Association for Information Systems
AIS Electronic Library (AISeL)

AMCIS 2004 Proceedings

Americas Conference on Information Systems (AMCIS)

December 2004

Investigating Deception in Cyberspace

Holtjona Galanxhi-Janaqi
University of Nebraska-Lincoln

Fiona Nah
University of Nebraska-Lincoln

Investigating Deception in Cyberspace

Holtjona Galanxhi-Janaqi
University of Nebraska-Lincoln
hgalanxh@unlnotes.unl.edu

Fiona Fui-Hoon Nah
University of Nebraska-Lincoln
fnah@unlnotes.unl.edu

ABSTRACT

The use of anthropomorphic avatars provides Internet users the opportunity and freedom to present their desired identity. As cyberspace becomes a heaven for deceptive behavior, human-computer interaction research will need to be carried out to study and understand these deceptive behaviors. The objective of this research is to investigate the behavior of deceivers and truthtellers in the cyberspace environment and to examine how their intention to deceive may influence the choice of their avatars. The effect of using avatars to support communication will also be studied. This research provides insights on communication and trust issues in avatar environments.

Keywords

Avatars, communication, deception, trust

INTRODUCTION

The use of anthropomorphic avatars provides new ways for people to interact. An avatar is a virtual representation of oneself that other users can see or interact with in a virtual environment. Avatars were first used in computer gaming, and are now used in a variety of environments beyond entertainment, including education, virtual meeting rooms, sales activities, fashion, trade shows, army and government.

Identity is essential for assessing the trustworthiness of the source and the reliability of information. The beauty, and sometimes misfortune, of the Internet is that it offers the opportunity for people to experiment with their identity (Suler, 1999). Often the cyberspace becomes a heaven for deceptive behaviors. It is, therefore, important to study these deceptive behaviors in order to not only better educate the users but also provide virtual environments where the design discourages or inhibits deceptive behaviors.

In this research, we investigate the behavior of deceivers and truthtellers in the cyberspace environment and study how the intention to deceive may influence one's choice of avatars. The differences between text-only and avatar-supported computer-mediated communication (CMC) will also be studied in relation to resultant trust perceptions.

BACKGROUND

Social interaction on the Internet has unique characteristics: 1) anonymity, due to the fact that relationships may be formed without the ubiquitous requirements of physical presence and proximity; 2) the individual can choose when to log on and off and repeatedly rewrite what he/she wishes to say, which gives him/her far greater control than is usual for a relationship happening in real time. Identity construction in cyberspace is directly related to the nature of the interactions and opportunities offered by the environment (Talamo and Ligorio, 2001). In communication, knowing the identity of those with whom you communicate is essential for understanding and evaluating an interaction. In most chat forums, participants can select from a variety of forms, including male, female, neutral, animals, cartoons, or some hybrid form. The anonymity is very different than that found in text-only chat environments, where only the name you have chosen publicizes your online identity. Multimedia programs allow people to express their identity visually, rather than through written words only.

Although visual signals may indicate aspects of a person's character and real life identity, they are too easily manipulated. Online identities can be as real or as ambiguous as a person wants them to be. An avatar can be seen as emphasizing some visual aspects, or hiding others, similar to a mask worn during a party (Talamo and Ligorio, 2001).

CMC in no way guarantees that a user's declared identity is real (Riva, 2002). In CMC, the user can intentionally mislead in many different ways or can behave chaotically. Although deceiving someone about your identity in the real world is difficult,

it seems to be far easier online, since there are unlikely to be relevant identity clues available offline (Riva, 2002). In the physical world, there is an inherent unity to the self, for the body provides a compelling and convenient definition of identity.

Research evidence from the interpersonal realm reveals that as much as one-third of daily conversations include some form of deception – broadly construed to include concealed, evasive, ambiguous, or exaggerated information as well as outright lies (Buller, Burgoon, Buslig and Roiger, 1996; Ekman, 1996). Deception, then, stands as a threat to successful collaborative work (Burgoon, Stoner, Bonito and Dunbar, 2003). Deception can be defined as the active transmission of messages and information to create a false conclusion (Buller and Burgoon, 1996) such as identity. Deception isn't irrational but definitely rational, being the result of strategic reasoning (Barzel, 1998). Some deceptive behaviors are quite harmful to individuals or to the community; others are innocuous, benefiting the performer without injuring the group. Some are clearly deceptions, meant to provide a false impression; others are more subtle identity manipulations, similar to the adjustments in self-presentation that are typical in many real world situations.

THEORETICAL BACKGROUND AND HYPOTHESES

The majority of existing deception research was carried out in non-interactive environments. Because interactive deception is a dynamic rather than static event, it allows deceivers to modify and adapt their communicative performance to changing circumstances (Burgoon et al., 2003). Interpersonal deception theory (Buller and Burgoon, 1996) contends that interactive deception differs in fundamental ways from non-interactive deception. When deception occurs interactively, i.e., when the deceivers create the messages "real time" during interaction with the targeted message recipient (for example, in chat communication), the resultant communication and judgment processes differ from those of non-interactive deception. The participation and interaction between a sender and a receiver should confer a net advantage for the deceiver (Burgoon, Buller and Floyd, 2001). Senders with higher participation (e.g., in dialogue rather than monologue) are able to better adapt to receivers' suspicions and create a more involving, dominant, and pleasant demeanor, making them more successful at evading detection. According to the theory, when communication senders attempt to manipulate messages so as to be untruthful, apprehension arises because their false communication may be detected. Simultaneously, communication receivers try to unveil or detect the validity of that information, causing suspicion about whether or not the sender is deceiving. In this research, we focus on studying interactive (i.e., synchronous) CMC rather than non-interactive (i.e., asynchronous) CMC.

Intention to deceive and choice of avatar

Lying behavior includes manipulating information, strategically controlling behavior, and image management (Buller and Burgoon, 1996). Strategic deception demands mental effort. Among other things, this strategic behavior may affect decisions concerning the choice of avatars. Do deceiving intentions influence the choice of avatars (e.g., gender swapping, neutral form) of the deceiver?

Deceivers are likely to experience cognitive anxiety from the possibility of being detected (Zhou, Twitchell, Qin, Burgoon and Nunamaker, 2003). One study (White and Burgoon, 2001) showed that deceivers felt more anxious and were more concerned about self-presentation than truthtellers prior to the interaction. When people tell lies, there are four underlying mechanisms at work (Zuckerman, DePaulo and Rosenthal, 1981):

- *Arousal*: Lying causes anxiety and arousal, either because of dissonance at conflicting values and behavior, or due to the fear of getting caught.
- *Behavior control*: When one is lying, one tries to control body language that might give it away.
- *Emotion*: Our emotions change when we are lying. For example, *duping delight*, where the liar is secretly pleased at their perceived success. Guilt may also appear.
- *Thinking*: To lie, we usually have to think a lot harder, such as to ensure coherence in our arguments.

This four-factor model of deception ascertains that lying causes psychological arousal, and emotions of guilt and anxiety are predominant felt emotions. Consequently, the following hypothesis was formulated:

Hypothesis 1: Deceivers will face a higher anxiety level than truthtellers.

As suggested by the four-factor model of deception, in order to overcome the higher anxiety level, deceivers are expected to choose avatars that are very different from themselves. Hence, we hypothesize that:

Hypothesis 2: Deceivers will choose to use avatars that are more different from themselves when compared to truthtellers.

Effect of communication medium on trust

According to the media richness theory, richer media should enable users to more quickly communicate and better understand equivocal information. Media richness is measured on a continuum against four criteria: feedback (asking questions or making corrections), multiple cues (transmitting voice inflection, body language, numbers, and symbols), language variety (range of meaning that can be encoded in language symbols), and personal focus (transmitting feelings and emotions) (Daft and Lengel, 1986).

People trust pictures (images) more than printed words and they are more likely to trust television images than newspaper articles (Rubin and McHugh, 1987). Research by several investigators has shown that trust is higher under face-to-face (richer medium) than CMC, and should produce more favorable attributions about another's sincerity and honesty (Burgoon et al., 2003). The study also shows that even in CMC, trust and mutuality can be established without meeting face-to-face (Burgoon et al., 2003). Since chats with avatars are richer and they offer more clues (i.e., voice, facial expressions), feelings and emotions, chats with avatars are expected to be perceived by recipients to be more trustworthy than text-only chats. Thus, we hypothesize that avatar chats will induce a higher level of trustworthiness in the recipients than text-only chats:

Hypothesis 3: The perception of trust by recipients will be higher for avatar chat than for text-only chat.

METHODOLOGY

The study will be carried out as a 2 (deceptive vs. truthful condition) x 2 (communication mode: text-only or avatar) experimental design. The subjects will be randomly assigned to dyads, which will be randomly assigned to one of the four experimental conditions.

A training session will be provided to all participants in order to avoid any nuisance factor related to the novelty of the communication modality. Before the experiment, all participants will answer demographic questions and pre-interaction measurements of communication anxiety. Written instructions will inform the participants that the purpose of the study was to examine how people interact through the use of computers.

The study comprises two synchronous communication modalities – text-only chat vs. avatar chat – and two conditions for interactive communication – deception and truthful. In the deceptive condition, one participant (deceptive participant - Person A in Figure 1) in each dyad will be randomly selected and explicitly instructed to deceive their partner about their thoughts regarding the discussion issues. In the truthful condition, participants taking the role of the truthful person (i.e., Person A in Figure 1) will be instructed to be completely honest as they discuss these issues. The participants that are instructed to deceive will be told that one of the goals of the research is to understand how people go about presenting views and opinions which they do not really believe, and that the ability to do so may sometimes be interpersonally necessary (adopted from White and Burgoon, 2001).

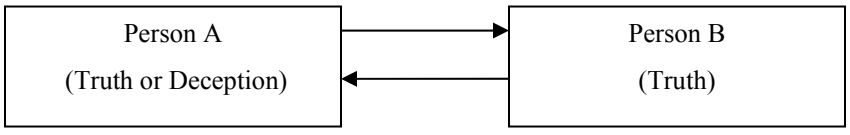

Figure 1: The dyad interaction

The subjects will discuss two topics: 1) calling in sick to school, and 2) giving an opinion about creative endeavor of a friend. These two tasks are adopted (and modified) from White and Burgoon (2001). The order of the topics discussed will be randomized. The subjects will fill out a questionnaire at the end of each task to evaluate their perceptions of the avatar chosen, their anxiety level, and the trustworthiness of their conversation partner.

DISCUSSION OF PILOT RESULTS

We conducted a pilot study with 17 pairs of undergraduate business students. Of these 17 pairs, 9 were in the deceptive condition and 8 in the truthful condition. Each pair was randomly assigned to text-only or avatar environment and to truthful or deception condition.

The pilot results partially support the first hypothesis. Subjects who communicated through the text-only medium experienced higher anxiety levels under the deception condition than the truthful condition, but the same phenomenon was not observed in the avatar environment. The pilot results support the second hypothesis – deceivers chose to use avatars that were more different from themselves when compared to truthtellers. The pilot results do not support the third hypothesis – there was no difference in trust estimates between the text-only vs. avatar-supported chat environment.

LIMITATIONS AND CONCLUSIONS

Consequently, several enhancements are necessary. We noticed that the subjects lacked experience with the software and use of avatars. A more comprehensive training session may be needed before the experiment. Providing better explanations and more specific guidelines to subjects would also help to better achieve the objectives of the study. We are in the process of refining our experimental training materials and procedure based on what we have learned from the pilot study. We plan to conduct the full experiment during the summer and present the results of our study at AMCIS'04.

REFERENCES

1. Barzel, A. (1998), "The Perplexing Conclusion: The Essential Difference between Natural and Artificial Intelligence is Human Beings' Ability to Deceive," *Journal of Applied Philosophy*, 15, 2, 166-178.
2. Buller, D.B., Burgoon, J.K. (1996), "Interpersonal deception theory", *Communication Theory*, 6, 203–242.
3. Buller, D.B., Burgoon, J.K, Buslig, A., Roiger, J. (1996), "Testing interpersonal deception theory: The language of interpersonal deception", *Communication Theory*, 6, 268-289.
4. Burgoon, J. K., Buller, D. B., Floyd, K. (2001), "Does participation affect deception success? A test of the interactivity principle", *Human Communication Research*, 27, 503–534.
5. Burgoon, J.K., Stoner, G.M., Bonito, J.A., Dunbar, N.E. (2003), "Trust and Deception in Mediated Communication", *Hawaii International Conference on System Sciences*.
6. Daft, R., Lengel, R. (1986), "Organizational information, media richness and structural design", *Management Science*, 32, 554–571.
7. Ekman, P. (1996), "Why don't we catch liars?" *Social Research*, 63, 801-818.
8. McKenna, K. Y. A., Bargh, J. A. (1998), "Coming out in the age of the Internet: Identity 'de-marginalization' from virtual group participation", *Journal of Personality and Social Psychology*, 75, 681-694.
9. Riva, G. (2002), "The Sociocognitive Psychology of Computer-Mediated Communication: The Present and Future of Technology-Based Interactions", Cyberpsychology & Behavior, 5, 6, 581-598.
10. Rubin, B., McHugh, M.P. (1987), "Development of Parasocial Interaction Relationships", *Journal of Broadcasting and Electronic Media*, 31, 3, 279-292.
11. Suler, J. (1999), "Gender switching in cyberspace", *The Psychology of Cyberspace,* http://www.rider.edu/~suler/psycyber/genderswap.html (article orig. pub. 1996).
12. Talamo, A., Ligorio, B. (2001), "Strategic Identities in Cyberspace", *Cyberpsychology & Behavior*, 4, 1, 109-122.
13. White, C.H., Burgoon, J.K (2001), "Adaptation and communicative design: Patterns of interaction in truthful and deceptive conversation", *Human Communication Research,* 27, 9-37.
14. Zhou, L., Twitchell, D.P., Qin, T., Burgoon, J.K., Nunamaker, J.F. (2003), "An Exploratory Study into Deception Detection in Text-only Computer-Mediated Communication", *Hawaii International Conference on System Sciences.*
15. Zuckerman, M., DePaulo, M., Rosenthal, R. (1981), "Verbal and nonverbal communication of deception", in L. Berkowitz (editor), *Advances in Experimental Social Psychology*, Academic Press, New York, 1–59.