Does the Contingency of Agents' Nonverbal Feedback Affect Users' Social Anxiety?

Sin-Hwa Kang[1], Jonathan Gratch[2], Ning Wang[2], and James H. Watt[1]

[1]Social and Behavioral Research Lab
Rensselaer Polytechnic Institute
110 8th St. Troy, NY 12180, USA
{kangs, wattj}@rpi.edu

[2]Institute for Creative Technologies
University of Southern California
13274 Fiji Way, Marina Del Rey, CA 90292, USA
{gratch, nwang}@ict.usc.edu

ABSTRACT

We explored the association between users' social anxiety and the interactional fidelity of an agent (also referred to as a virtual human), specifically addressing whether the contingency of agents' nonverbal feedback affects the relationship between users' social anxiety and their feelings of rapport, performance, or judgment on interaction partners. This subject was examined across four experimental conditions where participants interacted with three different types of agents and a real human. The three types of agents included the Non-Contingent Agent, the Responsive Agent (opposite to the Non-Contingent Agent), and the Mediated Agent (controlled by a real human). The results indicated that people having greater social anxiety would feel less rapport and show worse performance while feeling more embarrassment if they experience the untimely feedback of the Non-Contingent Agent. The results also showed people having more anxiety would trust real humans less as their interaction partners. We discuss the implication of this relationship between social anxiety in a human subject and the interactional fidelity of an agent on the design of virtual characters for social skills training and therapy.

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]: Intelligent agents
J.4 [**Social and Behavioral Sciences**]: Psychology

General Terms

Experimentation, Human Factors

Keywords

Contingency of nonverbal feedback, Social anxiety, Rapport, Virtual humans, Agents, Evaluation

1. INTRODUCTION

Virtual humans may allow "outsourcing" of tedious activity (e.g. customer service) and make cumbersome interfaces somehow more socially adroit, but their most intriguing (and controversial) application of such interfaces is the use of virtual humans to teach real humans how to behave more socially. Recent years have seen an explosion in the application of virtual humans toward problems of teaching a variety of social skills including childhood literacy [22], negotiation tactics [28], cultural awareness [29], and prevention of school bullying [15]. A particularly interesting use of the technology is in helping individuals who, through a variety of social disorders, find it difficult to function normally in social situations, such as treating social phobias [16,17,24] or autism spectrum disorder [21,26]. In this sense, virtual humans can mediate social interactions among people who have difficulty in shaping interpersonal relationships or help improve their social skills by interaction with virtual practice.

Social anxiety disorder (also social phobia) has been reported to be "one of the most frequent chronic psychological disorders" with a prevalence of up to 16% in western countries [8]. Roberson-Nay and her colleagues [20] define social anxiety disorder as "overwhelming anxiety and excessive self-consciousness in social situations." Myers [14] describes social anxiety as a condition in which "some people, especially those who are shy or easily embarrassed, feel anxious in almost any situation in which they might be evaluated." Herbelin [8] also quotes a study by the American Psychiatric Association [4]: "Social anxiety disorders or social phobias are characterized by intense and persistent fear of social performance situations in which embarrassment may occur, typically fear of public speaking and/or situations where interactions with others will occur."

Virtual humans and interactive virtual environments (IVEs) have shown promise as therapeutic approaches for psycho-social disorders, such as social anxiety. Herbelin [8] points out that virtual reality exposure therapy (VRET) has demonstrated empirical success in treating social phobias, although current systems must confront a number of technical and theoretical limitations and are confined to very restrictive social situations. For example, such therapy requires placing a user in an anxiety provoking situation and, given the limited social capacities of most virtual human technology, most systems address impoverished social situations, such as speaking in front of a passive audience. Research informing such systems has primarily focused on how interfaces (e.g. head-mounted display vs. desktop), appearances (photo-real vs. cartoonish characters) and simple motor behaviors (e.g. eye gaze patterns or random

Does the Contingency of Agents' Nonverbal Feedback Affect Users' Social Anxiety?, Sin-Hwa Kang, Jonathan Gratch, Ning Wang, and James H. Watt, *Proc. of 7th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2008)*, Padgham, Parkes, Müller and Parsons (eds.), May 12-16, 2008, Estoril, Portugal, pp. XXX-XXX.

body motion) impact the users' feelings of being together and connected. Such research has not investigated virtual humans that are able to respond in meaningful social ways to human subjects.

In this paper, we investigate the relationship between social anxiety in human subjects and the interactional behaviors of virtual humans. Rather than, as with past work, focusing on the physical or behavioral fidelity of virtual humans, here we address the question of *interactional fidelity*: To what extent must virtual humans respond moment-to-moment to a subject's behavior in the way that a human interaction partner might, if placed in the same situation? In this study, we consider the relationship between social anxiety and nonverbal feedback associated with rapport – a phenomenon whereby people tend to mirror each other's behavior and provide rapid positive feedback during the course of an interaction. We demonstrate a significant interaction between social anxiety and different methods of providing such feedback, thereby providing key insights to inform the design of virtual characters for social skills training and therapy.

2. RELATED WORK AND RESEARCH QUESTIONS

Social Anxiety and Co-presence with Virtual Humans

Simply being watched by a virtual human can provoke feelings of anxiety and impair performance on a variety of social tasks [10,11,16,17,19,24,30]. Mirroring findings in human-to-human interaction, this effect depends on several factors. Performance degrades in the presence of a monitoring virtual character on difficult tasks – a virtual audience can actually enhance performance on easy tasks. Further, it seems to require the belief that one is in the presence of a cognizant, potentially judgmental, being and the feeling of being with the other – a factor referred to as *co-presence* in the virtual reality (VR) community [16,24,30]. For example, Hoyt and her colleagues [10] found greater anxiety and poorer performance when a virtual human was presented as an agent controlled by a real human (also referred to as an avatar) but not if described as an agent controlled by a computer.

Of course, virtual humans can do more than simply watch – they can smile and praise or frown and criticize a user's performance – and such feedback plays an important role in this sense of presence and social anxiety. In a series of studies [11,16,17,24], Slater and his colleagues have explored the impact of virtual characters in head-mounted display on users' "fear of public speaking," one major type of social anxiety problems. They investigated the impact of people's interactions with a small group of virtual characters in virtual environments (VEs). They specifically dealt with people who have public speaking anxiety. In these studies, people's rating of self-performance was different based on the level of co-presence that they felt for their "perceived audience interest," when virtual characters were shown on a computer monitor (non-immersed) or in a virtual reality by head-mounted display (immersed), between which there was no significant difference in their co-presence. Participants showed a higher level of anxiety responses to the negative audience than to the positive one. Therefore, they conclude that there would be no difference in people's response to the type of audience between virtual humans and real persons, but would be a difference in their response to the kind of audience's feedback. People would feel more discomfort when they interact with a negative audience while feeling more comfort with a positive audience. They further came to the same conclusion when they conducted another experiment where participants were forced to be more involved in social interactions with virtual humans in a "socially negative manner" [11]. The researchers suggest providing the proper timing of virtual characters' responses to preserve the speakers' co-presence.

Overall, existing studies have investigated how the consequence of social interaction occurs when people interact with virtual humans in immersive or non-immersive VRs measuring users' co-presence and performance or anxiety itself if they feel anxiety under being monitored by those virtual humans as well as by real humans. These studies demonstrate that virtual humans affect people's feeling of being together and connected, regardless of whether they are displayed on a desktop monitor (non-immersive VR) or in head-mounted environments (immersive VR), and the type of feedback determines the degree of users' social anxiety. However, those studies investigated non-social interactions [19,30] or multiple audience interactions [11,16,17,24], not one-on-one social interactions. More importantly, no studies have investigated whether the proper timing of nonverbal feedback of a virtual human is critical to determine people's sense of being together and connected in their interactions while affecting their social anxiety.

Previous work [6] has considered the role of the proper timing of nonverbal feedback of a virtual human, but not its relation to social anxiety. In the following sections, we define the proper timing of feedback as the *contingency* of feedback. We now describe this earlier work and discuss its relationship to social anxiety.

Rapport and Contingent Feedback of Virtual Humans

Rapport -- a construct developed in the study of human-to-human social interaction -- is a feeling of connectedness that seems to arise from rapid and contingent positive feedback between interactants in social interactions [27]. Rapport is posited to increase mutual trust, persuasiveness and feelings of connection between interaction partners, and thus bears similarity to the notion of co-presence studied in VR settings [16,24,30]. It differs through its emphasis on the *joint* timing of behaviors (one interactant's behavior elicits the other's rapid nonverbal feedback such as head nods or postural mirroring) and the assumption that such joint behaviors implicitly convey mutual positive evaluations between interactants. The concepts are different in the sense that one might feel co-presence with a realistically virtual human that stared aggressively at you, but certainly not rapport.

Recent research suggests that virtual humans can establish something akin to rapport with people by producing rapid nonverbal feedback that is elicited by (i.e., contingent on) behaviors produced by the human interaction partner [5,6,7]. Mirroring general findings on rapport, these studies illustrate that the contingency of nonverbal feedback of virtual humans is crucial for interactants' sense of rapport. For example, Gratch and his colleagues [7] created two virtual humans, one that gave contingent feedback to a human storyteller (e.g., head nods and

postural mirroring) and the other that provided essentially random feedback, which was generated independent of the storyteller's behavior, and showed a significant impact on indices of rapport.

Based on the previous conclusion and the literature review, we propose to explore whether the contingent nonverbal feedback of a simple virtual human would promote people's feeling of being connected to the others and engaged in their social interactions while decreasing their social anxiety in one-on-one social interactions.

We will answer the next research questions through this study:

> When people with greater social anxiety experience the contingent nonverbal feedback of an agent, do they feel more rapport and less embarrassment? Do they also show better performance? Further, do they trust and like their interaction partners more?

3. EXPERIMENTAL DESIGN

To investigate the importance of feedback form and contingency, we studied two kinds of virtual characters: one, a "good virtual listener" (the "Responsive" condition) used an autonomous agent, the Rapport Agent, to synthesize head gestures and posture shifts in response to features of a real human speaker's speech and movements, and the other, a "virtual representation of a real listener" (the "Mediated" condition), which reproduced the actual head movements and posture shifts of a real human listener. To investigate whether these two characters could engender feelings of rapport in human speakers comparable to that of real human listeners, we added a "face-to-face" condition, in which speakers spoke directly to real human listeners, for comparison. In a fourth condition, we created a non-contingent response virtual listener (the "Non-Contingent" condition) that provided positive feedback that was unsynchronized with the speaker's movements and speech. Equivalency in feedback frequency across conditions was created by experimental design.

The study was designed with four conditions: Face-to-face (n = 40: 20 speakers, 20 listeners), Mediated (n = 40: 20 speakers, 20 listeners), Responsive (n = 24), and Non-contingent (n = 24), to which participants were randomly assigned using a coin flip. A confederate listener was used in the Responsive and Non-Contingent conditions.

Face-to-Face Condition. In the Face-to-Face condition, the participant talked to a human listener face-to-face.

Mediated Condition. In the Mediated condition, the participant interacted with a virtual character whose head movements and posture reflected the movements of a real human listener. Through the use of a stereo camera and image-based tracking software, the head position and orientation of the listener were captured and displayed by a virtual human character to the speaker in real time. Facial expression feedback was not recognized or displayed.

Responsive Condition. In the Responsive condition, the participant interacted with a virtual character displaying proper listening behaviors. These behaviors were contingent on the recognition of features of the participant's speech (acquired by microphone) and head movements (acquired by a stereo camera) and driven according to predefined behavior-mapping rules. For example, certain prosodic contours in the speaker's voice would cause the character to nod. Facial expressions were not generated.

Non-Contingent Condition. Finally, in the Non-contingent condition, the participant interacted with a virtual character whose behaviors were non-contingent on the behaviors of the speaker. Each participant was presented with a pre-recorded behavior sequence. Equivalency in feedback frequency across conditions was created by experimental design: Following the "yoking" design of Bailenson and Yee [2], the behavior corresponded to what was seen by the previous speaker in the Responsive condition.[1]

3.1 Participants

One-hundred thirty-one people (61% women, 39% men) from the general Los Angeles area participated in this study. They were recruited using Craigslist.com and were compensated $20 for one hour of their participation. On average, the participants were 37.5 years old.

3.2 Procedure

Participants in groups of two entered the laboratory and were told they were participating in a study to evaluate a communicative technology. The experimenter informed participants:

The study we are doing here today is to evaluate a communicative technology that is developed here. An example of the communicative technology is a web-camera used to chat with your friends and family.

Participants signed the consent form, and then the experimenter asked both participants "what's your favorite animal?" The participant whose answer came first alphabetically was assigned the speaker role and the other participant was assigned the listener role. In the Responsive and Non-contingent conditions, the confederate always gave the answer "zebra" to ensure their being assigned to the listener role.

Next, participants were led to two separate side rooms to fill out the pre-questionnaire, which asked for their demographic information and social anxiety related questions.

After both participants completed the pre-questionnaire, participants were led into the computer room. The experimenter then explained the procedure and introduced participants to the equipment used in the experiment.

Next, the speaker remained in the computer room while the listener was led to a separate side room to wait. The speaker then viewed a short segment of a video clip taken from the Edge Training Systems, Inc. Sexual Harassment Awareness video. Two video clips were selected and were merged into one video: The first, "CyberStalker," is about a woman at work who receives unwanted instant messages from a colleague at work (CLIP 1), and the second, "That's an Order!", is about a man at work who is confronted by a female business associate, who asks him for a foot massage in return for her business (CLIP 2).

[1] In the case where duration of the Non-Contingent session was longer than the last Responsive session, the system would loop to the beginning of the recording.

After the speaker finished viewing the video, the listener was led back into the computer room, where the speaker was instructed to retell the stories portrayed in the clips to the listener.

Speakers in all conditions (except the Face-to-Face condition) sat in front of a 30-inch computer monitor and approximately 8 feet apart from the listener, who sat in front of a 19-inch computer monitor. They could not see each other, being separated by a screen. The speaker saw an animated character displayed on the 30-inch computer monitor. Speakers in all conditions (but the Face-to-Face condition) were told that the avatar on the screen displayed the actual movements of the human listener. While the speaker spoke, the listener could see a real time video image of the speaker retelling the story displayed on the 19-inch computer monitor (see Figure 1). The monitor was fitted with a stereo camera system and a camcorder. For capturing high-quality audio, the participant wore a lightweight close-talking microphone and spoke into a microphone headset.

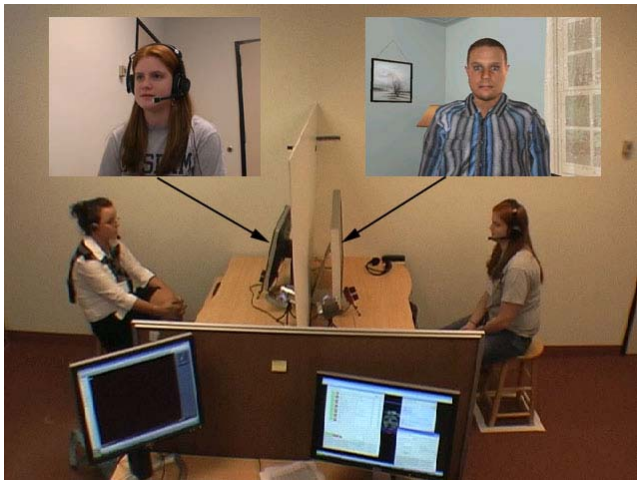

Figure 1. The setup for the experimental conditions

Next, the experimenter led the speaker to a separate side room. The speaker completed the post-questionnaire while the listener remained in the computer room and spoke to the camera what s/he had been told by the speaker.

Finally, participants were debriefed individually and probed for suspicion about the listener using the protocol from Aronson, Ellsworth, Carlsmith, and Gonzales [1]. No participants indicated that they believed the listener was a confederate in the study.

3.3 Equipment

To produce listening behaviors used in the Responsive condition, the Rapport agent first collected and analyzed the features from the speaker's voice and upper-body movements (See Figure 2). Two Videre Design Small Vision System stereo cameras were placed in front of the speaker and listener to capture their movements.

Watson, an image-based tracking library developed by Louis-Phillipe Morency, uses images captured by the stereo cameras to track the participants' head position and orientation [13]. Watson also incorporates learned motion classifiers that detect head nods and shakes from a vector of head velocities. Both the speaker and listener wore a headset with microphone. Acoustic features are derived from properties of the pitch and intensity of the speech signal using a signal processing package, LAUN, developed by Mathieu Morales [5].

Three Panasonic PV-GS180 camcorders were used to videotape the experiment: one was placed in front the speaker, one in front of the listener, and one was attached to the ceiling to record both speaker and listener. The camcorder in front of the speaker was connected to the listener's computer monitor for displaying video images of the speaker to the listener.

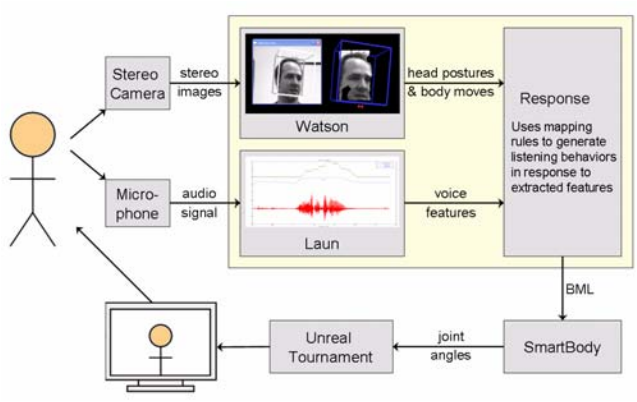

Figure 2. The system architecture of the Rapport Agent

Four desktop computers were used in the experiment: two DELL Precision 670 computers to run Watson and record stereo camera images, one for speaker and one for listener; one DELL Precision 690 to run the experiment system; and one DELL Precision 530 to store logs.

The animated agent was displayed on a 30-inch Apple display to approximate the size of a real life listener sitting 8 feet away. The video of the speaker was displayed on a 19-inch Dell monitor to the listener. All conditions used the same male virtual character (See Figure 1).

3.4 Measurements

3.4.1 Response Variables

Self-Reported Rapport. We constructed a 10-item rapport scale (Cronbach's alpha = .89), presented to speakers in the post-questionnaire. Sample items include: "I think the listener and I established a rapport" and "I felt I was able to engage the listener with my story." Scales ranged from 0 (disagree strongly) to 8 (agree strongly).

Self-Performance. Speakers' self-assessed performance in the speaking task was measured using the scale we constructed (Cronbach's alpha = .85). Sample items include: "I think I did a good job telling the story" and "I had difficulty explaining the story." This scale was issued in the post-questionnaire. Scales ranged from 0 (disagree strongly) to 8 (agree strongly).

Embarrassment. We constructed a 4-item scale which measured the speaker's degree of embarrassment due to feelings of social anxiety (Cronbach's alpha = .81). Sample items include: "I felt awkward telling the story to the listener" and "I felt comfortable

telling the story to the listener." This scale was issued to speakers in the post-questionnaire. Scales ranged from 0 (disagree strongly) to 8 (agree strongly).

Trustworthiness and Likableness. Speakers from all conditions were asked to evaluate the listener on these traits, using the items 'likeable' and 'trustworthy' taken from the dependent measure used in the Krumhuber, Manstead, Cosker, Marshall, and Rosin study [12]. This scale was issued to speakers in the post-questionnaire. Scales ranged from 0 (disagree strongly) to 8 (agree strongly).

3.4.2 Explanatory Variables

Shyness (Social Anxiety). The pre-questionnaire packet included questions about one's shyness [3,9]. Scales ranged from 1 (disagree strongly) to 5 (agree strongly). Sample items include: "I feel tense when I'm with people I don't know well" and "I feel inhibited in social situations."

Private Self-Consciousness. The pre-questionnaire packet contained questions about one's private self-consciousness [23]. Sample items include: "I'm always trying to figure myself out" and "I generally pay attention to my inner feelings." Scales ranged from 1 (disagree strongly) to 5 (agree strongly).

Public Self-Consciousness. The pre-questionnaire packet contained questions about one's public self-consciousness [23]. Sample items include: "I care a lot about how I present myself to others" and "I'm usually aware of my appearance." Scales ranged from 1 (disagree strongly) to 5 (agree strongly).

Researchers [9,14,18] introduced shyness as a concept correlated with social anxiety (or social phobia). According to this definition, social anxiety and shyness could be used to identify the same concept. Hopko and his colleagues [9] found that the revised Cheek and Buss shyness scale is a reliable and stable measurement to measure social anxiety, such as inhibition and discomfort in the presence of others. We used this scale to measure participants' social anxiety in this study.

Furthermore, Scheier and Carver [23] investigated the measurement of self-consciousness and found a positive correlation with social anxiety. Myers [14] stated "shyness is a form of social anxiety characterized by self-consciousness and worry about what others think," so social anxiety contains some features of self-consciousness. In this study, we measured the two constructs separately. Shyness included items indicating some form of aversion to the communication situation that might impede communication and other behaviors, while self-consciousness contained items specifying merely awareness of oneself without direct implications to behavior. Therefore, following conventional practice, we controlled for the effect of self-consciousness variables (both private and public) while assessing the association between people's social anxiety and their rapport and other related constructs after their participation in the experiment.

4. RESULTS

We conducted a hierarchical multiple regression relating Shyness (social anxiety) to Self-performance, Embarrassment, Trustworthiness of the listener, Likableness of the listener, and Self-reported Rapport in each condition, controlling for the potential effect of Private Self-Consciousness and Public Self-Consciousness by introducing them first in the hierarchy [25].

4.1 Results for Self-Performance

The results show that Shyness significantly reduces Self-Performance when people interact with the *Non-Contingent Agent*.

In the condition of the *Non-Contingent Agent*, addition of Shyness to the equation significantly increases the prediction of Self-Performance ($sr^2 = .40$). The ANOVA result specifies that the model as a whole is significant [$F(3, 21) = 6.43$, $p<.05$]. Shyness (beta=-.67, $p<.01$) statistically significantly decreases Self-Performance, when the potential effects of the controlled variables (Private Self-Consciousness, Public Self-Consciousness) are statistically removed.

In the other conditions (Responsive Agent, Mediated Agent, Face-to-Face), there were no statistically significant associations between Shyness and Self-Performance, when Private Self-Consciousness and Public Self-Consciousness were statistically controlled.

4.2 Results for Embarrassment

The results indicate that Shyness significantly increases the amount of Embarrassment when people interact with the *Non-Contingent Agent*.

In the condition of the *Non-Contingent Agent*, addition of Shyness to the equation significantly raises the prediction of Embarrassment ($sr^2 = .44$). The ANOVA result indicates that the model as a whole (which includes both blocks of variables) is significant [$F(3, 21) = 5.90$, $p<.01$]. The variable of Shyness (beta=.69, $p<.01$) makes a statistically significant contribution to the increase in Embarrassment, when the overlapping effects of the controlled variables (Private Self-Consciousness, Public Self-Consciousness) are statistically eliminated.

In the other conditions (Responsive Agent, Mediated Agent, Face-to-Face), there were no statistically significant associations between Shyness and Embarrassment, when Private Self-Consciousness and Public Self-Consciousness were statistically controlled.

4.3 Results for Trustworthiness of the Listener

The results demonstrate that Shyness statistically significantly reduces the perceived Trustworthiness of the Listener, when people interact with real humans.

In the condition of *Face-to-Face*, addition of Shyness to the equation significantly enhances the prediction of perceived Trustworthiness of the Listener ($sr^2 = .19$). The ANOVA result indicates that the model as a whole (which includes both blocks of variables) is significant [$F(3, 14) = 5.09$, $p<.05$]. Shyness (beta=-.51, $p<.05$) significantly diminishes perceived Trustworthiness of the Listener, when the potential effects of the controlled variables (Private Self-Consciousness, Public Self-Consciousness) are statistically eliminated.

In the other conditions (Non-Contingent Agent, Responsive Agent, Mediated Agent), there were no statistically significant associations between Shyness and Trustworthiness of the Listener, when Private Self-Consciousness and Public Self-Consciousness were statistically controlled.

4.4 Results for Self-Reported Rapport

The results show that Shyness is significantly associated with a decrease in people's sense of Rapport when they interact with the *Non-Contingent Agent*.

In the condition of the *Non-Contingent Agent*, addition of Shyness to the equation significantly increases the prediction of Self-reported Rapport (sr^2 = .17). Shyness (beta=-.44, p<.05) significantly reduces Self-reported Rapport when the overlapping effects of the controlled variables (Private Self-Consciousness, Public Self-Consciousness) are statistically eliminated, although the ANOVA result is not significant [F (3, 21) = 1.682, p>.05].

In the other conditions (Responsive Agent, Mediated Agent, Face-to-Face), there were no statistically significant associations between Shyness and Self-reported Rapport, when Private Self-Consciousness and Public Self-Consciousness were statistically controlled.

In addition, the results did not reveal statistically significant associations between Shyness and Likableness of the Listener in any of the four conditions, when Private Self-Consciousness and Public Self-Consciousness were statistically controlled.

5. DISCUSSION AND FUTURE DIRECTIONS

5.1 Summary of Results

In this study, we explored whether the contingency of agents' nonverbal feedback has an effect on the relationship between users' social anxiety and their feelings of rapport, performance, or judgment of interaction partners. The results were compared to those obtained when participants interacted with a real human. Unlike previous work, this comparison in this study was conducted using a one-on-one social interaction and non-immersive environment.

The results indicated that, even after controlling for public and private self-consciousness, users' social anxiety (shyness) significantly decreased their self-performance and self-reported rapport, while increasing their embarrassment in the condition of the *Non-Contingent Agent*. If we look at the results for the Non-Contingent condition in Figure 3, we see the Beta weights for Shyness on Self-Performance and Self-reported Rapport are negative, while the Beta weights for Shyness on Embarrassment are positive. This implies that people having more social anxiety (shyness) show worse performance and feel less rapport, while feeling more embarrassment, when they interact with the Non-Contingent Agent.

The results also demonstrate that users' social anxiety (shyness) considerably reduces their perceived trustworthiness of the listener in the condition of *Face-to-Face* interaction. This indicates that people having more social anxiety (shyness) trust a real human less as their interaction partner, lending support to the contention that virtual humans and virtual environments can facilitate social interactions among people with social anxiety disorders.

Figure 3 also shows the Beta weights for Shyness representing no significant result to provide a comparison with the significant results in the other conditions. (Figure 3 does not show the Beta weights for Shyness associated with Likableness of the Listener since the results did not reveal statistically significant association for this variable in any of the four conditions.)

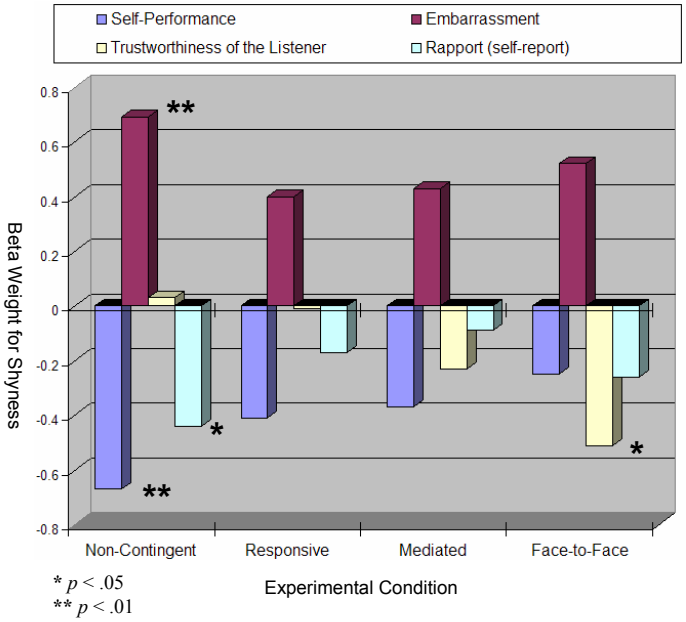

Figure 3. Associations between Social Anxiety (Shyness) and Dependent Variables (Self-Performance, Embarrassment, Trustworthiness of the Listener, and Self-reported Rapport) in each of four experimental conditions

5.2 Implications and Future Work

Previous studies suggest that, when interacting with an agent, the contingency of the agent's feedback has significant influence on the speaker's fluency. The results of this study indicates that the contingency of feedback is especially critical to people who are socially anxious. The results showed strong and significant negative association between social anxiety and perceived self-performance when the agent's feedback was not contingent to users' behavior. In the Non-Contingent condition, high social anxiety also reduced the sense of rapport and made participants feel more embarrassed. These significant associations were only found when the agent provided non-contingent feedback.

However, does this imply that an agent with contingent nonverbal feedback would establish more rapport with people who have greater social anxiety or increase their perceived self-performance? The Responsive Agent is designed to provide positive contingent feedback through backchannel continuers, postural mirroring, and mimicry of certain head movements of a human speaker as suggested by findings from previous studies in human-to-human interaction. The results of this study did not show that there were statistically significant associations between people's social anxiety and their sense of rapport, self-performance, judgment on the listener, and embarrassment when they interacted with the Responsive Agent. This indicates that contingent feedback did not improve the interaction, but neither did it hinder the interaction between the agent and those who

have social anxiety. It also implies that the Responsive Agent performed about the same as a real human listener.

The results further showed that when the speaker interacted with a real human listener, there was a significant negative association between speaker's social anxiety and perceived trustworthiness of the listener. We did not find this association to be significant in any other conditions, especially in the one where the speaker interacted with the Responsive Agent. This implies that people who have greater social anxiety may trust a virtual human more than a real human in social interaction involving storytelling, such as social skills training and therapy.

Several issues require future study. Hoyt and her collaborators [10] report that people showed less anxiety interacting with agents (controlled by a computer), as compared to avatars (controlled by a real human), since they did not think that they were monitored by the agents. In this study, all participants were informed that virtual humans were mediated by real humans, but in fact some of them (Responsive Agents, Non-Contingent Agents) were not. We would like to further investigate whether the sense of being monitored would interact with the effect of social anxiety. Secondly, this study assessed participants' level of social anxiety in indirect ways using the measurements of self-performance, embarrassment, etc. For future work, we also propose to elaborate the experimental design to better investigate the effect of social influence on social interaction with a virtual human by measuring people's sense of social anxiety both before and after their interactions. Third, James and his colleagues [11] point out the importance of participants' repeated exposures in simulated conditions that provoke anxiety. Desensitization from repeated exposure to the situation would lead to a decrease in their anxiety. Further studies that explore the effect of multiple exposures on associations between people's social anxiety and their sense of rapport, performance, and judgment on the listener would be valuable. Finally, we would like to construct behavioral measurements of *Rapport* that would better examine the effect of social interaction between socially anxious people and virtual humans. This additional study would include exploring other related notions such as co-presence and social presence whose measurements have been utilized to evaluate the effect of social interaction between real humans and virtual characters in mediated environments.

Overall, the study presented here adds to our understanding of the relationship between agents' nonverbal feedback, users' social anxiety, and perceived self-performance. Agents with untimely feedback can reduce users' perceived self-performance in people who have higher levels of social anxiety, specifically shyness. In terms of establishing trust, agents with contingent feedback may be in some sense better than real human listeners for people who are shyer. This adds further evidence for the assertion that the contingency of agents' nonverbal feedback matters, especially for shy populations.

6. ACKNOWLEDGEMENTS

We would like to thank Dr. Nicole Kraemer, Dr. Jeremy Bailenson, Dr. James Blascovich, and Dr. Brian Lickel for their helpful comments on this study. This work was sponsored by the U.S. Army Research, Development, and Engineering Command (RDECOM), and the content does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

7. REFERENCES

[1] Aronson, E., Ellsworth, P. C., Carlsmith, J. M., & Gonzales, M. (1990). *Methods of Research in Social Psychology* (2nd Edition ed.). New York: McGraw-Hill.

[2] Bailenson, J. N. & Yee, N. (2005). Digital Chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological Science, 16*, 814-819.

[3] Cheek, J.M. (1983). *The Revised Cheek and Buss Shyness Scale (RCBS)*. Wellesley College: Wellesley MA.

[4] DSM-IV. (1994). *Diagnostic and Statistical Manual of Mental Disorders*. 4 edn. American Psychiatric Association, Washington DC.

[5] Gratch, J., Okhmatovskaia, A., Lamothe, F., Marsella, S., Morales, M., van der Werf, R., et al. (2006). *Virtual Rapport.* Paper presented at the 6th International Conference on Intelligent Virtual Agents, Marina del Rey, CA.

[6] Gratch, J., Wang, N., Gerten, J., Fast, E., & Duffy, R. (2007). Creating Rapport with Virtual Agents. *7th International Conference on Intelligent Virtual Agents,* Paris, France 2007.

[7] Gratch, J., Wang, N., Okhmatovskaia, A., Lamothe, F., Morales, M., & Morency, L.-P. (2007). Can virtual humans be more engaging than real ones? *12th International Conference on Human-Computer Interaction*, Beijing, China 2007.

[8] Herbelin, B. (2005). Virtual reality exposure therapy for social phobia. Ph.D. Thesis. Institut des systèmes informatiques et multimédias, Ecole Polytechnique Federale de Lausanne.

[9] Hopko, D., Stowell, J., Jones, W., & Armento, M. (2005). Psychometric properties of the revised Cheek and Buss shyness scale. *Journal of Personality Assessment*, 84(2), 185-192. Lawrence Erlbaum Associates, Inc.

[10] Hoyt, C., Blascovich, J., & Swinth, K. (2003). Social inhibition in immersive virtual environments. *Presence*, Vol. 12, No. 2, April 2003, 183-195. Massachusetts Institute of Technology.

[11] James, L., Lin, C., Steed, A., Swapp, D., & Slater, M. (2003). Social anxiety in virtual environments: Results of a pilot study. *CyberPsychology & Behavior*, Vol. 6, No. 3, 2003, Mary Ann Liebert, Inc.

[12] Krumhuber, E., et al. (2005). Temporal aspects of smiles influence employment decisions: A comparison of human and synthetic faces. *11th European Conference Facial Expressions: Measurement and Meaning*. Durham, United Kingdom.

[13] Morency, L.-P., Sidner, C., Lee, C., & Darrell, T. (2005). *Contextual Recognition of Head Gestures.* Paper presented at the 7th International Conference on Multimodal Interactions, Torento, Italy.

[14] Myers, D. (1999). *Social Psychology*, McGraw-Hill College.

[15] Paiva, A., Dias, J., & Aylett, R. (2005). Learning by Feeling: Evoking Empathy with Synthetic Characters. *Applied Artificial Intelligence special issue on "Educational Agents - Beyond Virtual Tutors", 19*(3-4), 235-266.

[16] Pertaub, D.-P., Slater, M., & Barker, C. (2001). An Experiment on public speaking anxiety in response to three different types of virtual audience. *Presence: Teleoperators and Virtual Environments* 11:68–78.

[17] Pertaub, D.-P., Slater, M., & Barker, C. (2002). An experiment on public speaking anxiety in response to three different types of virtual audience. *Presence-Teleoperators and Virtual Environments*, 11 (1), 68-78.

[18] Rapee, R. M. (1998). *Overcoming shyness and social phobia: A step-by-step guide*. Northvale, NJ: Aronson.

[19] Rickenberg R. & Reeves B. (2000). The effects of animated characters on anxiety, task performance, and evaluations of user interfaces. In *Proc. CHI 2000*, ACM Press.

[20] Roberson-Nay, R., Strong, D., Nay, W., Beidel, D., & Turner, S. (2007). Development of an abbreviated social phobia and anxiety inventory (SPAI) using item response theory: The SPAI-23. *Psychological Assessment*, 2007 Mar; 19 (1): 133-145.

[21] Robins, B., Dautenhahn, K., Boekhorst, R. t., & A. Billard. (2005). Robotic Assistants in Therapy and Education of Children with Autism: Can a Small Humanoid Robot Help Encourage Social Interaction Skills? *Special issue "Design for a more inclusive world" of Universal Access in the Information Society, 4*(2).

[22] Ryokai, K., Vaucelle, C., & Cassell, J. (2003). Virtual Peers as Partners in Storytelling and Literacy Learning. *Journal of Computer Assisted Learning, 19*(2), 195-208.

[23] Scheier, M. F. & Carver, C. S. (1985). The Self-Consciousness Scale: A revised version for use with general populations. *Journal of Applied Social Psychology, 15*, 687-699.

[24] Slater, M., Pertaub, D-P., and A. Steed (1999) Public Speaking in Virtual Reality: Facing and Audience of Avatars, IEEE Computer Graphics and Applications, 19(2), March/April 1999, p6-9.

[25] Tabachnick, B. & Fidell, L. (2001). *Using Multivariate Statistics*. Allyn & Bacon.

[26] Tartaro, A. & Cassell, J. (2006). Authorable virtual peers for Autism Spectrum Disorders. Paper presented at *the Combined Workshop on Language-Enabled Educational Technology and Development and Evaluation of Robust Spoken Dialogue Systems at the 17th European Conference on Artificial Intelligenc*e, Riva del Garda, Italy.

[27] Tickle-Degnen, L. & Rosenthal, R. (1990). The nature of rapport and its nonverbal correlates. *Psychological Inquiry*, Vol. 1, No. 4, 285-293.

[28] Traum, D., Swartout, W., Marsella, S., & Gratch, J. (2005). Fight, Flight, or Negotiate. Paper presented at *the Intelligent Virtual Agents*, Kos, Greece.

[29] Zachary, W., Le Mentec, J.-C., Miller, L. C., Read, S. J., & Thomas-Meyers, G. (2005). Human behavioral representations with realistic personality and cultural characteristics. Paper presented at *the Tenth International Command and Control Research and Technology Symposium*, McLean, VA.

[30] Zanbaka, C., Ulinski, A., Goolkasian, P., & Hodges, L. (2007). Social responses to virtual humans: Implications for future interface design. In *Proc. CHI 2007*. ACM Press (2007).