Computers in Human Behavior 91 (2019) 33–41

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Full length article

Context in a bottle: Language-action cues in spontaneous computer-mediated deception

Shuyuan Mary Ho[a,*], Jeffrey T. Hancock[b]

[a] College of Communication and Information, Florida State University, 250 Shores Building, 142 Collegiate Loop, Tallahassee, FL, 32306-2100, USA
[b] Department of Communication, Stanford University, Building 120, 450 Serra Mall, Stanford, CA, 94305-2050, USA

ARTICLE INFO

Keywords:
Human computer interaction
Computer-mediated deception
Language-action cues
Sociotechnical systems
Mixed model ANOVA

ABSTRACT

Context always matters in online deception. This study conceptualizes language-action cues as a dynamic and interactive representation of context, and explores indicators of interpersonal deception in spontaneous computer-mediated communication. An online game was designed and developed to simulate computer-mediated deception scenarios in pairwise interactive dialogue that is difficult to capture in the real world. The study was conducted with 40 research participants in 80 game sessions during 2014 and 2015. Research results demonstrate the efficacy of language-action cues in revealing deception during pairwise interaction. The study contributes a novel perspective on viewing interaction using language-action cues to define context, with a sociotechnical research design that can computationally differentiate between deceivers and truth-tellers in spontaneous pairwise interaction.

1. Introduction

As users of computer-mediated communication (CMC) are exposed to an increasing number and variety of risks associated with online deception, it has become more and more important—and increasingly challenging—for users to guard against deceptive activities (e.g., fake review, opinion spam, spear phishing, identity fraud). The key is the ability to correctly interpret and codify the underlying intent of the message sender. In face-to-face (F2F) communication, the assessment of context is informed by words that are accompanied by other physical cues such as body language and facial expressions. CMC, in comparison, is "cue lean." Thus, to derive context, a message receiver has only words to consider during a message exchange. Context is often difficult to ascertain; however, as Dourish (2004) states, context is a relational—or even an occasioned property. Context is not static, but dynamically defined by activity and the associated interaction. In CMC, we may only derive context (and hence intent) from the interaction itself—including the words used, and any other associated non-verbal cues. Thus, our research question is: *How do language-action cues derive context, and further the identification of spontaneous computer-mediated deception*?

In this paper, we first discuss the creation of interaction context through language-action cues, and then argue that language-action cues can reveal and visualize communicative intent. Different types of communication modes and media can have an impact on computer-mediated deception, which is further examined. Our research considers and models a sociotechnical system that represents synchronous computer-mediated interaction. An online game was designed and developed to simulate interpersonal computer-mediated deception in pairwise interaction. We collected and analyzed the data to uncover patterns that can provide subtle cues to deceptive intent. Specifically, linear, logistic regression and mixed model ANOVA approaches were deployed to analyze the efficacy and accuracy of certain language-action cues for detecting computer-mediated deception in synchronous, spontaneous communications. The paper concludes with some reflections on implications, limitations, as well as potential directions for future research.

2. Computer-mediated deception

Online communication is interpersonal and interactive. It involves a sender and one or more receiver(s) who are engaged in an (more-or-less) interactive exchange. Each iteration of an interaction between a sender and a receiver helps to shape the context of communication. Within this exchange, there is often an opportunity for a message sender to influence the receiver(s) actions or beliefs. Accordingly, Miller, Deturck, and Kalbfleisch (1983) described deceptive communication as "… a general persuasive strategy that aims at influencing

* Corresponding author.
E-mail addresses: smho@fsu.edu (S.M. Ho), hancockj@stanford.edu (J.T. Hancock).

https://doi.org/10.1016/j.chb.2018.09.008
Received 26 July 2017; Received in revised form 24 January 2018; Accepted 12 September 2018
Available online 25 September 2018
0747-5632/

the beliefs, attitudes and behaviors of others by means of deliberate message distortions" (p. 99). Miller and Stiff (1993) and Stiff (1996) similarly characterized deceptive communication as an act involving the intentional use of persuasive strategies and activities to manipulate the receiver. In short, deceptive communication fundamentally means purposefully concealing the truth, either by omission or commission (Ekman and Friesen, 1969), and can be viewed as "a message knowingly transmitted by a sender to foster a false belief or conclusion by the receiver" (Buller and Burgoon, 1996). Buller and Burgoon's (1996) interpersonal deception theory (IDT) extends this understanding, explaining how a deceiver strategically shapes his/her communication behavior in response to the perceptions and suspicions of the receiver (s). The iterative interaction shapes the context of communication.

2.1. Interaction context

The "context" of a communication can include many different internal and external factors or elements. Dourish (2004) suggested that context can be interpreted subjectively by phenomenological theorists, broadly by critical theorists, or objectively by positivists. These different views infer that deriving context is not only a representational problem, but also a problem typically associated with interaction. Contemporary linguists such as Attardo (1994) suggested co-text analysis (interpreting linguistic materials in surrounding text) to identify the figurative, hidden interpretation of the human communication (p. 70). Co-text linguistic analysis is not the same as context codification (i.e., the references outside of the text through communicators' understanding of the genre, situations and discussion) as suggested by Dourish (2004). In human-computer interaction, Dourish (2004) first identified four basic assumptions to implicitly describe context as a representational problem. While context can be captured and represented as a form of information that is delineable, stable, and explicit, Dourish (2004) differentiated that "the kind of thing that can be modeled however is not the kind of thing that context is" (p. 22) because context is an interactional problem. Dourish (2004) further raised another four implicit assumptions about context. First, context is a *relational property*, not just an object (e.g., data files) or an activity (e.g., an act of posting online). Context refers to something in between an object and an activity that is relevant to a particular situation. Second, the scope of context is dynamic in nature; that is, context changes dynamically based on the interaction. Third, because of its dynamic nature, context is an *occasioned* property; that is, context can be momentary, and is often relevant only to a particular instance or action. Finally, context arises from the activity itself; that is, context does not typically exist apart from the interaction or activity (i.e., context does not exist until the interaction produces or creates it). This view of context as an interactional problem is especially salient to our investigation of language-action cues in computer-mediated deception (Ho and Hancock, 2018; Ho, Hancock, Booth, and Liu, 2016). Dourish (2004) highlighted that context can be found in embodied action: "assign (ing) a central role to the meanings that people find in the world and the meanings of their actions there in terms of the consequences and interpretations of those actions" (p. 28). Our theoretical stance and investigation thus is also anchored in the notion that the communication parties essentially create their own context as they communicate. Context is dynamic and ubiquitous, and is constantly being shaped by the interactivity of communication. Context can be represented by words along with non-verbal communication cues employed by communicating parties, often resulting in a sense of trust or distrust.

2.2. Language-action cues in the context of deceptive communication

Context, as characterized by Dourish (2004), is created by interactivity. One way in which its dynamic nature can be represented is through the language-action cues of communicating parties. This is especially evident when studying deceptive communication, where numerous studies have been focused on associated cues (behavioral, contextual, verbal or textual) in a wide range of facets and contexts. Ekman and Friesen, 1969; Ekman and Osullivan 1991) proposed the identification of deception using nonverbal cues and subtle changes in facial expression. Studies on deception range from everyday lies (DePaulo, Kashy, Kirkendol, Wyer, and Epstein, 1996; Smith, Hancock, Reynolds, and Birnholtz, 2014), to serious lies (DePaulo, Ansfield, Kirkendol, and Boden, 2004). Deception research also has discovered identifiable verbal (DePaulo et al., 2003; Vrij, 2015), nonverbal (Burgoon and Buller, 1994; Burgoon, Buller, Dillman, and Walther, 1995; DePaulo and Kashy, 1998), and text-based (Zhou, Burgoon, Twitchell, Qin, and Nunamaker, 2004) cues in deceivers' interpersonal communication. Research on deception strategies (Buller and Burgoon, 1994, 1996; Whitty, Buchanan, Joinson, and Meredith, 2012), beneficiary (Whitty and Carville, 2008), perceived credibility (George et al., 2014, 2016), as well as cultural influences and media choice (Furner and George, 2012; Lewis and George, 2008) have collectively informed our understanding of "how" and "why" people deceive.

Text-based CMC is "cue lean," in that it lacks physical cues to deception that are often available in F2F communication. However, certain communication cues can nonetheless still be observed and catalogued within a CMC environment (Hancock et al., 2007, 2009; Zhou & Zhang, 2004). These communication features and cues, such as first-person references, words of emotion, expressions of inhibition, prepositions, and conjunctions, have all been shown to be indicators that can differentiate deceivers from truth-tellers (Hancock et al., 2009). Use of more—or fewer—sensory or spatiotemporal words, and changes in the diversity and complexity of language have also been shown to be indicative of deception (Newman, Pennebaker, Berry, and Richard, 2003). As in F2F, level of detail (less or more) may be suggestive of CMC deception—although relevance of detail appears to be more significant than "detail" *per se*. That is, deceivers in CMC tend to be wordier than truth-tellers, but the additional words (i.e., details) provided are not necessarily relevant or meaningful (Zhou and Zhang, 2004).

Another language-action cue that is important in CMC (also in F2F communication) is *immediacy* (i.e., ways in which a speaker can associate, or distance him/herself from the content of his/her message (Zhou and Zhang, 2008). *Immediacy* (whether verbal or non-verbal, and whether created physically or psychologically) is particularly important in detecting deception. In the physical environment, nonverbal immediacy cues include eye contact, body language, facial expression, etc. While these specific cues were first studied in a F2F environment, certain cues—such as delay in response—are also present in CMC, and may operate similarly in both environments to create a psychological distance between deceiver and his/her communication partner (Whitty et al., 2012).

Examining these cues for the purpose of developing an automated process that can codify language patterns to indicate deceptive intent, we find that many of them—including quantity and consistency of detail—are measureable in a dynamic exchange of text messages by focusing on specific psychological features such as insights, cognitive, affective and social processes. These cues are measureable in a dynamic exchange by focusing on specific communication features. Linguistic analysis programs, such as the Linguistic Inquiry and Word Count (LIWC) developed by Pennebaker and King (1999) can facilitate this automatic assignment of words into psychological constructs. Although classifying psychological states can get very complicated, it is possible to benchmark verbal indicators (e.g., word count and details of information disclosed) so as to capture certain nonverbal behaviors (latency and usage of expression words) in a CMC environment, which can then be statistically computed (Zhou and Zhang, 2004).

2.3. Communication modes and media

People interact differently when using different communication

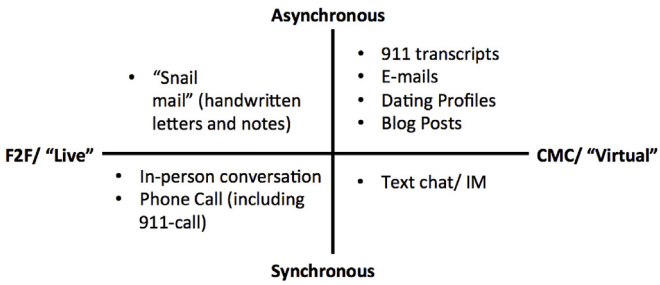

Fig. 1. Communication contexts.

modes and media. The mode and medium of the communication thus provide a mechanism that shapes people's communication behavior. Here, communication mode refers to whether the parties are interacting in real time (synchronous) or are communicating via messages exchanged back-and-forth over time (asynchronous). Moreover, the mode of any communication—whether it takes place in a "virtual" CMC environment, or in a "real" F2F environment—can also influence how one interacts and communicates.

Communication medium represents an important enabler of context. As illustrated in Fig. 1, certain media lends itself to a synchronous communication mode (e.g., instant-messaging/chat in CMC, and an in-person conversation in F2F settings). Other media options are designed primarily for an asynchronous mode of communication (e.g., e-mail in CMC, and a handwritten letter in F2F). Context is thus enabled through a combination of the medium and the mode used for communication.

2.4. Computer-mediated deception in context

Research into CMC deception cues has examined a variety of media types and modes of communication, while exploring the role of media choice as well as mode of communication. Both the mode of communication (i.e., synchronous (Zhou and Zhang, 2008) or asynchronous (Zhou et al., 2003)) and the specific medium chosen, may provide insight into to the type (i.e., planned or on-the-fly) and severity (i.e., serious or inconsequential) of a deceptive act (Whitty et al., 2012). One notable early study exploring this problem at a high level was done by Hancock, Thom-Santelli, and Ritchie (2004). Participants were asked to log their interactions and lies for seven (7) consecutive days. The results created a foundation for Hancock et al. (2004) "feature-based" model for deceptive CMC, which attempts to derive cues to deceptive communication by examining specific features—particularly looking at the communication mode (i.e., synchronous, asynchronous, or either/both) associated with the medium. Other salient factors that facilitate CMC-based deception include whether the medium records the communication, and whether the communication is distributed. A fundamental assumption of this model is that deception is often spontaneous, and therefore more likely to occur when media is "synchronous and distributed, but non-recordable" (Whitty et al., 2012). Here, we briefly outline studies that focus on both asynchronous and synchronous media types and communication modes in providing context.

2.4.1. Asynchronous communication

According to social distance theory (DePaulo et al., 1996), the social disapproval of deception and the psychological discomfort experienced by deceivers often results in deceivers' attempts to distance or separate themselves from their deception and the individual(s) they are attempting to deceive. Therefore, according to social distance theory, deceivers will tend to choose media that offers fewer cues to the receiver of the communication. As such, deceivers will be more likely to use an asynchronous mode of communication. Some studies examining deceptive language-action cues in asynchronous communication include deception in e-mails (Zhou et al., 2003), deceiving online dating profiles (Hancock et al., 2007; Toma and Hancock, 2010, 2012), fake hotel reviews (Ott, Choi, Cardie, and Hancock, 2011) and deception found in the transcripts of 911-calls (Burns and Moffitt, 2014).

A notable study examining language-action cues captured from e-mails was conducted by Zhou, Burgoon, et al. (2003) and Zhou, Twitchell, et al. (2003). This study evaluated deceptive cues in a team-based "desert survival scenario" game, and found that deceivers tended to be wordier when compared to truth-tellers—particularly in terms of using more verbs, modifiers and nouns in peripheral expressions that provide useless or irrelevant information. Moreover, Zhou and Zhang (2004) found that, in the context of asynchronous online communication, deceivers tend to be more active in language usage, while taking shorter pauses between messages. They were also more non-immediate than truth-tellers, using more group references and modal verbs. That said however, Zhou and Zhang (2004) made an important distinction between synchronous and asynchronous communication: the exercise in their particular experiment—which employed asynchronous communication (specifically, e-mail)—involved an element of persuasiveness on the part of the deceptive actor in the dyad. The exercise required the deceptive actors to be strategic, and the deception to be planned, which was better facilitated by an asynchronous means of communication. This finding emphasizes the importance of an appropriate understanding of context in codifying language patterns that can identify deceptive communication.

Another study examining CMC deception in an asynchronous mode looked at the truthfulness of online dating profiles. Toma and Hancock (2010) found that language-action cues involving emotion were more powerful in predicting deception than cognitive cues, with the notable exception that word count (a cognitive cue) was also highly significant. In this regard, Toma, Hancock, and Ellison (2009) revealed that participants tended to use deceptive information in their profiles sparingly to capture the emotion of profile viewers (Toma et al., 2009), but cognitive words appeared to be non-predictive. The discrepancy in these research findings reflects the fact that deception is context specific. In these highly controllable media environments, intentional deceivers may try to alleviate the cognitive burden and stimulate the emotional response associated with deception.

Ott et al. (2011) examined online hotel reviews as asynchronous CMC, and investigated the linguistic features that were most indicative of an authentic review. The results of their study suggest that truthful reviews included more "sensorial and concrete language" (especially concerning spatial configurations—i.e. overall room space and space usage), while deceptive/fake reviews included more superlatives. Such deceptive opinion spam is deliberately written to sound authentic. Ott, Cardie, and Hancock (2012) proposed a deception classifier to compare deception in six online review communities, and suggested that the prevalence of deception depends on the signaling costs (e.g., posting requirements). That is, the review communities with low signal costs tend to have more deception than communities with comparatively high signal costs. Ott, Cardie, and Hancock (2013) further identified the correlation between sentiment and deception in deceptive opinion spams. Both *positive* and *negative* opinion spam messages are used to manipulate perceptions of truth in user-generated online reviews.

Language-action cues in transcripts of 911 (emergency) telephone calls has also been studied by Burns and Moffitt (2014), to determine which cues are most indicative of bogus or fake calls. This study examines an asynchronous review of text collected in the course of a synchronous/real-time F2F interaction. Of note in the results from this study is that deceptive callers were found to show more "inhibition" (for example, delaying or telling the dispatcher to "hold on a minute"). In processing the transcriptions, deceptive callers were also found to use more words associated with immediacy (1st person pronouns) and non-immediacy (3rd person pronouns).

2.4.2. Synchronous communication

Deceptive communication is equivocal in nature (i.e., intentionally ambiguous), and thus open to interpretation by the receiver (i.e., the

individual a communicator is attempting to deceive). According to media richness theory, deceptive actors will tend to choose media types that provide for multiple cues, an opportunity for personalization, and immediate feedback (i.e., synchronous and spontaneous)—allowing them to ensure the equivocal nature of their message and adjust their deceptive communication strategy 'on the fly,' so as to obfuscate their deceptive intent (Daft, Lengel, and Trevino, 1987; Trevino, Lengel, and Daft, 1987). Thus, this theoretical framework—media richness theory—seems applicable to explicate the spontaneity of deception in synchronous communication. Four factors are used in determining the richness of the medium: feedback (spontaneous, immediate or delayed); number of cues available to the receiver (including social cues); language variety (i.e., the type and variety of symbols used to convey the particular message); and personal focus (i.e., infusing the message with personal feeling/emotions (Daft et al., 1987)). The richer the medium, the better it is able to convey equivocal messages.

Studies investigating CMC deception in a synchronous mode of communication have tended to focus on instant-messaging/chat. For example, Hancock, Curry, Goorha, and Woodworth (2008) found that—as with asynchronous communication—word count can be a significant predictor of deception. Participants in the dyad (consisting of one truthful and one deceptive partner) were given time (5 min) before play was to begin in which to review the fixed set of questions to be used, and plan their responses. Thus, the deceptive player was, in essence, given an opportunity to plan and prepare a strategy for implementing the intended deception(s). Further, Toma and Hancock (2010) suggest that—consistent with findings in the context of asynchronous communication (Hancock et al., 2008)—first person pronouns were used more by truth-tellers than deceivers, and deceivers use fewer self-references in CMC synchronous chat, but more third-person references. These findings also support social distance theory that suggests deceivers tend to create distance with their communicator.

Although chat features were incorporated in the above synchronous communication studies, the element of *spontaneity* was not designed into their research methodology. As the ability to discern and detect deception in any environment varies depends on many different factors, the importance of *spontaneity* is particularly important in how people respond and interact in real-time—with no pre-planning of any deceitful stories. Thus, as the availability of cues is reduced (limited to text) in CMC environment, our ability in identifying the spontaneous, synchronous computer-mediated deception has become particularly challenging (Ho, Hancock, et al., 2016).

3. Research design

To investigate *context-sensitive* and *spontaneity* in computer-mediated deception, we modeled *interaction* to better understand the role of context (Dourish, 2004). An online game system was designed and developed to mimic *spontaneity* in interpersonal deception scenarios, and to capture players' interactive conversations during pairwise interaction. This system creates a conceptual basis for further exploration of the dynamics of intentional deception. The use and efficacy of language-action cues in detecting context-sensitive deception, specifically in spontaneous synchronous chat-based CMC, was the focus of our investigation. Our approach further developed specific metrics for language cues and word choice, to analyze communication patterns that represent both deceptive and truthful communications. Specifically, the identification of text-based cues from these scenarios provides a means for understanding and measuring the linguistic parameters needed to detect online deception. It also enables us to observe how people lie successfully in different circumstances, and how deceivers create and manipulate communicative context to hide deceptive intent.

This interactive online game is called "*Real or Spiel*," which presents players with real-time interactive scenarios that require them to exchange either deceptive or truthful statements while using instantaneous, synchronous communication channels (Ho, Hancock, Booth, Liu, et al., 2016). The flow of player interaction and game play is depicted in Fig. 2. Each game scenario involves two participants (i.e., players) who are placed in randomly assigned pairings by the research team, and then randomly assigned an "*outer*" role (either initiating *speaker* or *detector*) in each gaming session. The initiating speaker in each scenario is also randomly assigned an "*inner*" role—either *saint* (truthful) or *sinner* (deceptive). The speaker establishes the *ground truth* prior to each scenario by truthfully answering questions on a selected topic from common knowledge domains. A total of 92 conversational topics—representing circumstances—were deployed in the online game, encompassing seven (7) categories: personal experience, finance, travel, entertainment, food, skills, and life style. A copy of the 92 topics is included in Appendix B. This provides a baseline for assessing the truthfulness or deceptiveness of subsequent responses to questions posed by the detector during the game scenario. At the end of each session, the detector tries to determine whether the speaker was being deceptive or truthful based on the question-and-answer exchanges.

Fig. 3 is a screenshot taken from a live game, including role assignment and a ground truth question.

4. Method

A mixed method was employed in this research. Qualitative data (i.e., players' conversations and interactions) was captured and analyzed quantitatively.

4.1. Data collection

Data for this study was collected in 2014 and 2015.[1] Each game session consisted of two players; a speaker and a detector. Data was collected across a total of 80 game sessions. 40 participants (22 males and 18 females) were randomly assigned into pairs, with each pair playing a total of 4 game sessions. The identity of each player's partner was not revealed to the player either before or after play, so each player was essentially interacting with a stranger. Players were all between 18 and 68 years of age, and were recruited using convenience sampling strategies, primarily from the student population of Florida State University. Players' names were replaced with pseudo-names to maintain privacy. Each game session lasted approximately 30 min, and consisted of approximately 4 role-play exchanges. At the end of each exchange, the players' alternate between roles automatically.

4.2. Data cleaning process

The data collected was cleaned and spell-checked.[2] The spell checker corrected most of the spelling errors in chat text. Common instant-messaging abbreviations were converted to their corresponding full written forms (e.g., LOL becomes "laugh/ing out loud"). We also excluded any pair's message containing fewer than 50 words in total exchange. The final dataset used in the analysis consisted of a total of 2196 lines of chat and 7271 words.

5. Data analysis

After the data had been cleaned, the linguistic cues from the data were extracted according to the categories established in the Linguistic Inquiry and Word Count (LIWC) tool (Newman et al., 2003; Pennebaker & King, 1999).

[1] The Florida State University's Institutional Review Board has approved human subject data collection (Protocols #2014.13490 and #2015.15885).

[2] All players communicated in English to play the game, although English may or may not have been their first language.

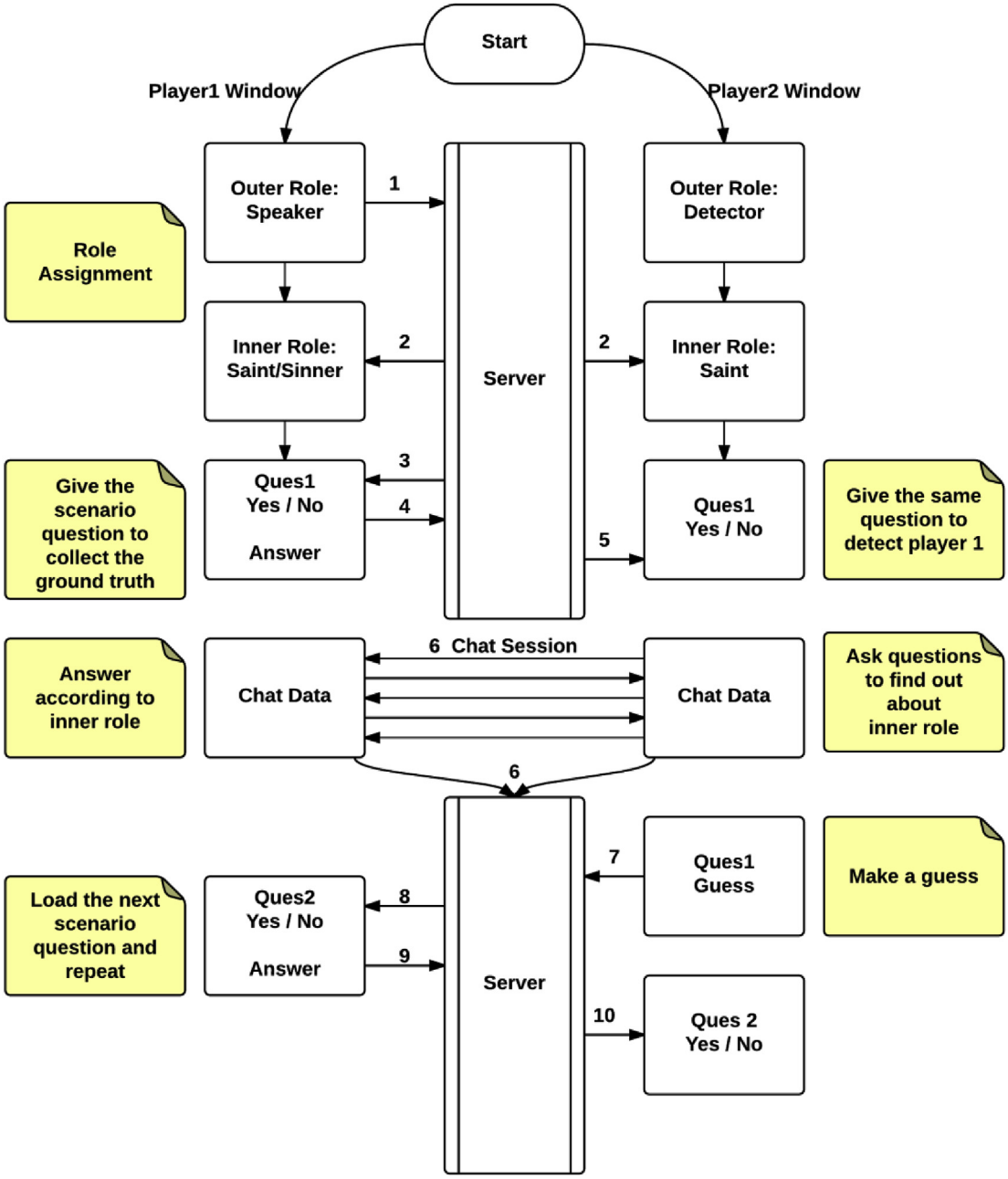

Fig. 2. Game design.

5.1. Context representation

The specific LIWC categories we included in our analysis are those set forth in Table 1. In order to avoid or minimize possible multicollinearity problems, cues directly corresponding to the "main" LIWC headings of "cogmech" and "affect" were ultimately excluded, in favor of including cues from their respective subcategories (i.e., posemo, negemo, certain, incl, excl, etc.).

The extracted cues included overall *word count*, words connoting *affective* processes (i.e., positive vs. negative emotion), and words connoting *cognitive* processes (i.e., words conveying certainty, inclusivity, exclusivity, discrepancy, insight and causation). Also examined were specific instances of *pronouns* (1st vs. 2nd person). Other cues explored further use of words of negation (i.e., no/not), quantifiers (i.e., few, many), and words connoting tentativeness. Finally, the role of *immediacy* was explored by examining response *time-lag*—i.e., the time between when the detector asks a question and the speaker answers. We note that the LIWC tool codes only individual words and does not take polysemy or abbreviations of words into account (e.g., I'd for 'I would' is not counted as discrepancy).

5.2. Linear regression analysis

We first performed a linear regression analysis to identify certain combinations of language-action cues as being predictive of deception. Three models of different combinations of cues were derived with statistical significance (Table 2). Overall, Model 3 seems to be preferable ($F = 3.103$, $p = 0.02$).

However, Table 3 illustrates the results of removing insignificant cues (as variables), and these results indicate that the three models are very similar in terms of predictive power. The R square for each model is very close. In fact, the adjusted R square is almost the same for both Model 2 and 3. The standard error of the estimate is very close between all three models. Table 4 illustrates the regression coefficients for both

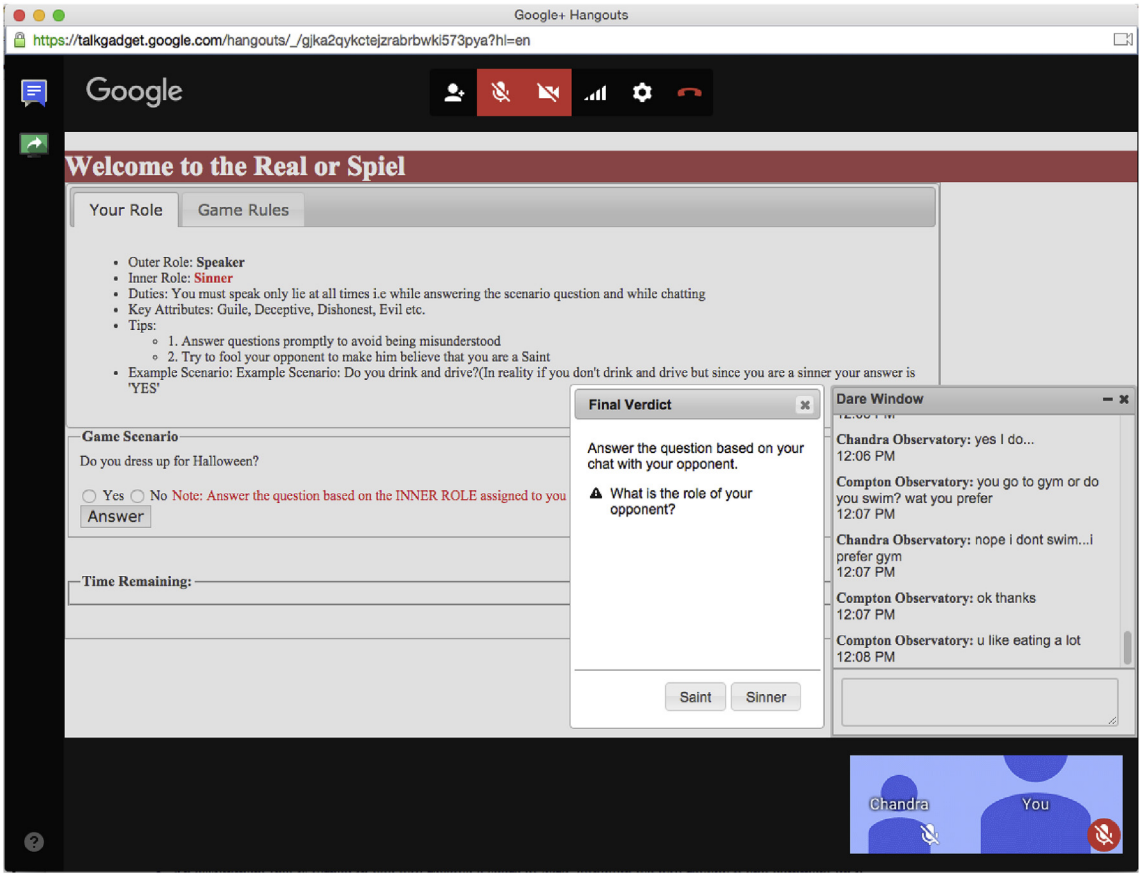

Fig. 3. Game interface.

Table 1
Language-action Cues extracted by LIWC.

LIWC Categories	CODING SCHEMA	Examples
Affective Process	affect	happy, cried, abandon
Positive Emotion	posemo	love, nice, sweet
Negative Emotion	negemo	hurt, ugly, nasty
Cognitive Process	cogmech	cause, know, ought
Certainty	certain	always, never
Inclusive	incl	and, with, include
Exclusive	excl	but, without, exclude
Discrepancy	discrep	should, would, could
Insight	insight	think, know, consider
Causation	cause	because, effect, since
Negations	negate	no, not, never
Pronouns	pronoun	
1st person singular	self-reference	I, me, myself
2nd person	other reference	you
Quantifiers	quant	few, many, much
Social	social	family, friends, humans
Tentative	tentat	maybe, perhaps, guess
Word Count	WC	n/a

Table 2
ANOVA summary for Models 1, 2 & 3.

ANOVA[a]						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6.894	16	.431	2.071	.021[b]
	Residual	13.106	63	.208		
	Total	20.000	79			
2	Regression	6.830	12	.569	2.896	.003[c]
	Residual	13.170	67	.197		
	Total	20.000	79			
3	Regression	6.685	11	.608	3.103	.002[d]
	Residual	13.315	68	.196		
	Total	20.000	79			

[a] Dependent Variable: Deceiver.
[b] Predictors: (Constant), insight, discrep, cause, I, WC, negemo, certain, Time_lag, incl, posemo, tentat, excl, quant, negate, social, you.
[c] Predictors: (Constant), insight, I, WC, negemo, Time_lag, incl, posemo, tentat, excl, quant, negate, you.
[d] Predictors: (Constant), insight, I, WC, negemo, incl, posemo, tentat, excl, quant, negate, you.

Word Count and *Insight* cues to be statistically significant across three models; however, the rest of the language-action cues, individually, were not statistically significant in predicting or identifying deception. Finally, we derived the collinearity diagnostics (Table 5) and identified that Model 1 has the highest eigenvalues (9.880) indicating the most viable model with appropriate dimensionality.

5.3. Logistic regression analysis

Based on the results of linear regression analysis, we chose Model 1 with the highest eigenvalue (9.880) that contains the dimensionality of all language-action cues as predictor variables, and ran a logistic regression analysis. Logistic regression analysis was based on the dichotomous nature of the outcome variable "Deceiver" (0 = truth-teller/ 1 = deceiver). The rule of parsimony was applied to assess the model fitness. The chi-square difference test determines the model that explains the best fit with the fewest variables. We ran three logistic regression models at different cutoff values, and all three models derive optimal chi-square of $\chi^2 = 35.8$, $p < 0.01$ (Table 6). All three logistic models were a good fit.

Similar to the results from linear regression analysis, the

Table 3
Linear regression model summary.

Model Summary[d]									
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.587[a]	.345	.178	.456	.345	2.071	16	63	.021
2	.584[b]	.342	.224	.443	-.003	.077	4	63	.989
3	.578[c]	.334	.227	.443	-.007	.741	1	67	.392

[a] Predictors: (Constant), insight, discrep, cause, I, WC, negemo, certain, time_lag, incl, posemo, tentat, excl, quant, negate, social, you.
[b] Predictors: (Constant), insight, I, WC, negemo, time_lag, incl, posemo, tentat, excl, quant, negate, you.
[c] Predictors: (Constant), insight, I, WC, negemo, incl, posemo, tentat, excl, quant, negate, you.
[d] Dependent Variable: Deceiver.

independent/predictor variables such as *Word Count* ($p = 0.008$) and *Insight* ($p = 0.02$) cues were found to be statistically significant (Table 7). However, the rest of the language-action cues, individually, were not statistically significant in predicting or identifying deception.

While multicollinearity may provide one possible explanation for this apparently contradictory result (i.e., non-significant predictors, but strong overall model), we had already eliminated any variables that might have created such a problem. Instead, we attribute this phenomenon to the nature of communication in the context of computer-mediated deception. Certain combinations of words provide a context that can be indicative of deception, and we found that the patterns created by these word combinations are more significant than the words themselves. Thus, even if individual specific language-action cues are not significant, a particular *combination* of language-action cues may be significant in its efficacy to predict deception.

5.3.1. Model 1 at 0.5 cutoff

The initial logistic regression on this model was run with a (default) cut value of 0.5. Model 1 yields an overall accuracy of 75% in correctly classifying “0's” (i.e., truth-tellers) and “1's” (i.e., deceivers). This is a good indication that the model is appropriate for classifying and predicting truth-tellers’ statements vs. deceivers' statements. We also note that because the focus of our study is on identifying deceivers (i.e., classification as a “1”), the accuracy of the model specifically in categorizing “1's” is equally important as overall accuracy. In this instance, Table 8 shows the model to be 75% accurate in correctly classifying both “1's” as deception and as “0's” as truthful statements, as well as having 75% overall accuracy. Fig. 4 summarizes the logistic regression classification output, at the default (0.5) cutoff level.

5.3.2. Model 2 at 0.6 cutoff

The second iteration of the logistic regression analysis (depicted in Table 9) was run with a cut value of 0.6. This model yielded an accuracy rate in classifying “1's” (i.e., deceivers) of only 65%, although the model classifies “0's” (i.e., truth-tellers) with 85% accuracy, and the overall model had an accuracy of 75% (which was not different from the results where the cut-off value was set to 0.5). Fig. 5 summarizes the logistic regression classification output at the 0.6 cutoff level. Fig. 5 shows the observed groups and predicted probabilities output for Model 2. It illustrates more “0's” (i.e., truth-tellers) were correctly classified in this model than in Model 1, while fewer “1's” (i.e., deceivers) were correctly classified in this model.

5.3.3. Model 3 at 0.4 cutoff

Finally, we ran a third model, at a cut value of 0.4 (depicted in Table 10). Model 3 yields an accuracy rate for classification of “1's” as being 82.5%, with a slightly lower overall accuracy of 74% when classifying “0's” as compared to the 75% accuracy of Table 10. Table 10 summarizes the logistic regression classification output at the 0.4 cutoff level.

Fig. 6 shows the observed groups and predicted probabilities output from the logistic regression results for Model 3. It illustrates how Model 3 parsed “0's” (i.e., truth-tellers) and “1's” (i.e., deceivers) at cutoff 0.4. As can be seen, more “1's” were correctly categorized in this model than in Models 1 and 2, but fewer “0's” were correctly classified in this model than in Model 1.

5.4. Mixed model ANOVA

We further ran a mixed model ANOVA with language-action cues extracted as repeated measures (i.e., independent variables) for each subject. Both within-subjects and between-subjects factor tests were performed. Based on the Wilks’ lambda distribution, the observed power for the multivariate tests is 0.985, $F(22) = 2.313$, $p = 0.006$ (Table 11). The partial eta squared value is $\eta^2 = 0.472$ (47.2%) of the dependent variable explained by the deceiver, which suggests a medium effect size (Table 11).

The significance of the Mauchly's test of sphericity is $p < 0.000$ (Table 12); this means that the sphericity was violated. A repeated measures ANOVA with Greenhouse-Geisser correction thus determined that the mean for deceivers differed statistically significantly between factor points ($F(1.32, 102.73) = 11.56$, $p < 0.001$), and the effect size partial eta square ($\eta^2 = 0.129$) is medium (Table 13). We also measure the tests of within-subject contrasts, and linear results for the factors were found to have statistically significant results, $p < 0.01$ (Table 14).

Table 15 illustrates the tests of between-subjects effects. The observed power for the univariate tests is 0.939 with statistically significant $F = 13.47$, $p < 0.001$. The effect size partial eta square value is $\eta^2 = 0.147$ (14.7%) of the dependence variables explained by the deceiver is medium.

Post hoc tests using Bonferroni correction revealed that the differences between the truth-tellers (i.e., saints) and deceivers (i.e., sinners) were statistically significant ($p < 0.001$). We found a statistically significant difference ($p < 0.001$) between each pair of participants (Table 16).

6. Summary and discussion

One of our objectives in undertaking this study was to demonstrate the efficacy that language-action cues can collectively represent in the context of information exchange, which can be indicative to revealing a deceiver. We first performed linear regressions to examine the efficacy of the dataset. Three models were generated, and each model consists of different combinations of language-action cues, thereby demonstrating the efficacy of contextual illustration. Language-action cues are able to illustrate the dynamics and interactivity of communication, and further identify statistical significance in cases of deceptive communication. The context did not change materially when some cues (as variables) were removed in Model 2 and Model 3. Accordingly, we opted to retain the entire dataset with the variables from Model 1, which is the most

inclusive and hence provides the greatest amount of context for the communications. We thus assert that the language-action cues collectively represent context for predicting computer-mediated deception more effectively than any one individual cue alone.

An additional objective was to identify the most accurate set of cues for identifying a deceiver in spontaneous CMC—in relation to a binary outcome: deceivers ("1's") or truth-tellers ("0's"). The chi-square difference tests showed statistical significance ($p = 0.003$) across three logistic regression models. As our objective is to identify models that can correctly classify deceivers from truth-tellers, we compared results from different cutoff points, and suggest that Model 3 (with cut value at 0.4) provides the optimal combination for overall accuracy, and specific accuracy with respect to identifying deceivers. 74% of the responses were correctly classified in this model, which is virtually the same as the other two models (each had a 75% classification success rate). However, Model 3 correctly classified 82.5% of the deceivers, while Model 1 and Model 2 correctly classified only 75% and 65% of the deceivers respectively.

Our final objective was to consider language-action cues as repeated measures for a mixed-design analysis of variance model, and test whether pairwise interaction exists between each pair of the deceiver and truth-teller. A post hoc test using Bonferroni correction revealed that the differences between groups (saints vs. sinners) were statistically significant ($p < 0.001$). The difference for repeated measures between each pair of truth-teller (i.e., saint) and deceiver (i.e., sinner) in pairwise comparison was found to be statistically significant ($p < 0.001$). Comparatively speaking, deceivers were found to be less expressive, but used more decorative words per message exchanged. Deceivers showed more negative emotions and appeared more anxious when communicating with truth-tellers. Deceivers took less time to respond, but interestingly, they used more words of insight (e.g., think, know) and gave assurance by using more words of certainty (e.g., always, never) when interacting with truth-tellers. Truth-tellers, on the other hand, used more words of speculation (e.g., perhaps, guess) and took a longer time to respond to inquiries. Truth-tellers also tend to provide more reasoning to their ideas by using words of causation (e.g., because) and expressed more reflective thinking by using words of discrepancy (e.g., should, could).

7. Limitations and future work

Our research demonstrates the effectiveness of adopting Google + Hangout as a communication platform for this experiment; however, technical limitations were also experienced. In particular, players sometimes experienced technical problems in logging into the Google + pseudo accounts created for the game, and when launching the game interface. These difficulties not only confused and distracted players, but also reduced the overall amount of time spent in the game itself, thus reducing our ability to collect data. Our future work will include reconstructing the game on a more user-friendly stand-alone platform. We also plan to design and develop an automated participant assignment (i.e., pairing) system within the new platform.

Another technical limitation involves timing and the switching of roles. We noticed in several instances that the detector had not yet submitted his/her guess as to the speaker's inner role before the scenario changed. We plan to address this issue by redesigning the game so that the next new scenario will not start until the detector has actually submitted his/her guess on the speaker's role. Future work may also include incorporating additional survey questions to solicit cognitive responses from the detector, asking about the type of cues or behaviors that the speaker led her/him to make a final evaluation.

We believe that research can be carried out with the benefits of providing both learning and a fun experience to participants. We thus plan to increase the strength of competitive aspect of the game by systematically reporting participants' guesses (i.e., correct or incorrect answers) during the game. Providing more feedback to the participants may help them make better decisions. At the same time, we can also observe how research participants make both deception decisions as well as detection decisions. The game design will be enhanced to deliver a more enjoyable, competitive, and educational experience for participants. Future designs will incorporate features such as systematically reporting on participants' guesses (i.e., correct or incorrect answers) and generating respective "scores."

Another limitation lies in how data is processed. Currently, the LIWC tool assumes the use of specific words to imply certain cognitive or affective processes in the minds of the speaker. However, this assumption is hypothetical and cannot be answered by simple linguistic analysis alone. Co-text analysis may encompass important aspects in identifying the figurative, hidden interpretation of the human communication.

The final noteworthy limitation involves the sample size of the dataset, which we acknowledge is fairly small. We anticipate that in our future work, we will recruit from a broader population of potential participants in order to create a larger dataset. Nonetheless, we submit that the statistical significance of the current study is suggestive and encouraging.

8. Contributions and conclusion

Communicators' actions are embedded in their language-action cues in the iteration of message exchange, and these actions define the interactivity that shapes the dynamic context of computer-mediated communication. From a holistic perspective, the words and other nonverbal language-action cues, as well as the mode and medium used for communication, create the context of that interaction—which can be a viable measure to derive the communicative intent of interacting parties. Rather than a single word choice each speaker chooses, it is the collective pattern of language-action cues that are most salient in spotting deception within the interactivity context of spontaneous computer-mediated communication. The research provides empirical evidence that language-action cues collectively demonstrate efficacy in identifying computer-mediated deception in synchronous, spontaneous text-based CMC communication. Context is an interactional problem (Dourish, 2004), and this study identifies context not as static but as interactive information created and exchanged dynamically illustrating deceptive intent. Context as embedded in language-action cues differentiates dynamic communicative intent, and statistical significance was found between deceivers and truth-tellers within their interaction.

This study demonstrates unique perspectives in the collective language-action cues embedded in the landscape of the interactive and dynamic context. The spontaneous interaction context was operationalized in a sociotechnical research system that mimics computer-mediated deception scenarios. Deceivers' communicative intent was conceptualized and visualized by the language-action cues. As deceivers tend to adopt dynamic strategies to conceal their communicative intent, the operationalization of this study enables the identification of communicator's deceptive intent as embedded in language-action cues that can be statistically computed with high accuracy and power. Our results rigorously demonstrate that the overall combination of language-action cues—rather than specific words used to deceive—can represent interaction context, which can be statistically indicative of computer-medicated deception (Ho and Hancock, 2018).

Patterns in words and language-action cues can assist in identifying spontaneous computer-mediated deception. The study further demonstrates potency in computationally identifying spontaneous computer-mediated deception, and informs effectiveness and power in adopting computational approaches to identify computer-mediated deception (Ho, Hancock, et al., 2016; Ho, Liu, et al., 2016). Moreover, the merit of the game design specifically emphasizes not just synchronous communication, but *spontaneity* within synchronous communication (Ho et al., 2015). That is, the results of this study provide insight into patterns of spontaneous computer-mediated deception in synchronous CMC, and

may be able to inform the design of an online polygraph system—or a prototype detection system for computer-mediated deception when face-to-face interaction is not available.

Acknowledgements

The authors gratefully acknowledge the grants from National Science Foundation (#1347113 and #1347120, 09/01/13–08/31/15), Florida Center for Cybersecurity Collaborative Seed Grant (#2108-1072-00-O, 03/01/15–02/28/16), and Florida State University Council for Research and Creativity Planning Grant (#034138, 12/01/13–12/12/14). The authors acknowledge and appreciate the research efforts and contributions from Cheryl Booth, Sai Surya Shashanka Timmarajus, Kashyap Vemura, and Aravind Hariharan.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chb.2018.09.008.

References

Attardo, S. (1994). *Linguistic theories of humor*. Berlin: New York: Walter de Gruyter & Co 978-3-11-014255-6.
Buller, D. B., & Burgoon, J. K. (1994). Deception: Strategic and nonstrategic communication. *Strategic Interpersonal Communication,* 191–223.
Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6*(3), 203–242.
Burgoon, J. K., & Buller, D. B. (1994). Interpersonal deception: III. Effects of deceit on perceived communication and nonverbal behavior dynamics. *Journal of Nonverbal Behavior, 18*(2), 155–184.
Burgoon, J. K., Buller, D. B., Dillman, L., & Walther, J. B. (1995). Interpersonal deception. IV. Effects of suspicion on perceived communication and nonverbal behavior dynamics. *Human Communication Research, 22*(2), 163–196.
Burns, M. B., & Moffitt, K. C. (2014). Automated deception detection of 911 call transcripts. *Security Informatics, 3*(8), 1–9.
Daft, R. L., Lengel, R. H., & Trevino, L. K. (1987). Message equivocality, media selection, and manager performance: Implications for information systems. *MIS Quarterly, 11*(3), 355–366.
DePaulo, B. M., Ansfield, M. E., Kirkendol, S. E., & Boden, J. M. (2004). Serious lies. *Basic and Applied Social Psychology, 26*(2&3), 147–167.
DePaulo, B. M., & Kashy, D. A. (1998). Everyday lies in close and causal relationships. *Journal of Personality and Social Psychology, 74*(1), 63–79.
DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70*(5), 979–995.
DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Pyschological Bulletin, 129*, 74–112.
Dourish, P. (2004). What we talk about when we talk about context. *Journal of Personal and Ubiquitous Computing, 8*(1), 19–30.
Ekman, P., & Friesen, W. B. (1969). Nonverbal leakage and clues to deception. *Psychiatry, 32*, 88–106.
Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar? *American Phychologist, 46*(9), 913–920.
Furner, C. P., & George, J. F. (2012). Cultural determinants fo media choice for deception. *Computers in Human Behavior, 28*(4), 1427–1438.
George, J. F., Giordano, G., & Tilley, P. A. (2016). Website credibility and deceiver credibility: Expanding prominence-interpretation theory. *Computers in Human Behavior, 54*, 83–93.
George, J. F., Tilley, P. A., & Giordano, G. (2014). Sender credibility and deception detection. *Computers in Human Behavior, 35*, 1–11.
Hancock, J., Birnholtz, J., Bazarova, N., Guillory, J., Perlin, J., & Amos, B. (2009). Butler lies: Awareness, deception and design. *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 517–526). New York, NY: ACM.
Hancock, J. T., Curry, L. E., Goorha, S., & Woodworth, M. (2008). On lying and being lied to: A linguistic analysis of deception in computer-mediated communication. *Discourse Processes, 45*(1), 1–23.
Hancock, J. T., Thom-Santelli, J., & Ritchie, T. (2004). Deception and design: The impact of communication technologies on lying behavior. *Proceedings of the conference on computer human interaction* (pp. 130–136). New York, NY.
Hancock, J., Toma, C., & Ellison, N. (2007). The truth about lying in online dating profile. *Proceedings of the SIGCHI 2007 conference on human factor in computing systems* (pp. 449–452). New York: ACM.
Ho, S. M., & Hancock, J. T. (2018). Computer-mediated deception: Collective language-action cues as stigmergic signals for computational intelligence. *Proceedings of the 2018 51st Hawaii international conference on system sciences (HICSS-51)* (pp. 1671–1680). Big Island, Hawaii: IEEE Computer Society.
Ho, S. M., Hancock, J. T., Booth, C., & Liu, X. (2016a). Computer-mediated deception: Strategies revealed by language-action cues in spontaneous communication. *Journal of Management Information Systems, 33*(2), 393–420.
Ho, S. M., Hancock, J. T., Booth, C., Liu, X., Liu, M., Timmarajus, S. S., et al. (2016b). Real or Spiel? A decision tree approach for automated detection of deceptive language-action cues. *Proceedings of the 2016 49th Hawaii international conference on system sciences (HICSS-49), 3706-3715*. IEEE Computer Society.
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,* 157–159 IEEE.
Ho, S. M., Liu, X., Booth, C., & Hariharan, A. (2016c). Saint or Sinner? Language-action cues for modeling deception using support vector machines. In K. S. Xu, D. Reitter, D. Lee, & N. Osgood (Eds.). *Social, cultural, and behavioral modeling (SBP-BRiMS'16), LNCS 9708* (pp. 325–334). Springer.
Lewis, C. C., & George, J. F. (2008). Cross-cultural deception in social networking sites and face-to-face communication. *Computers in Human Behavior, 24*(6), 2945–2964.
Miller, G. R., Deturck, M. A., & Kalbfleisch, P. J. (1983). Self-monitoring, rehearsal, and deceptive communication. *Human Communication Research, 10*(1), 97–117.
Miller, G. R., & Stiff, J. B. (1993). *Deceptive communication*, *Vol. 14*, Newbury Park, CA: Sage142.
Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richard, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin, 29*(5), 665–675.
Ott, M., Cardie, C., & Hancock, J. T. (2012). *Estimating the prevalence of deception in online review communities.* ACM201–210.
Ott, M., Cardie, C., & Hancock, J. T. (2013). *Negative deceptive opinion spam*, *Vols. 497–501*. Association for Computational Linguistics.
Ott, M., Choi, Y., Cardie, C., & Hancock, J. T. (2011). Finding deceptive online spam by any stretch of the imagination. *Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies* (pp. 309–319). Stroudsburg, PA: Association for Computational Linguistics.
Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology, 77*(6), 1296–1312.
Smith, M. E., Hancock, J. T., Reynolds, L., & Birnholtz, J. (2014). Everyday deception or a few prolific liars? The prevalence of lies in text messaging. *Computers in Human Behavior, 41*, 220–227.
Stiff, J. B. (1996). Theoretical approaches to the study of deceptive communication: Comments on interpersonal deception theory. *Communication Theory, 6*(3), 289–296.
Toma, C. L., & Hancock, J. T. (2010). *Reading between the lines: Linguistic cues to deeption in online dating profiles.* ACM5–8.
Toma, C. L., & Hancock, J. T. (2012). What lies beneath: The linguistic traces of deception in online dating profiles. *Journal of Communication, 62*(1), 78–97.
Toma, C. L., Hancock, J. T., & Ellison, N. (2009). Separating fact from fiction: An examination of deceptive self-presentation in online dating profiles. *Personality and Social Psychology Bulletin, 34*(8), 1023–1036.
Trevino, L. K., Lengel, R. H., & Daft, R. L. (1987). Media symbolism, media richness and media choice in organizations. *Communication Research, 14*(5), 553–574.
Vrij, A. (2015). Verbal lie detection tools: Statement validity analysis, reality monitoring and scientific content analysis. In P. A. Granhag, A. Vrij, & B. e. Verschuere (Eds.). *Detecting deception: Current challenges and cognitive approaches* (pp. 3–35). Malden, MA: Wiley-Blackwell.
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, 63*(1), 208–216.
Whitty, M. T., & Carville, S. E. (2008). Would I lie to you? Self-serving lies and other-oriented lies told across different media. *Computers in Human Behavior, 24*(3), 1021–1031.
Zhou, L., Burgoon, J. K., & Twitchell, D. P. (2003). A longitudinal analysis of language behavior of deception in email. *Proceedings of the 1st NSF/NIJ conference on Intelligence and security informatics: Vols. 102–110*Springer-Verlag Berlin Heidelberg.
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*(4), 27.
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. *Proceedings of the 36th Hawaii international conference on systems sciences*. IEEE.
Zhou, L., & Zhang, D. (2004). Can online behavior unveil a deceiver? *Proceedings of the 37th Hawaii international conference on system sciences*. IEEE Press.
Zhou, L., & Zhang, D. (2008). Following linguistic footprints: Automatic deception detection in online communication. *Communications of the ACM, 51*(9), 119–122.

Shuyuan Mary Ho is an associate professor at School of Information, at Florida State University. Her research focuses on computer-mediated deception, trusted human computer interaction, cloud forensics, online games for research, and sociotechnical behavioral experiments. Her work appears in over 40 journal articles and conference proceedings and has been funded by U.S. National Science Foundation.

Jeffrey T. Hancock is a professor of in the Department of Communication at Stanford University. He works on understanding psychological and interpersonal processes in social media by using computational linguistics and behavioral experiments to examine deception and trust, emotional dynamics, intimacy and relationships, and social support. His research has been published in over eighty journal articles and conference proceedings, and has been supported by funding from the U.S. National Science Foundation and the U.S. Department of Defense. His work has also been featured in the popular press, including the *New York Times*, *CNN*, *NPR*, *CBS* and the *BBC*.