should be noted that, in addition to our findings on the use of words associated with cognitive processes, our results from both Study 1 and Study 2 point to statistically significant differences in use of language-action cues associated with achievement. That is, words associated with achievement were more prevalent in groups with a deceptive actor than groups without a deceptive actor.

H2(b): Partially Supported

Hypothesis 2(b) states that the use of language-action cues associated with negation and affective processes in a group's interactions will increase after an actor within the group has accepted an incentive for deception. A paired-sample *t*-test was conducted to determine if there was any significant difference in the category counts before the bait was introduced to, and after accepted by, the deceptive actor.

Table 7 indicates the differences in the use of words associated with negation (coded as "negate") were found to be statistically significant in Study 1. On the other hand, Table 8, which presents our results for this same category in Study 2, shows the use of words associated with affective process (coded as "affect") and positive emotion (coded as "posemo") were also found to be statistically significant. We submit that the context and subject of discussions, as well as the participants themselves, would tend to drive the overall usage of words from a particular LIWC category in any case, so it is not necessarily to be expected that particular cues (e.g., affect, posemo, or negate) would be represented in a particular conversation. Accordingly, we suggest that the results of each study provide only partial support for H2(b).

Both individually and collectively, the results of Study 1 and Study 2 show a definite shift in the groups' interaction after bait is introduced. This can be observed in the language-action cues exhibited within the treatment groups. As noted, the Study 1 results showed that after bait was introduced, these groups tended to use more negation category words, suggesting disagreement within and among the team members and the Team Leader (data in Days 3–5). The negative *t* values in Table 7 indicate the increased usage of negation words after the bait was introduced. Similarly, the results from Study 2 indicate that the use of language-action cues related to affective process—which includes words related to positive emotions (such as happiness and love) and negative emotions (such as sadness, anxiety, fear, and hate)—increased after the bait was introduced, as shown in the negative *t* value in Table 8. This suggests that i) the Team Leader was attempting to establish emotional connections with team members in order to build and/or maintain trust, and/or ii) team members were somehow vaguely aware of something "going on" (either with the Team Leader specifically, or in general) and were concerned—or upset—about it. Unknowingly, team members were sensitive to the bait offered to their Team Leader.

Groups Influenced by Active Versus Passive Deceptive Actors

McCornack (1992) stated that deception involves manipulation of sensitive information, and proposed that deception

DOI: 10.1002/asi

TABLE 8. Paired samples *t*-test Study 2.

		Paired samples test					
		Paired differences					
		Mean	Std. deviation	Std. error mean	*t*	df	Sig. (2-tailed)
Pair 1	Affect - Affect_2	−2.22786	1.95652	.52290	−4.261	13	.001
Pair 2	Posemo - Posemo_2	−2.22071	1.90621	.50946	−4.359	13	.001
Pair 3	Negate - Negate_2	.04357	.75647	.20217	.216	13	.833

falls into one of two typologies: active (deception by acts of commission) or passive (deception by acts of omission). In the case of active deception, the deceptive actor uses affirmative actions to instigate the deceptive act, which may include lies, as well as intentional concealment of relevant information or providing partial/slanted information. In contrast, passive deception involves the deceptive actor failing to carry out an expected act. In either case, the deceptive actor faced an ethical dilemma and betrayed her/his trusted position. The results discussed above, taken collectively, provide a further insight into the group behaviors containing deceptive actors. We note first that in both Study 1 and Study 2, all Team Leaders engaged in deception—that is, none of them refused the "bait." However, the deceptive actors in each study engaged in different types of deception. We observed that all the deceptive Team Leaders from Study 1 were actively deceptive. Specifically, these Team Leaders took proactive steps to sabotage the team (e.g., intentionally submitting wrong answers despite the team's having arrived at correct answers). In contrast, all of the deceptive Team Leaders from Study 2 were passively deceptive. Specifically, the Team Leader's deceptive intent was often manifested by silence or failure to act (e.g., knowing that the answer from the team was wrong, but neglecting to correct it). Our results suggest that deception can be revealed by different linguistic cues in groups with an actively deceptive Team Leader versus groups whose Team Leaders are passively deceptive. This is illustrated by H2(a) and H2(b): language-action cues associated with insight, certainty, and negation were statistically significant in the case of the groups containing *active* deceptive actors (i.e., from Study 1), but not as significant for the groups containing *passive* deceptive actors (i.e., from Study 2). In contrast, language-action cues associated with cognition, exclusivity, inclusivity, affective process, and positive emotion were significant for groups containing *passive* deceptive actors (i.e., in Study 2) but not as significant for groups containing *active* deceptive actors (i.e., in Study 1).

H3(a) and H3(b): Not Supported

Hypothesis 3(a) states that the language-action cues of a deceptive actor's communications with his/her group will differ from those of a nondeceptive actor. An independent sample *t*-test was run to determine if any such differences could be observed, and the results indicate that no significant differences were observed. Likewise, Hypothesis 3(b) posits that language-action cues used by a deceptive actor in communicating with her/his group will be different before an incentive for deception has been introduced, when compared with after an incentive has been accepted. A paired-sample *t*-test was run to determine if any such differences could be observed and, again, the results indicate that no significant differences were observed.

Although neither H3(a) nor H3(b) were supported, this is consistent in general with the literature. These two hypotheses are essentially examining how a deceptive actor manages to "act natural" during deception. As to H3(a), our observations suggest that deceptive actors were able to conceal their deceptive intent such that their language-action cues are effectively the same as those of nondeceptive actors. And, as to H3(b), our observations indicate that deceptive actors were able to maintain a consistent communication style and language-action cues usage with their respective groups before and after accepting the "bait," so group members would not sense any change in their interactions, thereby concealing their deceptive intent.

H4(a): Supported

Hypothesis 4(a) states that a motivated deceptive actor will use more language-action cues associated with negation and cognitive process in the course of communicating with an overseer (i.e., Game Master) than a nondeceptive actor will. To test this hypothesis, private "shadow" communications between Game Master and deceptive Team Leaders were compared with the communications between the Game Master and nondeceptive Team Leaders.

In Study 1, we ran an independent samples *t*-test (assuming equal variance) on the categories of insight ($t(2) = -9.000$, $p = .012$) and causation ($t(2) = -8.043$, $p = .015$) between deceptive and nondeceptive actors and found a statistically significant difference between these cues as used by the two sets of actors (Table 9). Likewise, in Study 2 we found the categories causation ($t(2) = -4.919$, $p = .039$) and negation ($t(2) = -7.00$, $p = .020$) to be significant for the same hypothesis (Table 10). Thus, H4(a) is supported by these data.

In contrast to Hancock et al.'s (2008) findings, the results show that deceptive actors use words from the categories of causation and negation more frequently than nondeceptive actors do. In other words, when negotiating and making a deal with the Game Master, the deceptive actors tend to use more words related to thinking and reasoning in negative tones than did nondeceptive actors. The negative *t* value also indicates a comparative increase in the usage of the

DOI: 10.1002/asi

TABLE 9. Independent samples test Study 1.

	Independent samples test				
	t-test for equality of means				
	t	df	Sig. (2-tailed)	Mean difference	Std. error difference
Insight_2	−9.000	2	.012	−18.00000	2.00000
	−9.000	1.000	.070	−18.00000	2.00000
Cause_2	−8.043	2	.015	−14.50000	1.80278
	−8.043	1.742	.022	−14.50000	1.80278
Negate_2	−3.677	2	.067	−13.00000	3.53553
	−3.677	2.000	.067	−13.00000	3.53553

TABLE 10. Independent samples test Study 2.

	Independent samples test				
	t-test for equality of means				
	T	df	Sig. (2-tailed)	Mean difference	Std. error difference
Insight	−1.800	2	.214	−4.50000	2.50000
	−1.800	1.000	.323	−4.50000	2.50000
Cause	−4.919	2	.039	−5.50000	1.11803
	−4.919	1.471	.069	−5.50000	1.11803
Negate	−7.000	2	.020	−3.50000	.50000
	−7.000	1.000	.090	−3.50000	.50000

negation, insight, and causation words from the deceptive actors when compared to those of the nondeceptive actors during private negotiations with the Game Master.

H4(b): Partially Supported

Hypothesis 4(b) states that a deceptive actor will use more language-action cues associated with affective process and cognitive process in communicating with an overseer after having accepted an incentive for deception. In other words, we can expect that deceptive actors would alter their behavior by using more inclusive and cognition words after an incentive for deception has been introduced. To examine this question, we first divided the data from the deceptive actors' shadow communications (when making a deal with the Game Master) into pre- and postincentive subsets.

In Study 1, the paired-samples *t*-test was conducted, and the results indicate a statistically significant difference in the use of words associated with cognitive processes ($t(1) = -16.600$, $p = .038$), affective process ($t(1) = -26.00$, $p = .024$), and positive emotion ($t(1) = -15.00$, $p = .042$) as between pre-and postincentive communications (Table 11). Likewise, in Study 2 our findings suggest that there is a statistically significant difference in the use of inclusivity ($t(1) = -15.00$, $p = .042$) as between pre- and postincentive communications (Table 12). Thus, H4(b) is supported.

Our findings with respect to H4(a) and H4(b) collectively indicate a behavioral shift in deceptive actors after the bait has been introduced. Deceptive actors used more cognition and inclusive words during negotiations with the Game Master, while the negative *t* value on the inclusive language-action cues suggests that deceptive actors were friendlier with the Game Master after the bait was introduced.

Summary of Results

Our findings from this study lead us to conclude first and foremost that it is possible to recognize an actor's deceptive activity based on subtle distinctions in the words they use. During group interaction, deceptive actors tended to use more *negation* words than other members in the same group (H1). When deceptive actors are facing an ethical dilemma (i.e., deciding whether or not to accept the bait), the group's interaction with them tends to include more words involving *cognitive process*, *insight*, *certainty*, *inclusivity*, *exclusivity*, and *achievement* (H2).

The results also indicate that, consistent with the literature, there is no statistically significant difference in a deceptive actor's behavior (as manifested in language-action cues) before and after an incentive for deception is introduced (H3[a] and H3[b]). This implies that deceptive actors were fairly effective in concealing their deceptive intent and maintaining consistent behaviors and use of language action cues.

Finally, our results suggest that there is indeed a behavioral shift on the part of a deceptive actor in his/her interactions (negotiations) with the Game Master/overseer, following the actor's acceptance of an incentive to deceive her/his team. In particular, deceptive actors tended to use more words from the categories of *insight*, *causation*, and *negation* when interacting with the Game Master than nondeceptive actors did (H4a). In addition, deceptive actors tended to use more *cognition* and *inclusive* words while

DOI: 10.1002/asi

TABLE 11. Paired samples test Study 1.

		Paired samples test					
		Paired differences					
		Mean	Std. deviation	*t*	df	Sig. (2-tailed)	Std. error mean
Pair 1	Cog - Cog_2	−124.50000	10.60660	7.50000	−16.600	1	.038
Pair 2	Inc - Incl_2	−24.50000	3.53553	2.50000	−9.800	1	.065
Pair 3	Aff - Aff_2	−52.00000	2.82843	2.00000	−26.000	1	.024
Pair 4	PoEmo - PoEmo_2	−45.00000	4.24264	3.00000	−15.000	1	.042

TABLE 12. Paired samples test Study 2.

		Paired samples test					
		Paired differences					
		Mean	Std. deviation	*t*	df	Sig. (2-tailed)	Std. error mean
Pair 1	Cog - Cog_2	−31.00000	18.38478	13.00000	−2.385	1	.253
Pair 2	Inc - Inc_2	−7.50000	.70711	.50000	−15.000	1	.042
Pair 3	Aff - Aff_2	−17.00000	39.59798	28.00000	-.607	1	.653
Pair 4	PoEmo - PoEmo_2	−17.50000	34.64823	24.50000	-.714	1	.605

negotiating the incentive for deception offered by the Game Master (H4b).

Limitations

As suggestive as our results are, we must nonetheless acknowledge certain limitations with respect to sample size and game management. In particular, in terms of our analysis of H3(a) and H3(b), our sample size was insufficient to "rule out" the potential significance of individual-level interaction cues in detecting deception. However, these insufficient data reflect the fact that when an organizational insider who faces a dilemma within a group (e.g., Snowden's dilemma at NSA [Toxen, 2014]), it is rather challenging to detect such a passive act of deception. Nonetheless, our research design is innovative to mimic ethical dilemma in a computer-mediated collaborative group context. Moreover, this design examines group/interpersonal interactions in the context of an ongoing act of deception. This design allows our relatively small sample size to be generalizable to a larger population.

In terms of game management, we note again that in Study 1 all deceptive actors were *actively* deceptive, while in Study 2 they were all *passively* deceptive. We believe that this is a by-product of the specific implementation of the game rules by the particular overseer (i.e., Game Master) in each case. We also believe that it would be of interest to confirm our findings concerning the differences between passive and active deceptive actors by attempting to redesign the game so that we have an opportunity for a mix. Specifically, future iterations of the study will involve a tiered incentive structure—so that deceptive Team Leaders who are seen to be actively deceptive would be offered and would receive a larger incentive than those who are passively deceptive.

Lastly, we must note perhaps the largest limitation: As much as we attempted to make the interactions in this game reflect or mimic the real world, it is nonetheless still just a simulation in a world supported by CMC technologies. The deceptive actor knows, even if only subconsciously, that the consequences of undermining his/her team in this context are materially different from doing the same thing in the real world. In short, a research simulation can only go so far to help us predict what will happen in the "real" world. We willingly accept this limitation not only for reasons of convenience, but for ethical considerations.

Conclusion, Contribution, and Future Works

From time to time, we have all faced ethical dilemmas that force us to choose whether to pursue a course of action that would result in personal benefit at the expense of others, or to pursue a course of action resulting in a benefit to the larger group, perhaps at our own expense. Particularly when we choose the former, concealment and deceiving the group as to our intentions are critical to our success. While identifying deceptive intent is challenging in CMC, the results of our study illustrate that we can identify certain patterns and shifts in the occurrences of certain categories of words used by a focal actor that may be an early indicator of intent to deceive. Our findings confirm that we can detect shifts in a deceptive actor's intent when facing an ethical dilemma in a collaborative group context through language-action cues.

The methodology presented in this paper is both complex and novel. In particular, we note that the online game design—in both studies—was developed and deployed in-house, specifically for the study. We compared data across platforms (Blackboard vs. Google Hangouts) over several years (2008 vs. 2014). We processed data that were not participants' subjective perceptions, but their actual textual

DOI: 10.1002/asi

conversations. This is a true representation of group information behavior and interactions while incorporating an individual's subtle ethical dilemma. The group interaction was objectively represented by the cleaned, truncated, and categorized language-action cues. The novelty of this study and its contribution resides in its use of linguistic analysis tools to objectively examine users' information behavior in the context of a violation of group trust by a deceptive actor. Specifically, our research design achieves this by capturing data at both the group interaction level (words/conversations between and among group members before and after the focal actor commits to deception), as well as the individual interaction level (the communications between the Team Leader and the overseer). Thus, our data collection essentially provides a 360° view of group dynamics during interaction with deceptive actors. We can therefore compare the statistical significance of group-level cues and individual-level cues side-by-side to evaluate the availability and prevalence of certain cues to deception.

In sum, we believe our research provides useful insights into the detection of shifts and changes of deceptive intent in a CMC environment as reflected in language-action cues. Use of words connoting negation, words associated with cognitive processes (specifically insight and causation), or words associated with affective processes (specifically, positive emotions) were found to be among the more indicative language-action cues for this purpose. Further, our research supports the feasibility of an automated process or system to detect deceptive communication in CMC for the purpose of identifying a potential deceptive insider before they act, and this could provide the basis for an unobtrusive means to preempt insider threats.

Future work will seek to identify additional language-action cues that suggest deceptive intent within the context of group communication. In addition, our future studies will explore different machine-learning approaches of logistic regression modeling, decision trees, and support vector machine (SVM) analysis with larger sample sizes to identify significant deception-based language-action cues and patterns. This study contributes to the domain of social computing, behavioral modeling, and the prediction of computer-mediated deception. Humans alone are not very effective at determining deception on their own, but the analogy of crowd-sourced humans as sensors in an interactive network for identifying potential deception is supported by our findings.

Acknowledgments

Shuyuan Mary Ho acknowledges the National Science Foundation Secure and Trustworthy Cyberspace (SaTC) Program for the support of EAGER #1347113 award (09/01/13—08/31/15), the Florida State University Council of Research & Creativity for support of Planning Grant PG #034138 (12/01/13—11/30/14) and the First Year Assistant Professor Grant FYAP #033114 (05/10/13—08/10/13). Professor Dr. Jeffrey T. Hancock acknowledges the National Science Foundation Secure and Trustworthy Cyberspace (SaTC) Program for support of EAGER #1347120 award (09/01/13—08/31/15).

References

Allen, D.G., Mahto, R.V., & Otondo, R.F. (2007). Web-based recruitment: Effects of information, organizational brand, and attitudes toward a Web site on applicant attraction. Journal of Applied Psychology, 92, 1696–1708.

Austin, J.L. (1962). Toward the developmetn of a psycholinguistic-based measure of insider threat risk of focusing on core word categories used in social media. In Proceedings of the Nineteenth Americas Conference on Informations, Chicago, Illinois, 15–17 August (pp. 1–8). Association for Information Systems.

Brown, C.R., Greitzer, F.L., & Watkins, A. (2013). Toward the development of a psycholinguistic-based measure of insider threat risk focusing on core word categories used in social media. In Proceedings of the Nineteenth Americas Conference on Informations, Chicago, Illinois, 15–17 August (pp. 1–8). Association for Information Systems.

Brown, C.R., Watkins, A., & Greitzer, F.L. (2013). Predicting insider threat risks through linguistic analysis of electronic communication. In *Proceedings of the 2013 46th Hawaii International Conference on System Sciences* (pp. 1849–1858). IEEE. doi:10.1109/HICSS.2013.453

Buller, D.B., & Burgoon, J.K. (1996). Interpersonal deception theory. Communication Theory, 6, 203–242.

DePaulo, B.M., Ansfield, M.E., Kirkendol, S.E., & Boden, J.M. (2004). Serious lies. Basic and Applied Social Psychology, 26, 147–167.

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

DePaulo, B.M., Kashy, D.A., Kirkendol, S.E., Wyer, M.M., & Epstein, J.A. (1996). Lying in everyday life. Journal of Personality and Social Pyschology, 70, 979–995.

DePaulo, B.M., Lindsay, J.J., Malone, B.E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. Pyschological Bulletin, 129, 74–112.

Ekman, P., & Friesen, W.V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage and coding. Semiotica, 1, 49–98.

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

Greitzer, F.L., Kangas, L.J., Noonan, C.F., Brown, C.R., & Ferryman, T. (2013). Psychosocial modeling of insider threat risk based on behavioral and word use analysis. e-Service Journal, 9, 106–138.

Greitzer, F.L., Kangas, L.J., Noonan, C.F., Dalton, A.C., & Hohimer, R.E. (2012). Identifying at-risk employees: Modeling psychosocial precursors of potential insider threats. In *Proceedings of the 2012 45th Hawaii International Conference on System Sciences* (pp. 2392–2401). IEEE. doi:10.1109/HICSS.2012.309

Hancock, J., Birnholtz, J., Bazarova, N., Guillory, J., Perlin, J., & Amos, B. (2009). Butler lies: Awareness, deception and design. In *Proceedings of the 2009 SIGCHI Conference on Human Factors in Computing Systems (CHI'09)* (pp. 517–526). New York: ACM. doi: 10.1145/1518701.1518782.

Hancock, J., Curry, L.E., Goorha, S., & Woodworth, M. (2008). On lying and being lied to: A linguistic analysis of deception in computer-medicated communication. Discourse Process, 45, 1–23.

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, 327–343.

Hirschberg, J. (1993). Pitch access in context: Predicting intonational prominence from text. Artificial Intelligence, 63, 305–340.

Hirschberg, J. (2002). Communication and prosody: Functional aspects of prosody. Speech Communication, 36, 31–43.

Hirschberg, J., & Litman, D. (1993). Empirical studies on the disambiguation of cue phrases. Computational Linguistics, 19, 501–503.

Ho, S.M., & Benbasat, I. (2014). Dyadic attribution model: A mechanism to assess trustworthiness in virtual organizations. Journal of the American Society for Information Science and Technology, 65, 1555–1576.

Ho, S.M., Fu, H., Timmarajus, S.S., Booth, C., Baeg, J.H., & Liu, M. (2015). Insider threat: Language-action cues in group dynamics. In *Proceedings of the 2015 ACM SIGMIS Conference on Computers and People Research* (pp. 101–104). New York: ACM.

Ho, S.M., Hancock, J.T., Booth, C., Burmester, M., Liu, X., & Timmarajus, S.S. (2016). Demystifying insider threat: Language-action cues in group dynamics. In *Proceedings of the 2016 49th Hawaii International Conference on System Sciences (HICSS-49)* (pp. 2729–2738). IEEE.

Ho, S.M., Hancock, J.T., Booth, C., & Liu, X. (2016). Computer-mediated deception: Strategies revealed by language-action cues in spontaneous communication. Journal of Management Information Systems, 33, doi:10.1080/07421222.2016.1205924

Ho, S.M., Hancock, J.T., Booth, C., Liu, X., Liu, M., Timmarajus, S.S., & Burmester, M. (2016). Real or Spiel? A decision tree approach for automated detection of deceptive language-action cues. In *Proceedings of the 2016 49th Hawaii International Conference on System Sciences (HICSS-49)* (pp. 3706–3715). IEEE.

Ho, S.M., Hancock, J.T., Booth, C., Liu, X., Timmarajus, S.S., & Burmester, M. (2015). Liar, liar, IM on fire: Deceptive language-action cues in spontaneous online communication. IEEE Intelligence and Security Informatics (pp. 157–159). IEEE.

Ho, S.M., & Hollister, J.M. (2013). Guess who? An empirical study of gender deception and detection in computer-mediated communication. In Proceedings of the 76th American Society for Information Science and Technology (pp. 1–4). American Society for Information Science.

Ho, S.M., Liu, X., Booth, C., & Hariharan, A. (2016). *Saint or Sinner*? Language-Action Cues for Modeling Deception Using Support Vector Machines. In: Xu K., Reitter D., Lee D., Osgood N. (eds) Social, Cultural, and Behavioral Modeling. Lecture Notes in Computer Science, vol 9708. (pp. 325–334). Springer, Cham.

Ho, S.M., & Warkentin, M. (2017). Leader's dilemma game: An experimental design for cyber insider threat research. Information Systems Frontiers, 19, 377–396. doi:10.1007/s10796-015-9599-5

Hoffman, A.M. (2002). A conceptualization of trust in international relations. European Journal of International Relations, 8, 375–401.

Hosmer, L.T. (1995). Trust: The connecting link between organizational theory and philosophical ethics. Academy of Management Review, 20, 379–403.

Joinson, A.N., & Dietz-Uhler, B. (2002). Explanations for the perpetration of and reactions to deception in a virtual community. Social Science Computer Review, 20, 275–289.

Jones, S., & Marsh, S. (1997). Human-computer-human interaction: Trust in CSCW. ACM SIGCHI Bulletin, 29, 36–40.

McCornack, S.A. (1992). Information manipulation theory. Communication Monographs, 59, 1–16.

Mitnick, K.D., Simon, W.L., & Wozniak, S. (2002). The Art of Deception: Controlling the Human Element of Security (pp. 1–368). In C. Long, N. Stevenson, & J. Atkins (Eds.). Indianapolis, IN: John Wiley & Sons.

Pennebaker Conglomerates. (2007). Linguistic inquiry and word count (LIWC) software: Pennebaker Conglomerates, Inc. Retrieved from http://www.liwc.net/

Pennebaker, J.W., & King, L.A. (1999). Linguistic styles: Language use as an individual difference. Journal of Personality and Social Pyschology, 77, 1296–1312.

Pennebaker, J.W., Mehl, M.R., & Niederhoffer, K.G. (2003). Psychological aspects of natural language use: Our wods, our selves. Annual Review of Psychology, 2003, 547–577.

Rotter, J.B. (1967). A new scale for the measurement of interpersonal trust. Journal of Personality, 35, 651–665.

Rotter, J.B. (1980). Interpersonal trust, trustworthiness and gullibility. American Psychologist, 35, 1–7.

Rubin, V.L. (2010). On deception and deception detection: Content analysis of computer-mediated stated beliefs. In Proceedings of the American Society for Information Science and Technology (pp. 1–10). Wiley.

Rubin, V.L., & Camm, S.C. (2013). Deception in video games: Examining varieties of griefing. Online Information Review, 37, 369–387.

Rubin, V.L., & Lukoianova, T. (2015). Truth and deception at the rhetorical structure level. Journal of the Association for Information Science and Technology, 66, 905–917.

Taylor, P.J., Dando, C.J., Ormerod, T.C., Ball, L.J., Jenkins, M.C., Sandham, A., & Menacere, T. (2013). Detecting insider threats through language change. Law and Human Behavior, 37, 267–275.

Toma, C., & Hancock, J. (2010). Reading between the lines: Linguistic cues to deception in online dating profiles. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work (CSCW'10)* (pp. 5–8). ACM.

Toma, C., & Hancock, J. (2012). What lies beneath: The linguistic traces of deception in online dating profiles. Journal of Communication, 62, 78–97.

Toxen, B. (2014). The NSA and Snowden: Securing the all-seeing eye. Communication of the ACM, 57, 44–51.

Welch, E.W., Hinnant, C.C., & Moon, M.J. (2005). Linking citizen satisfaction with e-government and trust in government. Journal of Public Administration Research and Theory, 15, 371–391.

Whitty, M.T., Buchanan, T., Joinson, A.N., & Meredith, A. (2012). Not all lies are spontaneous: An examination of deception across different modes of communication. Journal of the American Society for Information Science and Technology, 49, 208–216.

Wright, R.T., & Marett, K. (2010). The influence of experiential and dispositional factors in phishing: An empirical investigation of the deceived. Journal of Management Information Systems, 27, 273–303.

Zhou, L., Burgoon, J.K., Twitchell, D.P., Qin, T., & Nunamaker, J.F. Jr. (2004). A comparison of classification methods for predicting deception in computer-mediated communication. Journal of Management Information Systems, 20, 139–165.

Zhou, L., Twitchell, D.P., Qin, T., Burgoon, J.K., & Nunamaker, J.F. Jr. (2003). An exploratory study into deception detection in text-based computer-mediated communication. In *Proceedings of the 2003 36th Hawaii International Conference on Systems Sciences (HICSS-36)* (pp. 1–10). IEEE.

Zhou, L., & Zhang, D. (2004). Can online behavior unveil a deceiver? An exploratory investigation of deception in instant messaging. In Proceedings of the 2004 37th Hawaii International Conference on System Sciences (HICSS-37) (pp. 1–9). IEEE.

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.