

8-2019

The effect of conversational agent skill on user behavior during deception

Ryan M. Schuetzler

University of Nebraska at Omaha, ryan.schuetzler@byu.edu

G. Mark Grimes

University of Houston, gmggrimes@bauer.uh.edu

Justin Scott Giboney

Brigham Young University, justin_giboney@byu.edu

Follow this and additional works at: <https://digitalcommons.unomaha.edu/isqafacpub>

 Part of the [Databases and Information Systems Commons](#)

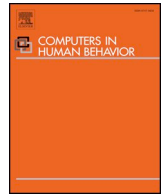
Please take our feedback survey at: https://unomaha.az1.qualtrics.com/jfe/form/SV_8cchtFmpDyGfBLE

Recommended Citation

Schuetzler, Ryan M.; Grimes, G. Mark; and Giboney, Justin Scott, "The effect of conversational agent skill on user behavior during deception" (2019). *Information Systems and Quantitative Analysis Faculty Publications*. 74.

<https://digitalcommons.unomaha.edu/isqafacpub/74>

This Article is brought to you for free and open access by the Department of Information Systems and Quantitative Analysis at DigitalCommons@UNO. It has been accepted for inclusion in Information Systems and Quantitative Analysis Faculty Publications by an authorized administrator of DigitalCommons@UNO. For more information, please contact unodigitalcommons@unomaha.edu.



Full length article

The effect of conversational agent skill on user behavior during deception

Ryan M. Schuetzler^{a,*}, G. Mark Grimes^b, Justin Scott Giboney^c^a University of Nebraska at Omaha, United States^b University of Houston, United States^c Brigham Young University, United States

ARTICLE INFO

Keywords:

Conversational agents

Relevance

Deception

Conversational skill

Chatbots

ABSTRACT

Conversational agents (CAs) are an integral component of many personal and business interactions. Many recent advancements in CA technology have attempted to make these interactions more natural and human-like. However, it is currently unclear how human-like traits in a CA impact the way users respond to questions from the CA. In some applications where CAs may be used, detecting deception is important. Design elements that make CA interactions more human-like may induce undesired strategic behaviors from human deceivers to mask their deception. To better understand this interaction, this research investigates the effect of conversational skill—that is, the ability of the CA to mimic human conversation—from CAs on behavioral indicators of deception. Our results show that cues of deception vary depending on CA conversational skill, and that increased conversational skill leads to users engaging in strategic behaviors that are detrimental to deception detection. This finding suggests that for applications in which it is desirable to detect when individuals are lying, the pursuit of more human-like interactions may be counter-productive.

1. Introduction

The goal of many modern information systems is to replace expensive human-human interactions with automated human-computer interactions to provide more cost effective, efficient, customizable, and reliable service. Recent advancements in natural language processing (NLP) have led to a surge of automated conversational agents (CAs)—computer systems designed to communicate in natural language with humans, as opposed to using predefined computer commands. CAs have recently been used in domains such as healthcare (Bickmore & Picard, 2005), education (Fridin & Belokopytov, 2014; Rodrigo et al., 2012), and as “digital personal assistants” (e.g., Siri and Alexa). These systems can provide improved convenience and service at a fraction of the cost of a human agent. Because of these benefits, companies are rapidly investing in CA technologies to supplement their existing customer interaction platforms (Lin & Chang, 2011).

One common goal in the design of such applications is to increase the social presence of the CA by manipulating characteristics of the CA to make it seem more human-like (Heerink, Kröse, Evers, & Wielinga, 2010). For example, characteristics of a CA such as gender, demeanor, dress (Nunamaker, Derrick, Elkins, Burgoon, & Patton, 2011), similarity (Pickard, Burns, & Moffitt, 2013), and perceived agency (Appel, Von Der Pütten, Krämer, & Gratch, 2012) affect how humans perceive and

interact with the CA. Social presence also improves perceptions of a system as users develop more connection with the system (Shin, 2013). This has been demonstrated in human-robot interaction, where the social capabilities of a robot strongly influence acceptance (Shin & Choo, 2011). Technologies that facilitate NLP enable computers to engage in more human-like conversations by analyzing the speech provided by the user and responding in a way similar to how another human would respond. While communicating with a computer in the same way one communicates with another human is often considered desirable and is frequently portrayed in popular culture as the future of computing (i.e., HAL from Space Odyssey, KITT from Knight Rider, or JARVIS from Iron Man), our pursuit of this type of natural communication with computers has perhaps surpassed our understanding of the impact of increasingly human-like communication on the people using the technology. To ensure that improved CA capabilities do not conflict with operational goals of CAs, we must gain a better understanding of the effect of increasing the capabilities of a CA on human behaviors. The goal of this research is to advance this understanding.

Eliciting complete and truthful information is important in many emerging applications for CAs, such as conducting investigatory interviews (Pollina & Barretta, 2014; Proudfoot, Boyle, & Schuetzler, 2016), detecting deception (Cunningham, 2017; Higginbotham, 2013; Warth, 2017), and discussing sensitive health matters (Schuetzler, Giboney,

* Corresponding author.

E-mail address: rschuetzler@unomaha.edu (R.M. Schuetzler).

Grimes, & Nunamaker, 2018; Steiner, 2012). Obtaining truthful information is especially important in contexts where people may have incentive to deceive, such as a job interview (HarQen, 2017) or insurance claim (Hibbeln, Jenkins, Schneider, Valacich, & Weinmann, 2014). Additionally, determining the veracity of information in these contexts is important to the efficacy of the decision-making process and can facilitate a system's ability to assess and appropriately respond to information provided to it.

Given our current understanding of CAs, these operating environments, and the process of human-human deception, it is unclear if increasing the human-likeness of a CA—and therefore its social presence—is universally desirable. Prior work has shown that socially anxious people disclose more information to a CA that is perceived to be a computer rather than one that is perceived to be human (Kang & Gratch, 2010). Similarly, Interpersonal Deception Theory (IDT, Buller & Burgoon, 1996) proposition 8 suggests that deceivers exhibit more *strategic deception behaviors*—actions that reflect “large-scale plans and intentions (as opposed to specific behavioral routines or tactics)” (Buller & Burgoon, 1996, p. 207) to make a deceptive message appear credible—to avoid detection as their relational and behavioral familiarity with their target increases. When an interlocutor is more socially present, this facilitates familiarity, thereby leading to an increase in the deceiver's strategic behaviors to avoid their deception being detected. Strategic behaviors include information management (modifying or manipulating message content), image management (attempts to maximize credibility through a competent and trustworthy demeanor), and behavior management (actions designed to suppress behaviors that might expose one's deception) (Buller & Burgoon, 1996). Some strategic behaviors, such as rehearsed lies, are planned and deliberately acted on, while others, such as controlling the length or speed of a response, are not planned or deliberate, but rather are natural responses used for behavior management (i.e., attempts to maintain a natural and unassuming flow in the communication). That is to say, not all strategic behaviors—such as maintaining a natural rate of speech—are conscious and deliberate actions, but rather they may be constituent behaviors used to maintain the overall goal of behavior management.

These applications and theories reflect potentially conflicting design principles that could be inferred from the desire to make CAs more human-like, to encourage truthful responses, and to detect deceptive responses. Many CA designs call for the CA to be more human-like to facilitate engagement. However, in contexts where a goal of the CA is to elicit truthful responses and detect when responses may be deceptive, appearing more human-like may cause individuals to engage in more strategic behaviors to avoid detection, thus making the CA's detection of deception more difficult. To establish the most effective type of interaction for eliciting and assessing the veracity of information, we must understand how a CA's characteristics influence humans' behavioral indicators of deception. In this research, we extend the traditional human-to-human interpersonal communication described by IDT to consider deceptive communication from humans interacting with virtual agents that have varying levels of conversational skill—that is, how capable of following human-like communication patterns the CA is. We aim to answer the following research question:

What impact does the level of conversational skill of a CA have on peoples' behavior during deceptive and truthful communication?

To answer this question, we develop a theoretical research model that describes the influence of improving a CA's conversational skill on two behavioral indicators of deception: response latency and response hesitations (DePaulo et al., 2003). In the following pages we provide a theoretical foundation for this work and describe a laboratory experiment used to test our hypotheses. Finally, we discuss the results of the study and the theoretical and practical implications.

2. Theoretical development

2.1. Human-likeness of conversational agents

Prior research has demonstrated that computers are often perceived and interacted with as social actors (Nass, Steuer, & Tauber, 1994). This view of computers as social actors is in part facilitated by the user's perception of the computer as a partner in the communication and leads to a “sense of human contact embodied in a medium” (Gefen & Straub, 1997, p. 390). Information systems that are perceived as having a more human-like touch facilitate trust (Gefen & Straub, 2004), enjoyment and perceived usefulness (Hassanein & Head, 2007), and self-efficacy beliefs (Baylor, 2009). Variations in characteristics such as gender and demeanor have also been shown to influence the connection users feel with the system (Nunamaker et al., 2011). Features such as these help to establish a sense of connection with the other participant, human or computer, through the medium (Schultze, 2010), and create perceptions of humanity.

Humanity for a CA is very broadly defined as the ability of the agent to act in human-like ways (Radziwill & Benton, 2017). By the most strict definitions, this would require the CA to pass the Turing test of artificial intelligence (Turing, 1950). For general applications, however, individual elements of a CA can reflect humanity, even if the Turing Test is not passed (Araujo, 2018; Hayes & Ford, 1995; Morrissey & Kirakowski, 2013). The Computers are Social Actors (CASA) paradigm posits that we often attribute humanity to computers, even if they demonstrate few or no human-like qualities (Nass et al., 1994; Nass & Moon, 2000). For example, people frequently personify their personal electronics by giving them names, attributing blame or emotions to them, and applying norms such as politeness and reciprocity (Nass, Moon, & Carney, 1999), despite the devices clearly falling short of the strict definition of humanity. This reflects the fluid nature of our relationship with computers as social actors, and the mindless way in which users anthropomorphize systems.

This anthropomorphism is often the result of a computer system having social abilities (Appel et al., 2012; Waytz, Heafner, & Epley, 2014) such as exhibiting *conversational skill* (Morrissey & Kirakowski, 2013). Participants in human-to-human conversations exhibit conversational skill by following conversational norms regarding the content, timing, and flow of the conversation. One of these norms is the maxim of relation (Grice, 1975), which describes the expectation that conversation partners will respond to each other during the conversation with responses that are tailored to the conversation. CAs can simulate human conversational skill by giving tailored responses, in which the CA adjusts its own dialogue based on what people say. Through these tailored responses, users perceive that the CA understands what they are saying and can adapt its conversation accordingly. This is referred to as the *conversational relevance* of responses—that is, that responses are appropriate and contingent on the messages they have received (Sundar, Bellur, Oh, Jia, & Kim, 2016).

Violations of the conversational relevance norm, such as responding with generic questions or comments, or frequently changing topics in order to mask a lack of conversational awareness, violate the maxim of relation, leading to a decrease in the perception of human-likeness (Kirakowski, O'Donnell, & Yiu, 2007). Maintaining conversational relevance, especially in computer-mediated communication (CMC), involves incorporating context into messages to maintain continuity from one message to the next (Herring, 2013). For a CA to be considered conversationally skilled, it must be capable of processing information in users' messages and carrying that information forward into its own messages (Morrissey & Kirakowski, 2013). It is the act of tailoring responses by incorporating conversational context that signals to users that the CA understands.

Other conversational norms relate to temporal features of the response such as response latency and hesitations (Miller, 1968). *Response latency* is the delay between when a user receives a message and when

they send a reply. In oral communication, response latency is a brief period of silence typically measured as the time between when one speaker is finished speaking, and the next speaker begins to speak (Bassili, 2000). This brief silence between speaking turns indicates that a speaker is finished and sets the expectation that the other conversational partner will take a turn. This period of silence is generally short as people attempt to minimize silence between conversational turns—on average roughly 250 ms for spoken communication (Stivers et al., 2009). In text-based communication, response latencies are generally longer as they must compensate for the length of time required to read and understand the message (Temple & Geisinger, 1990). Response latency can also be negative—for example, when a participant begins responding before their partner has finished sending their message—resulting in a phenomenon known as cross-talk (Taboada, 2006).

Hesitation time is another temporal nonverbal communication behavior that provides additional insight into the meaning of a message. Hesitations in oral communication are periods of silence within a communicator's turn that vary in both their duration and operationalization depending on the type of conversation or speech (Duncan, 1969). Increases in hesitation time are frequently associated with increases in cognitive effort (Berthold & Jameson, 1999). Hesitations in spoken dialogue include both filled pauses (e.g., “umm” or “uhh”) and unfilled pauses (e.g., silence) (Goldman-Eisler, 1961). While words typically associated with filled pauses may be found in typed text, they serve a different purpose in written as compared to spoken communication (Carey, 1980; Kalman & Gergle, 2014). Further, many of the paralinguistic features found in oral communication are not available in computer-mediated communication, thus typed communication has only unfilled pauses (Zhou, 2005). Both response latency and hesitations are conceptually similar to their oral counterparts (Schuetzler, Grimes, Elkins, Burgoon, & Valacich, 2016), and may signal that the sender is distracted, not engaged in the conversation or is uncertain of the message that is being sent (Burgoon et al., 2015).

Bringing together the essence of CASA and conversational norms, it should hold that when an individual's conversational partner demonstrates conversational skill—for example, by giving responses that are tailored to be relevant to the conversation—they will reciprocate by treating their partner more like a person and less like an object (Dennis & Kinney, 1998; Williams, 1977) and likewise adhere to conversational norms—for example, by being more temporally responsive. To this end, we hypothesize:

H1. People interacting with a more conversationally skilled CA will exhibit a) lower response latency and b) lower average hesitation time than people interacting with a less conversationally skilled CA will.

2.2. Deception

As CAs have become more integrated into interactions in which people might behave deceptively, the ability to assess the veracity of responses has become increasingly important. Thus, one area of interest is identifying changes in human behaviors that are the result of a human attempting to deceive a CA (Nunamaker et al., 2011; Zhou, Burgoon, Zhang, & Nunamaker, 2004). Deception may occur in interactions where one might lie to prevent embarrassment (e.g., doctor's office), avoid getting in trouble (e.g., interaction with a law enforcement officer), or to gain undesired benefits (e.g., interaction with a company or customer service representative).

Interpersonal Deception Theory (IDT, Buller & Burgoon, 1996) postulates how deception changes normal communication between at least two parties. While IDT was originally formulated with human-human interactions in mind, prior research has successfully applied IDT in CMC settings (Derrick, Meservy, Jenkins, Burgoon, & Nunamaker, 2013; Zhou, Burgoon, Twitchell, Qin, & Nunamaker Jr, 2004). Further, as CA technologies continue to improve and computers are able to communicate in more human-like ways, the relevance of the distinction

between humans and computers in dyadic communication is reduced. IDT makes 18 propositions related to deception, focusing on the deeply interconnected nature of interpersonal communication, and the importance of each participant in a deceptive communication. IDT posits that communication participants are actively encoding and decoding messages, that the communication is dynamic, and that the communicators have multiple goals, topics, and methods of transmission (Buller & Burgoon, 1996).

Communication, and especially deceptive communication, involves both strategic and nonstrategic behaviors (Buller & Burgoon, 1996; Kellermann, 1992; McCornack, 1992), and is governed by cognitive and behavioral forces. IDT describes deceptive communication as additionally having a central deceptive message, ancillary messages designed to enhance the perceived validity of the message, and inadvertent messages that imply the truth (i.e., messages that “leak” from the communicator) (Buller & Burgoon, 1996; Ekman & Friesen, 1969). Successful deception is dependent on producing appropriate strategic cues of truthfulness and managing nonstrategic cues that result from the emotional and mental effort associated with managing deceptive communication (Zuckerman, DePaulo, & Rosenthal, 1981).

In order to give the appearance of truthfulness, deceivers must put forth greater effort to appear credible (IDT proposition 3a). However, the increased effort associated with appearing credible produces other observable artifacts, including “performance decrements” (IDT proposition 3b). Such performance decrements include exhibiting longer response latency or hesitation time due to the effort required to produce and maintain a false story. Accordingly, extant deception research suggests that deceivers tend to exhibit increases in hesitations (Buller, Comstock, Aune, & Strzyzewski, 1989; Buller, Strzyzewski, & Comstock, 1991; Levine & McCornack, 1996) and response latency (Ho, Hancock, Booth, & Liu, 2016; Sporer & Schwandt, 2006).

Since most deception research to date has focused on oral communication, there remains a scarcity of research testing latency and hesitations for typed communication (e.g., Derrick et al., 2013). This is one of the first projects looking at typing behavior during deception with conversational agents. Based on the nonstrategic leakage propositions from IDT, we hypothesize:

H2. People engaging in deception will exhibit longer a) response latency and b) hesitation times when interacting with a CA, compared to truthful interactions.

2.3. The interactive nature of CA conversational skill and deception

As described in the prior two subsections, it is expected that a CA that demonstrates more human-like behavior by exhibiting more conversational skill will elicit shorter response latency and shorter hesitations from its human interlocutor than would a CA with low conversational skill, as conversational norms are reciprocated. However, according to deception theories, deceivers may display longer response latency and hesitations than truth-tellers as they experience the additional effort associated with deception (Dunbar et al., 2014; Sporer & Schwandt, 2006). There is obviously tension between these outcomes, as neither conversational norms nor deception exist in a vacuum. While deception will generally lead to increased response latency and hesitation, individuals who wish to mask their deception may engage in strategic behaviors, such as responding more quickly or reducing hesitations, in order to avoid detection. However, engaging in such strategic behaviors only makes sense if the human interlocutor perceives their communication partner as having the ability to evaluate their messages.

IDT proposition 4 suggests that increasing interactivity results in greater strategic behavior and reduced nonstrategic behavior as people work to manage their behavior to avoid detection. Applying this proposition to human-CA conversation would suggest that deceivers communicating with a CA that exhibits higher conversational skill by

providing tailored responses (i.e., more interactivity) would be more likely to strategically manage expressed information and control their nonstrategic behavior (Buller & Burgoon, 1996; Burgoon & Buller, 1994). To avoid detection, deceivers will be more strategic in their communication by attempting to control the expected increase in response latency and hesitations when interacting with a more skilled CA (Sporer & Schwandt, 2006). This phenomenon has been demonstrated in deceptive interpersonal communication (Dunbar et al., 2014). Thus, while we expect deceivers overall to show increases in response latency and hesitations (H2), we expect this relationship will be moderated by the presence of a more skilled (i.e., humanlike) CA as deceivers engage in strategic behaviors to avoid detection. As such, we propose the following hypothesis:

H3. The relationship between deception and a) response latency and b) hesitations will be moderated by CA conversational skill.

In summary, we hypothesize that a CA that exhibits high conversational skill will lead people to reciprocate conversational norms by likewise being more responsive, operationalized here as reduced response latency (H1a) and average hesitation time (H1b), as compared to interactions with a CA that exhibits low conversational skill. We expect deception to increase response latency (H2a) and average hesitation time (H2b), however, we expect these relationships to be moderated by CA conversational skill (H3a/b) as people strategically attempt to control their response in the presence of a CA with higher conversational skill. Our model is presented in Fig. 1.

3. Method

3.1. Developing the chatbots

Conversational agents are a broad category of interaction tools. One common type of CA is a chatbot. A chatbot is a text-based agent designed to carry on a conversation. A collection of custom chatbots were developed for this study using the ChatScript language (Wilcox, 2017). This approach afforded the research team several advantages for creating a robust environment for testing the hypotheses. First, the creation of custom chatbots allowed us to control the interaction between the participants and the CA and to capture additional features of the communication that are not readily available in existing chat software. JavaScript code embedded in the chat application captured precise timings of when messages were presented to the participants, when participants started responding, and time taken between each typed character. The participants' responses were split into two data streams—the content of the message, which was processed by natural language processing (NLP) algorithms used by ChatScript to formulate responses—and the keystroke timing, which was sent to a separate application for analysis.

Second, the use of ChatScript as a development platform facilitated the creation and organization of chat topics. The topics can be triggered by keywords from the user or selected by the chatbot if no matching topics are available. ChatScript also includes features such as automatic spelling correction and expansion of abbreviations (e.g., “IDK” to “I do not know”) and contractions (e.g., expanding “how’s” into “how is”).

This is important as people frequently make typographical errors or use shorthand that impedes NLP techniques. This capability eliminates the need for chatbot developers to write patterns that match many potential ways of communicating the same message.

Finally, ChatScript groups keywords and responses into *concepts*. Concepts are used for organizing words and phrases into groups with similar meaning so that the chatbot can respond appropriately to an idea that might be expressed in a variety of ways. For example, people may greet the bot in many ways, such as “hello,” “howdy,” or “hey.” No matter how people choose to say hello, the effective meaning is the same. ChatScript groups these greetings into a single concept (~emo-hello). Developers can use this concept (rather than the exact words) to match all greetings, rather than anticipating every possible way a user might begin a conversation.

3.2. Experiment design

We conducted a study at a large public university in the United States. Participants were recruited from an entry-level MIS course and instructed to report to our research lab at a specified time. Upon arrival at the lab, the participants completed an online survey capturing demographic information and computer use behavior. Upon completion of the survey, participants' web browsers were automatically redirected to the experiment where they were randomly assigned to a high conversational skill chatbot which gave tailored responses or a low conversational skill chatbot which gave generic responses. Before interacting with the chatbot, the participants were given the following instructions:

In this experiment you will be placed in an online chatroom and asked questions about a series of images. Your chat partner cannot see these images, and you will be instructed to lie about the content of some of the images – it is important to follow these instructions.

The instructions were worded in such a way that participants were not told whether their chat partner was human or computer. All chat interactions were in fact with a chatbot.

The experiment was designed to mimic a classic deception experiment in which participants are shown a series of images then asked to describe some images deceptively and others truthfully (Ekman & Friesen, 1974). Prior to beginning the interaction, participants were shown example screenshots of the interface, with instructions for how to proceed. Screenshots showed what the interface would look like when they were to lie and when they were to tell the truth. Alternating between lying and telling the truth is common for deception studies to be able to compare an individual's truthful behavior to their deceptive behavior (Elkins & Derrick, 2013; Giboney, Brown, Lowry, & Nunamaker, 2015; Nunamaker et al., 2011). For images with instructions to deceive, a bold message stated “For this image, please answer the questions as if the picture is of [something similar in topic but of opposite valence]”. For example, if the picture were of two dogs fighting, the instructions might ask the participant to describe the image as two dogs sleeping on a bed. Neither of the images used in the example screenshots were used in the experiment, and both example images were of neutral valence.

After the two examples, the participants clicked a link that took them to the chat interface. Participants were shown, and had a chat conversation about, twelve images—four each of positive, negative, and neutral valence—from the International Affective Picture System (Lang, Bradley, & Cuthbert, 2008). Participants were instructed to lie about the content of 6 of the 12 images shown—two from each valence category. For the other six images, participants were asked to tell the truth. Images were presented in two different orders to different groups of participants, with each group being asked to lie about different images. This was done to account for potential ordering effects or differences in the images.

Prior to beginning the chat, participants were given as much time as

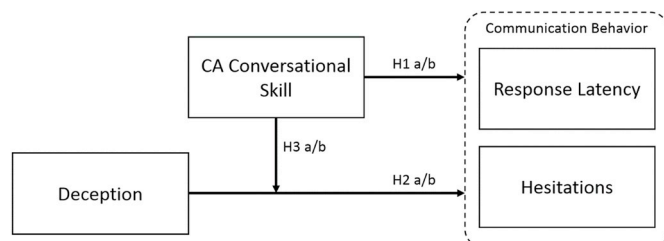


Fig. 1. Research model.

needed to review the image. Participants were instructed to click a button labeled “I’m ready” to indicate they were ready to begin the chat. Shortly after clicking the button, the first question from the chatbot appeared and participants began their conversation. In each step of the conversation, participants had as much time as they needed to construct their description (truthful or deceptive) of the image. After the set of questions for an image was completed, the chat window was temporarily disabled and a system message instructed them to click “Next” to move to the next image, then “I’m ready” after they had viewed the image and were ready to chat about it. An example of the chat interface is shown in Fig. 2. This research design resulted in one binary, between-subjects condition (high or low conversational skill CA) and one binary, within-subjects condition (answer truthfully or answer deceptively).

For each image shown, the CA followed a conversation stream similar to that presented in Table 1 in which the participant was asked two base questions and two follow up questions. In the high conversational skill condition, the CA was configured to look for key words in the participant’s message to formulate a response that was tailored to the message provided by the participant. For example, as illustrated in Fig. 3a, since the user said the key words “puppies” and “sitting” in response to the first inquiry, the CA responded with the question “What are the puppies sitting on?” As illustrated in Fig. 3b, participants in the low skill CA condition received a generic follow-up question each time. This is not to say that the responses were not appropriate, however. For example, despite “Why does it make you feel that way” being an appropriate response to “Very happy”, it does not signal any understanding of the message sent by the user. The same response from the CA would be equally appropriate if the user had said “Sad”, “Angry”, or “Elated”—thus we consider these responses to be “generic”. All participants received the same number of messages from the CA regardless of condition, keeping the conversation length roughly consistent between the conditions.

Response latency was measured in milliseconds (ms) from the time users received a question from the CA and when they began typing a reply. If users began typing before the question from the CA was received, this resulted in a negative response latency. Hesitations were measured as pauses longer than 500 ms after the first character was typed. Hesitations shorter than 500 ms were removed, as prior work has suggested hesitations of under 500 ms are indicative of normal typing latencies rather than hesitations for thinking (Joyce & Gupta, 1990).

4. Analysis

One-hundred twelve students were recruited from an entry-level MIS course. They participated in exchange for course credit. Data from five participants was excluded due to technical issues with the experiment site, and four participants were removed for reporting that they failed to follow instructions (they reported that they did not lie on the images they were instructed to lie about). We manually reviewed responses to verify that deception occurred. In total, 103 participants’ data (51 male, 52 female) were used in the final analysis. Of those participants, 88% were native English speakers with an average age of 19.5 years ($SD = 3.02$) ranging from 18 to 45 years.

We analyzed the results with a repeated measures nested generalized linear model comparison. We included deception and question as random effects within participants. The deception measure was a dummy code for each participant \times image indicating whether the participant was told to lie or tell the truth about the image. We first tested response latency. We removed response latencies from the first image of the experiment because it was significantly higher than the rest, presumably due to the novelty of the interface (See specific numbers in Appendix A).

The first variables in the model are controls for ordering effects and question effects. Order was not significant ($X^2(1) = 0.43, p = .51$), but there was a significant effect of question, $X^2(3) = 60.24, p < .001$. Next, we tested for the main effects of CA conversational skill and deception. We found a significant main effect of CA conversational skill, $X^2(1) = 36.32, p < .001$, but no additional main effect of deception, $X^2(1) = 0.86, p = .35$. Finally, we include the interaction between conversational skill and deception in the model, which is statistically significant, $X^2(1) = 4.60, p = .032$.

In the final model, the statistically significant main effect of conversational skill on response latency remains ($b = 1380, t = 6.93, p < .0001$, b is unstandardized in nested models), as does the interaction effect ($b = -243, t = -2.14, p = .03$). However, there is no significant main effect of deception ($b = 62, t = 0.80, p = .42$). The means plot of response latency differences is shown in Fig. 4. The skilled CA condition increased response latency by an average of 1.4 seconds ($SE = 0.2$). Response latency decreased during deception for those in the high skill condition, but increased slightly for those in the low skill condition.

Similar models were used to compare mean hesitation time across conditions. First we took the average hesitation length for each participant during each question. The mean of hesitation times for both truthful and deceptive responses were calculated. As the distribution of hesitation time was right-skewed, we used a log transformation to produce a more normal distribution. Nested linear model comparisons showed that, as with response latency, question number had a significant effect on hesitation length, $X^2(3) = 22.93, p < .001$. CA conversational skill, $X^2(1) = 0.39, p = .53$, deception, $X^2(1) = 0.007, p = .94$, and the interaction of skill and deception, $X^2(1) = 0.08, p = .77$, were all non-significant. The means plot in Fig. 5 shows that there was very little difference in hesitation length between conditions.

4.1. Post hoc analysis

The primary experiment analysis shows significant differences in behavior between participants assigned to the CA with high conversational skill that gave tailored (and inherently varied) responses and low conversational skill which gave generic, and ultimately invariant, responses. It is possible, however, that this difference is due to the repetitive nature of the responses in the low conversational skill condition, and that the mere presence of variance in the responses from the high conversational skill was the driving force of the effect, rather than the relevance of the responses. To eliminate the confounding effect caused by the fact that the responses from the low conversational skill CA were not only generic but also repetitive, we gathered additional data in which participants were assigned to interact with a bot that gave varied, but still generic, responses—that is, rather than giving the exact same response for every follow up question, the CA was configured to give a variety of similar, but still generic, responses.

In the conversationally varied condition, the CA randomly selected variations of the static follow-up questions given by the low conversational skill CA. For example, rather than asking the static follow-up question, “Please describe one more detail about the image,” as the second question for every image, the varied CA randomly selected a variation of this question such as “Next describe more detail about the picture.” or “Okay, now tell me another detail about the photo.” Similar variations were made to the second follow-up question. With this manipulation, we can isolate the effects of conversational relevance and variety from the high conversational skill CA. We collected data from 47 additional participants from the same subject population as the main study.

We analyzed the combined results with a repeated measures nested generalized linear model comparison, which yielded very similar results

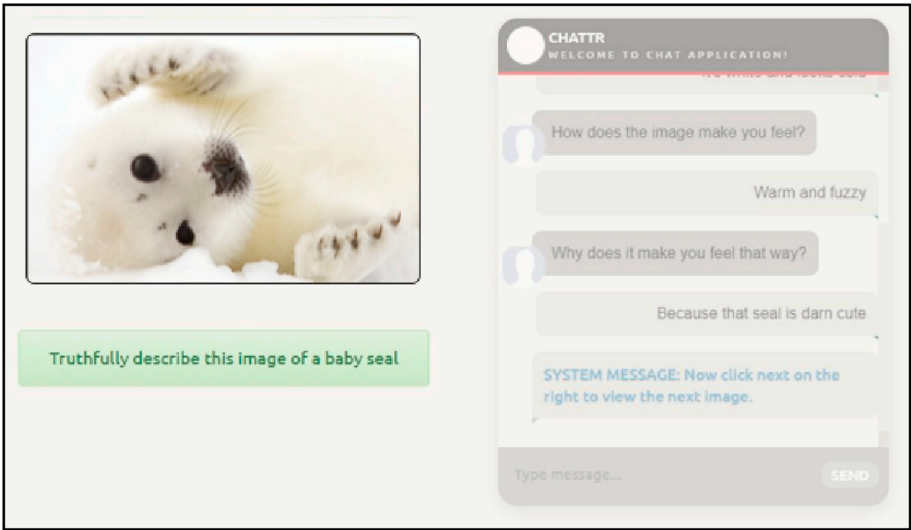


Fig. 2. Experiment interface (Left justified bubbles are text from the chatbot)¹.

Table 1
Tailored and generic interview flow.

(a) Tailored Responses (High skill)	(b) Generic Responses (Low skill)
1 _a . Please describe the contents of the image. 1 _b . [Question based on response to 1 _a] 2 _a . How does the image make you feel? 2 _b . [Question based on response to 2 _a]	1 _a . Please describe the contents of the image. 1 _b . Please describe one more detail. 2 _a . How does the image make you feel? 2 _b . Why does it make you feel that way?

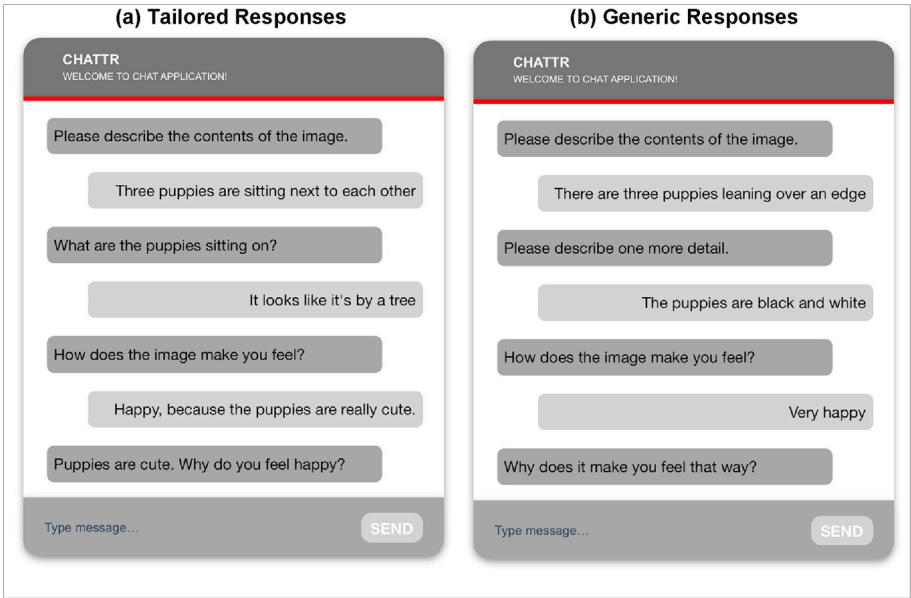


Fig. 3. Sample tailored and generic conversations.

to the previous analysis. Order was still not significant, $X^2(1) = 0.044$, $p = .83$. Question was still significant, $X^2(3) = 110.61$, $p < .001$. CA condition (tailored, variety, and generic) was significant, $X^2(1) = 33.86$, $p < .001$, and deception was still not significant as a main effect, $X^2(1) = 0.29$, $p = .59$. Finally, the interaction between condition and deception approached statistical significance, $X^2(1) = 5.22$, $p = .073$. In the final model, there was a significant effect for the

tailored condition ($b = 1384$, $t = 6.32$, $p < .0001$) but not for the variety condition ($b = 319$, $t = 1.44$, $p = .1522$). There was also a significant effect for the interaction between the tailored condition and deception ($b = -243$, $t = -2.12$, $p = .0344$) but not for the interaction between variety and deception ($b = -28$, $t = -0.24$, $p = .8083$).

Fig. 6 shows the means of response latency of the variety condition compared to the generic and the tailored conditions. The figure shows that the variety condition is more similar to the generic condition than to the tailored condition.

¹ “baby harp seal” by Flickr user CaroLa is licensed under CC BY 2.0/Resized and included in interface.

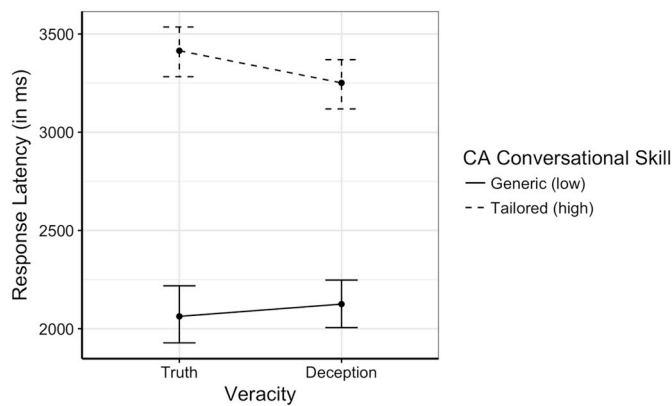


Fig. 4. Means plot of response latency.

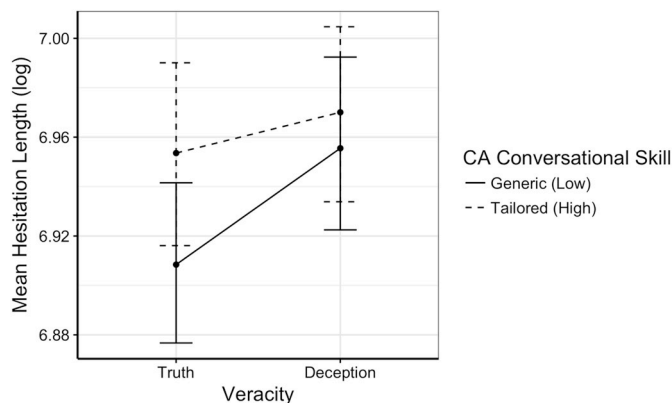


Fig. 5. Means plot of hesitation length.

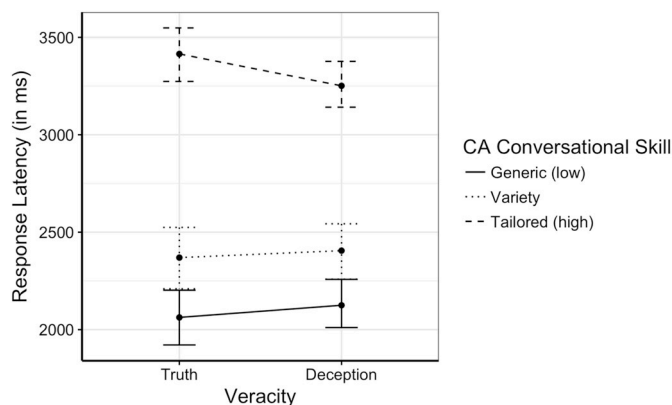


Fig. 6. Means plot of response latency for three CA types.

5. Discussion

5.1. Effects on behavior

Our research question focused on how CA conversational skill affects behavioral responses—operationalized here as response latency and hesitations in typing—during truthful and deceptive communication. We operationalized CA conversational skill with tailored responses and variety in responses. We see a main effect of tailored responses and an interaction between responses and deception on response latency. We did not, however, find any effect on hesitation time.

In our testing of H1, we showed a statistically significant difference

in the direction opposite our hypothesis. The reason for the significance in the opposite direction is likely largely because participants interacting with the low skill CA received four questions, but they were the same four generic questions each time. Participants in the tailored condition had a much more engaging conversation, which required more thought as they read and thought about the follow-up questions.

Our post-hoc analysis revealed one additional outcome that gives insight into the differences in behavior between the two conditions: the presence of response latencies below zero. Negative latency occurs when the user begins typing an answer to a question before the message from the CA is presented to them. In interpersonal communication, it is a violation of conversational norms to talk over someone, or to answer questions while they are still being asked. The same is not true of most human-computer interactions, or even computer-mediated interpersonal interactions, where turn-taking behavior is different than in oral communication (Garcia & Jacobs, 1999; Herring, 2013). As might be expected, negative response latency was more prevalent with the low skill CA, since, after just a few interactions, people knew what questions would be asked before they were displayed. Out of the 4275 response latencies collected for this analysis, 378 (8.8%) were negative. Of the responses with negative latency, 313 (82.8%) were from people in the generic condition, while only 65 (17.2%) came from people in the tailored condition, even though they still had two repeated questions for each image.

The driving force behind the response latency difference is the difference in the richness of the conversation. In the low skill condition, the CA was essentially nonconversational. Since there was no reciprocity in the conversation, participants treated the chat as a transactional interaction to be completed as quickly and efficiently as possible—they treated the CA like a computer system rather than as a social actor. On the other hand, the more skilled CA contributed to the conversation by following Grice's maxim of relation. The responses given by the CA provided feedback that the agent understood what the user said and wanted to know more. In this way, it created a dialogue, as opposed to the question-prompted monologue of the generic CA.

We find that for those in the skilled CA condition, response latency is lower during deception, while for the low skill CA response latency increases slightly relative to each condition's truthful baseline. These findings are particularly interesting as the statistical interaction suggests that the tailored CA questions led to strategic behaviors to avoid detection. Participants with the tailored CA took less time to formulate deceptive responses than truthful ones in an attempt to maintain a natural flow of conversation. This result is in line with IDT, which suggests that deceivers in interpersonal communication engage in strategic behavior to mask their deception (Buller & Burgoon, 1996). We see no such strategic behavior in the presence of the generic CA. We suggest that the reason for this finding is that participants in the generic CA condition felt they could take additional time to formulate their deception without being detected, since they were almost certainly interacting with a computer. Those in the tailored chatbot condition felt more pressured to respond in such a way as to avoid detection, since they were possibly interacting with a human.

We find no support for our hesitation hypothesis with deception (H2b). Similarly, prior research with spoken communication has shown inconsistent results for hesitations during deception (DePaulo et al., 2003). Our study does not find any support for any meaningful effect on pauses in typed communication with a CA.

5.2. Implications

Our results have implications for the development of systems where detecting deception is important. Automated screening systems are being developed for a variety of applications, including border crossings (Higginbotham, 2013; Nunamaker et al., 2011) and job interviews

(HarQen, 2017). Significant research effort is being directed toward making CAs more engaging and humanlike, for example, Google's Duplex system to conduct phone calls (Lomas, 2018). While interacting with a chatbot that gives more tailored responses may be more engaging, this effect may be counterproductive in deception detection applications. Our results indicate that tailored responses, as would be present in a more skilled CA, encourage strategic deception behavior. Therefore, it may be beneficial for developers to create CAs that are less humanlike in scenarios where the detection of deception is important. For example, if a company wanted to make a job screening CA that could detect deception, the findings here indicate that the conversational skill of the CA will be a factor influencing the behavior of respondents. While our results cannot be used to create a predictive model for deception, we have shown that CA behavior influences human behavior in important, measurable ways.

The CA technology described in this paper also paves the way for future advancements in the design of commercial and research-oriented CAs. While humans consider both the content of the message and the nonverbal behaviors accompanying the message when crafting a response (Burgoon, Guerrero, & Floyd, 2009), the current state of the art for CAs is to rely primarily on the content of the message for formulating responses, neglecting the extra behavioral information that may be helpful to interpret messages. Although CAs may not have access to the wide range of cues that humans use to infer meaning in messages, they do have access to novel cues that are not readily available to human observers. With enhancements to the CA described in this manuscript, a system could analyze the behavioral responses (response latency and hesitations) in real time and use this information to further inform the CA's responses.

There are also implications for the design of conversational agents for the completion of repetitive tasks. While the same number of questions were asked in both conditions, users who received the generic questions for each image not only demonstrated more negative response latency—indicative of their lack of engagement—but also rushed through the study, spending 10% less time completing the experiment (unskilled mean 17.0 min, skilled mean 18.9 min). Participant comments following the experiment expressed displeasure and frustration with the conversationally unskilled CA. The following is a representative comment from a user in the low conversational skill CA condition:

"I understand man power takes a lot of [sic] more work and effort. However, the chat was very boring and I probably would have felt more emotions if it wasn't the same exact questions. Very clear it was a generated program."

This frustration can lead to dissatisfaction with the system or task to be completed. For repetitive tasks such as this interview, a CA with low conversational skill may cause a negative experience for users. On the other hand, a CA that is created with the ability to respond in a tailored way to user messages may provide a more engaging and overall better conversation. Therefore, the design of the CA's conversational skill should reflect the goals of the system.

5.3. Limitations and future research

As with any research, there are limitations to this work. First, the deception in this study was limited by a lack of motivation for the deceivers. They were simply told to lie, with no incentive other than the instruction given to them. Future research should examine the impact

of motivation, or different types of deception. If the chatbot is conducting an interrogation rather than a simple Q&A, there is the potential that the social presence factor becomes an even greater driver. Participants were also not invested in the content of the lie. Lying about an image displayed on the screen is different than lying about personal experiences, for example (DePaulo et al., 2003). Personal lies invoke self-presentational instincts such as the desire to make a good impression (Goffman, 1959). This study is also limited to spontaneous lies, as compared to practiced lies. While we gave participants time to think about their deception before being interviewed, the experiment does not reflect many common deception scenarios in which individuals have time to practice and review their responses. Further research would need to investigate other types of lies, including those with a deeper connection to the deceiver.

It is also possible that the CA with low conversational skill gave responses that, through pure happenstance, were perceived by participants as demonstrating understanding. As previously described and illustrated in Fig. 3b, generic follow-up questions can be appropriate—and even relevant—without being tailored to the conversation at hand or intentionally demonstrating understanding. Thus our manipulation of conversational skill is not as clean as one would ultimately desire. However, any generic response that might have conveyed understanding in the generic CA condition would only lead to a weakening of our statistical results (i.e., closer means between the groups). Thus, while future work might consider ways to address this problem, from a practical perspective the differences between the groups are likely more significant than what is present in the current work.

Future research can expand the work here to integrate other aspects of social presence such as embodiment of the agent, or audio-based communication. The current research was text-only, which provides limited cues compared to the voice and body. Because text-to-speech and speech-to-text technology are active areas of research, the capabilities of these systems will continue to improve over the coming years. The addition of embodiment with a face and voice could provide a greater sense of social interaction, resulting in even greater social responses.

6. Conclusion

This research explores the influence of human-like traits in conversational agents by building on Interpersonal Deception Theory to explain the impact of a CA's conversational skill on a person's cues of deception when engaging with the CA. We proposed and tested a model demonstrating how a CA's adherence to Grice's maxim of relation (Grice, 1975) affects communication behavior. We show that introducing enhanced conversational skill decreases response latency during deception in human interviewees. The conversational skill of the CA also increases response latency overall. Those receiving tailored responses from the CA engage in more strategic behavior to manage their response latency during deception. Even small changes to the CA's ability to respond appropriately to users has real and significant effects on communication behavior.

Acknowledgements

This material is based upon work supported by the Nebraska Research Initiative.

Appendix A. Descriptive Statistics

Table 2
Response Latency (times in ms; standard deviations in parentheses).

	Truthful	Deceptive	Totals
Generic	2063 (2347)	2125 (2242)	2097 (2290)
Variety	2370 (2472)	2405 (2401)	2389 (2433)
Tailored	3414 (2125)	3251 (2079)	3326 (2101)
Totals	2597 (2389)	2579 (2293)	

Table 3
Mean Hesitation Length (times in ms; standard deviations in parentheses).

	Truthful	Deceptive	Totals
Generic	1141 (738)	1253 (1047)	1202 (922)
Variety	1267 (974)	1246 (875)	1255 (920)
Tailored	1221 (918)	1247 (971)	1235 (947)
Totals	1179 (831)	1250 (1011)	

References

- Appel, J., Von Der Pütten, A., Krämer, N. C., & Gratch, J. (2012). Does humanity matter? Analyzing the importance of social cues and perceived agency of a computer system for the emergence of social reactions during human-computer interaction. *Advances in Human-Computer Interaction*, 2012. <https://doi.org/10.1155/2012/324694>.
- Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior*. <https://doi.org/10.1016/j.chb.2018.03.051>.
- Bassili, J. N. (2000). Editor's introduction: Reflections on response latency measurement in telephone surveys. *Political Psychology*, 21(1), 1–6.
- Baylor, A. L. (2009). Promoting motivation with virtual agents and avatars: Role of visual presence and appearance. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 364(1535), 3559–3565. <https://doi.org/10.1098/rstb.2009.0148>.
- Berthold, A., & Jameson, A. (1999). Interpreting symptoms of cognitive load in speech input. In J. Kay (Ed.), *UM99 user modeling: Proceedings of the seventh international conference* (pp. 235–244). Vienna: Springer Vienna. https://doi.org/10.1007/978-3-7091-2490-1_23.
- Bickmore, T. W., & Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction*, 12(2), 293–327.
- Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory*, 6(3), 203–242.
- Buller, D. B., Comstock, J., Aune, R. K., & Strzyzewski, K. D. (1989). The effect of probing on deceivers and truth-tellers. *Journal of Nonverbal Behavior*, 13(3), 155–170.
- Buller, D. B., Strzyzewski, K. D., & Comstock, J. (1991). Interpersonal deception: I. Deceivers' reactions to receivers' suspicions and probing. *Communication Monographs*, 58(1), 1–24.
- Burgoon, J. K., & Buller, D. B. (1994). Interpersonal deception: III. Effects of deceit on perceived communication and nonverbal behavior mechanics. *Journal of Nonverbal Behavior*, 18(2), 155–184. <https://doi.org/10.1111/j.1468-2958.1995.tb00365.x>.
- Burgoon, J. K., Guerrero, L. K., & Floyd, K. (2009). *Nonverbal communication*. New York, NY, USA: Pearson Education, Inc.
- Burgoon, J. K., Mayew, W. J., Giboney, J. S., Elkins, A. C., Moffitt, K., Dorn, B., et al. (2015). Which spoken language markers identify deception in high-stakes settings? Evidence from earnings conference calls. *Journal of Language and Social Psychology*, 35(2). <https://doi.org/10.1177/0261927X15586792>.
- Carey, J. (1980). Paralanguage in computer mediated communication. *Proceedings of the 18th annual meeting on association for computational linguistics* (pp. 67–69). Stroudsburg, PA, USA: Association for Computational Linguistics. <https://doi.org/10.3115/981436.981458>.
- Cunningham, J. (2017). Your lying eyes. Retrieved <https://attheu.utah.edu/facultystaff/your-lying-eyes/>, Accessed date: 18 December 2017.
- Dennis, A. R., & Kinney, S. T. (1998). Testing media richness theory in the new media: The effects of cues, feedback, and task equivocality. *Information Systems Research*, 9(3), 256–274. <https://doi.org/10.1287/isre.9.3.256>.
- DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, 129(1), 74–118. <https://doi.org/10.1037/0033-2909.129.1.74>.
- Derrick, D. C., Meservy, T. O., Jenkins, J. L., Burgoon, J. K., & Nunamaker, J. F. (2013). Detecting deceptive chat-based communication using typing behavior and message cues. *ACM Transactions on Management Information Systems*, 4(2), 9.
- Dunbar, N. E., Jensen, M. L., Bessarabova, E., Burgoon, J. K., Bernard, D. R., Harrison, K. J., et al. (2014). Empowered by persuasive deception: The effects of power and deception on dominance, credibility, and decision making. *Communication Research*, 41(6), 852–876. <https://doi.org/10.1177/0093650212447099>.
- Duncan, S., Jr. (1969). Nonverbal communication. *Psychological Bulletin*, 72(2), 118.
- Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and cues to deception. *Psychiatry*, 32(1), 88–100.
- Ekman, P., & Friesen, W. V. (1974). Detecting deception from the body or face. *Journal of Personality and Social Psychology*, 29(3), 288–298. <https://doi.org/10.1037/h0036006>.
- Elkins, A. C., & Derrick, D. C. (2013). The sound of trust: Voice as a measurement of trust during interactions with embodied conversational agents. *Group Decision and Negotiation*, 22(5), 897–913. <https://doi.org/10.1007/s10726-012-9339-x>.
- Fridin, M., & Belokopytov, M. (2014). Acceptance of socially assistive humanoid robot by preschool and elementary school teachers. *Computers in Human Behavior*, 33, 23–31. <https://doi.org/10.1016/j.chb.2013.12.016>.
- Garcia, A. C., & Jacobs, J. B. (1999). The eyes of the beholder: Understanding the turn-taking system in quasi-synchronous computer-mediated communication. *Research on Language and Social Interaction*, 32(4), 337–367. https://doi.org/10.1207/S15327973rls3204_2.
- Gefen, D., & Straub, D. W. (1997). Gender differences in the perception and use of e-mail: An extension to the Technology Acceptance Model. *MIS Quarterly*, 21(4), 389–400.
- Gefen, D., & Straub, D. W. (2004). Consumer trust in B2C e-commerce and the importance of social presence: Experiments in e-products and e-services. *Omega*, 32(6), 407–424. <https://doi.org/10.1016/j.omega.2004.01.006>.
- Giboney, J. S., Brown, S. A., Lowry, P. B., & Nunamaker, J. F. (2015). User acceptance of knowledge-based system recommendations: Explanations, arguments, and fit. *Decision Support Systems*, 72. <https://doi.org/10.1016/j.dss.2015.02.005>.
- Goffman, E. (1959). *The presentation of self in everyday life*. New York: Anchor Books.
- Goldman-Eisler, F. (1961). A comparative study of two hesitation phenomena. *Language and Speech*, 4(1), 18–26. <https://doi.org/10.1177/002383096100400102>.
- Grice, H. P. (1975). Logic and conversation. In P. Cole, & J. L. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 41–58).
- HarQen (2017). *On-demand phone and video interviewing*. Retrieved from <https://harqen.com/index.php>.
- Hassanein, K., & Head, M. (2007). Manipulating perceived social presence through the web interface and its impact on attitude towards online shopping. *International Journal of Human-Computer Studies*, 65(8), 689–708. <https://doi.org/10.1016/j.ijhcs.2006.11.018>.
- Hayes, P., & Ford, K. (1995). Turing test considered harmful. *Proceedings of international joint conference on artificial intelligence* (pp. 972–977).
- Heerink, M., Kröse, B., Evers, V., & Wielinga, B. (2010). Relating conversational expressiveness to social presence and acceptance of an assistive social robot. *Virtual Reality*, 14(1), 77–84. <https://doi.org/10.1007/s10055-009-0142-1>.
- Herring, S. C. (2013). Relevance in computer-mediated conversation. In S. C. Herring, D. Stein, & T. Virtanen (Eds.), *Pragmatics of computer-mediated communication* (pp. 245–268). Berlin: Mouton de Gruyter. <http://info.ils.indiana.edu/~herring/relevance.pdf>.
- Hibbeln, M., Jenkins, J. L., Schneider, C., Valacich, J. S., & Weinmann, M. (2014). Investigating the effect of insurance fraud on mouse usage in human-computer interactions. *International conference on information systems, auckland, New Zealand, December 14–17* (pp. 1–16).
- Higginbotham, A. (2013). *Deception is futile when big brother's lie detector turns its eyes on you*. Wired Magazine. Retrieved from <https://www.wired.com/2013/01/ff-lie-detector/>.
- 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(2), 393–420. <https://doi.org/10.1080/07421222.2016.1205924>.

- Joyce, R., & Gupta, G. (1990). Identity authentication based on keystroke latencies. *Communications of the ACM*, 33(2), 168–176. Retrieved from <https://pdfs.semanticscholar.org/ba16/9305d66ae123b794246d1ba24d86f710bfff0.pdf>.
- Kalman, Y. M., & Gergle, D. (2014). Letter repetitions in computer-mediated communication: A unique link between spoken and online language. *Computers in Human Behavior*, 34, 187–193. <https://doi.org/10.1016/j.chb.2014.01.047>.
- Kang, P.-H., & Gratch, J. (2010). Virtual humans elicit socially anxious interactants' verbal self-disclosure. *Computer Animations and Virtual Worlds*, 21(May), 473–482. <https://doi.org/10.1002/cav>.
- Kellermann, K. (1992). Communication: Inherently strategic and primarily automatic. *Communication Monographs*, 59(3), 288–300.
- Kirakowski, J., O'Donnell, P., & Yiu, A. (2007). The perception of artificial intelligence as "human" by computer users. *Proceedings of the 12th international conference on human-computer interaction: Intelligent multimodal interaction environments* (pp. 376–384). Retrieved from <http://dl.acm.org/citation.cfm?id=1769632>.
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. Gainesville, FL: University of Florida Technical Report A-8.
- Levine, T. R., & McCornack, S. A. (1996). A critical analysis of the behavioral adaptation explanation of the probing effect. *Human Communication Research*, 22(4), 575–588.
- Lin, J.-S. C., & Chang, H.-C. (2011). The role of technology readiness in self-service technology acceptance. *Managing Service Quality*, 21(4), 424–444. <https://doi.org/10.1108/09604521111146289>.
- Lomas, N. (2018, May). *Duplex shows Google failing at ethical and creative AI design*. Retrieved March 19, 2019, from <https://techcrunch.com/2018/05/10/duplex-shows-google-failing-at-ethical-and-creative-ai-design/>.
- McCornack, S. A. (1992). Information manipulation theory. *Communication Monographs*, 59(1), 1–16. <https://doi.org/10.1080/03637759209376245>.
- Miller, R. B. (1968). Response time in man-computer conversational transactions. *Proceedings of the December 9–11, 1968, fall joint computer conference, Part I on - AFIPS'68 (fall, Part I)*, 267. <https://doi.org/10.1145/1476589.1476628>.
- Morrissey, K., & Kirakowski, J. (2013). 'Realness' in chatbots: Establishing quantifiable criteria. In M. Kurosu (Ed.). *Human-computer interaction. Interaction modalities and techniques: 15th international conference, HCI international 2013, Las Vegas, NV, USA, July 21–26, 2013, proceedings, Part IV* (pp. 87–96). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-39330-3_10.
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81–103. <https://doi.org/10.1111/0022-4537.00153>.
- Nass, C., Moon, Y. M., & Carney, P. (1999). Are people polite to computers? Responses to computer-based interviewing systems. *Journal of Applied Social Psychology*, 29(5), 1093–1110. <https://doi.org/10.1111/j.1559-1816.1999.tb00142.x>.
- Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. *Proceedings of the ACM CHI conference on human factors in computing systems* (pp. 72–78). New York, New York, USA, New York, USA: ACM Press. <https://doi.org/10.1145/259963.260288>.
- Nunamaker, J. F., Derrick, D. C., Elkins, A. C., Burgoon, J. K., & Patton, M. W. (2011). Embodied conversational agent-based kiosk for automated interviewing. *Journal of Management Information Systems*, 28(1), 17–48. <https://doi.org/10.2753/MIS0742-1222280102>.
- Pickard, M. D., Burns, M. B., & Moffitt, K. C. (2013). A theoretical justification for using embodied conversational agents (ECAs) to augment accounting-related interviews. *Journal of Information Systems*, 27(2), 159–176. <https://doi.org/10.2308/isis-50561>.
- Pollina, D. A., & Barretta, A. (2014). The effectiveness of a national security screening interview conducted by a computer-generated agent. *Computers in Human Behavior*, 39, 39–50. <https://doi.org/10.1016/j.chb.2014.06.010>.
- Proudfoot, J. G., Boyle, R., & Schuetzler, R. M. (2016). Man vs. Machine: Investigating the effects of adversarial system use on end-user behavior in automated deception detection interviews. *Decision Support Systems*, 22–33. <https://doi.org/10.1016/j.dss.2016.02.008>.
- Raddziwill, N. M., & Benton, M. C. (2017). *Evaluating quality of chatbots and intelligent conversational agents*. Retrieved from <http://www.masswerk.at/elizabot/eliza.html>.
- Rodrigo, M. M. T., Baker, R. S. J., Agapito, J., Nabos, J., Repalam, M. C., Reyes, S. S. J., et al. (2012). The effects of an interactive software agent on student affective dynamics while using an intelligent tutoring system. *IEEE Transactions on Affective Computing*, 3(2), 224–236.
- Schuetzler, R. M., Giboney, J. S., Grimes, G. M., & Nunamaker, J. F. (2018). The influence of conversational agent embodiment and conversational relevance on socially desirable responding. *Decision Support Systems*, 114(October), 94–102. <https://doi.org/10.1016/j.dss.2018.08.011>.
- Schuetzler, R. M., Grimes, G. M., Elkins, A. C., Burgoon, J. K., & Valacich, J. S. (2016). Talking and typing: A comparison of vocalic and keystroke dynamics features for deception detection. In M. Jensen, T. Meservy, J. K. Burgoon, & J. F. Nunamaker Jr. (Eds.). *Report of the HICSS-49 credibility assessment and screening technologies symposium*.
- Schultze, U. (2010). Embodiment and presence in virtual worlds: A review. *Journal of Information Technology*, 25(4), 434–449. <https://doi.org/10.1057/jit.2010.25>.
- Shin, D.-H. (2013). Defining sociability and social presence in Social TV. *Computers in Human Behavior*, 29(3), 939–947. <https://doi.org/10.1016/j.chb.2012.07.006>.
- Shin, D.-H., & Choo, H. (2011). Modeling the acceptance of socially interactive robotics: Social presence in human-robot interaction. *Interaction Studies*, 12(3), 430–460. <https://doi.org/10.1075/is.12.3.04shi>.
- Sporer, S. L., & Schwandt, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology*, 20(4), 421–446.
- Steiner, J. F. (2012). Rethinking adherence. *Annals of Internal Medicine*, 157(8), 580–585.
- Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., et al. (2009). Universals and cultural variation in conversation. *Proceedings of the National Academy of Sciences*, 106(26), 10587–10592.
- Sundar, S. S., Bellur, S., Oh, J., Jia, H., & Kim, H.-S. (2016). Theoretical importance of contingency in human-computer interaction. *Communication Research*, 43(5), 595–625.
- Taboada, M. (2006). Spontaneous and non-spontaneous turn-taking. *Pragmatics*, 16(2–3), 329–360. <https://doi.org/10.1075/prag.16.2-3.04tab>.
- Temple, D. E., & Geisinger, K. F. (1990). Response latency to computer-administered inventory items as an indicator of emotional arousal. *Journal of Personality Assessment*, 54(1–2), 289–297. <https://doi.org/10.1080/00223891.1990.9673994>.
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, LIX(236), 433–460. Retrieved from <http://www.jstor.org/stable/2251299>.
- Warth, G. (2017, July 9). *SDSU professor helps design lie-detecting animated kiosk*. The San Diego Union-Tribune. Retrieved from <http://www.sandiegouniontribune.com/news/education/sd-me-sdsu-robotics-20170629-story.html>.
- Waytz, A., Heafner, J., & Epley, N. (2014). The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology*, 52(0), 113–117. <https://doi.org/10.1016/j.jesp.2014.01.005>.
- Wilcox, B. (2017). *ChatScript*. Retrieved from <https://github.com/bwilcox-1234/ChatScript>.
- Williams, E. (1977). Experimental comparisons of face-to-face and mediated communication: A review. *Psychological Bulletin*, 84(5), 963–976. <https://doi.org/10.1037/0033-2909.84.5.963>.
- Zhou, L. (2005). An empirical investigation of deception behavior in instant messaging. *IEEE Transactions on Professional Communication*, 48(2), 147–160. <https://doi.org/10.1109/TPC.2005.849652>.
- Zhou, L., Burgoon, J. K., Twitchell, D., Qin, T., & Nunamaker, J. F., Jr. (2004a). A comparison of classification methods for predicting deception in computer-mediated communication. *Journal of Management Information Systems*, 20(4), 139–165.
- Zhou, L., Burgoon, J. K., Zhang, D., & Nunamaker, J. F. (2004b). Language dominance in interpersonal deception in computer-mediated communication. *Computers in Human Behavior*, 20(3), 381–402.
- Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. *Advances in Experimental Social Psychology*, 14(1), 2–51.