Using an Interactive Animated Avatar's Facial Expressiveness to Increase Persuasiveness and Socialness

Jennifer Hyde, Elizabeth J. Carter, Sara Kiesler, Jessica K. Hodgins
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213
{jdtam, ejcarter, kiesler, jkh}@cs.cmu.edu

ABSTRACT
Research indicates that the facial expressions of animated characters and agents can influence people's perceptions and interactions with these entities. We designed an experiment to examine how an interactive animated avatar's facial expressiveness influences dyadic conversations between adults and the avatar. We animated the avatar in realtime using the tracked facial motion of a confederate. To adjust facial expressiveness, we damped and exaggerated the avatar's facial motion. We found that ratings of the avatar's extroversion were positively related to its expressiveness. However, impressions of the avatar's realism and naturalness worsened with increased expressiveness. We also found that the confederate was more influential when she appeared as the damped or exaggerated avatar. Adjusting the expressiveness of interactive animated avatars may be a simple way to influence people's social judgments and willingness to collaborate with animated avatars. These results have implications for using avatar facial expressiveness to improve the effectiveness of avatars in various contexts. Adjusting the expressiveness of interactive animated avatars may be a simple way to influence people's social judgments and willingness to collaborate with animated avatars.

Author Keywords
Animated characters; avatars; facial motion; conversation; persuasiveness; computer-mediated communication.

ACM Classification Keywords
H.5.1 Information Interfaces and Presentation (e.g. HCI): Multimedia Information Systems

INTRODUCTION
Animated characters have the ability to influence how people behave [2] and think [3]. These characters include animated agents controlled by computer algorithms and avatars controlled by people. Researchers have found that agents and avatars are useful because they can improve learning and administer effective therapy (e.g., [3, 24]). When agents and avatars exhibit humanlike behaviors (e.g., facial expressions [4]), people usually respond well to them. These attributes make animated characters attractive as stand-ins for people in educational and therapeutic applications.

Avatars are especially useful in these applications because they enable people to craft new personas, maintain anonymity, and create a sense of social presence remotely. For example, virtual therapy with avatars allows people with social disorders to experience realistic social situations that are difficult to recreate otherwise. Clients and therapists can use avatars to practice various social situations and to act out different roles [24].

For agents and avatars to be effective educators and therapists, these characters may need to exhibit the traits of successful people in these roles. Students and clients prefer human educators and therapists who are warm and respectful [1, 41], and characters with typical cartoon appearances may not convey those qualities well. People may view cartoon characters animated using traditional practices [39], including exaggerated movements and facial expressions, as incompetent and silly. Exaggerating characters with a realistic appearance may result in even worse outcomes, such as unnerving and uncanny movements [19]. To gain people's respect and trust, agents and avatars may need to be somewhat realistic looking. They also may need to exhibit realistic movements and subdued and calm social behaviors. Because people make social judgments based on nonverbal facial cues, it may be beneficial to mimic human facial motion accurately or even with reduced movement.

In a prior experiment using cartoon and comparatively realistic animated characters [20], we measured viewers' perceptions of the characters' social traits, such as extroversion, warmth, and competence. The characters exhibited damped, normal, or exaggerated facial motion while delivering a monologue regarding a personal experience. Facial motion magnitude did not affect participants' ratings of realistic characters' warmth and competence. However, exaggerating the facial motion of cartoon characters resulted in lower ratings of competence. Additionally, facial motion magnitude was positively correlated with ratings of extroversion for both types of characters.

CHI 2015, April 18–23 2015, Seoul, Republic of Korea

ACM 978-1-4503-3145-6/15/04$15.00
http://dx.doi.org/10.1145/2702123.2702465

If facial motion magnitude influences perceptions of the personalities of interactive animated avatars, then animators could customize interactive avatars to suit different situations or user preferences. Customization could be useful for adapting interactive characters to different domains in which it is important for clients and providers to have relationships involving collaboration and trust, including education and therapy. Adjusting people's perceptions of personality by damping or exaggerating an avatar's facial movement could make it easier for therapists to act out different roles. However, it is unknown whether the perceptual effects of facial motion magnitude that exist for animated characters will apply to interactive animated avatars. When people interact and become acquainted with one another, their social judgments of each other are more accurate than when they passively observe one another (e.g., [9]). While interacting, people can create situations that cause other people to exhibit additional personality cues [27]. During collaborative tasks, people's perceptions may also be more accurate because they may care more about determining the credibility of others' information and the best ways to promote compromise.

In this paper, we examine whether the perceptual effects found in our prior work [20] are applicable when people actively interact with an animated avatar rather than passively observe a character. Although we predict that the facial motion magnitude of an interactive animated avatar will influence participants' judgments, other research suggests that other personality cues may override the effect of facial motion magnitude. We expect that participants' perceptions of the interactive animated avatar's extroversion will be positively related to its facial motion magnitude, and perceptions of the interactive avatar's warmth and competence will be negatively related to its facial motion magnitude.

PERSONALITY EFFECTS ON INTERACTION

To understand how avatar personality can be manipulated to influence human-avatar collaboration, we reference the literature in psychology and human-agent interaction. Researchers have found that people's personalities affect how well they collaborate with one another (e.g., [16]). For example, because extroversion is related to dominance, extroverts tend to be more influential and less persuasible than introverts (e.g., [33]). Researchers have also found that open individuals are more willing than closed individuals to consider new ideas and perspectives, making them more persuasible [33].

To improve people's interactions with computer agents, researchers have experimented with adjusting agent behavior and appearance, and these adjustments have influenced people's attitudes, willingness to collaborate, and persuasibility (for a review, see [14]). Nass and colleagues [32] examined participants while they interacted with a computer agent that used a dominant or submissive style of speech while collaborating with the participant using text only. The authors found that when the agents and participants had similar personalities, the agents were more persuasive than when the personalities of the two were complementary. Nass and Lee [31] reported further evidence supporting this similarity-causes-attraction phenomenon when they evaluated participants' attitudes towards extroverted and introverted agents represented by computer-synthesized speech and text. In that experiment, they found that participants preferred the agents with similar levels of extroversion to their own. Additionally, the authors found that participants found similarly extroverted agents more credible and convincing than dissimilar agents.

Studies with embodied agents have also investigated how similarity between agents and users can affect interaction. Researchers have evaluated how agents and users that vary in similarity on characteristics, such as physical resemblance, race, and gender (e.g., [30, 40]) affect collaboration and learning, but few studies have used embodied agents to investigate the effect of agent and user personality on collaboration. Isbister and Nass [23] used an embodied agent to investigate how the consistency of agent personality cues would influence agent-user collaboration. They also analyzed the interactions of users whose personalities were similar and dissimilar to the agents. Their agents exhibited dominant and submissive personalities through their speech and body posture cues. The authors found that agents using speech and posture cues that were consistent with the personality styles of the participants influenced the participants more than agents using inconsistent cues. In contrast with their own prior work, Isbister and Nass [23] found that participants preferred agents with complementary personalities to their own. The authors attributed this difference to the fact that their agents were more interactive and humanlike than the agents used in previous experiments [32, 31]. They reasoned that in prior psychology experiments, participants presented with textual descriptions of others exhibited similarity-attraction, but participants who interacted with others exhibited complementarity effects. The complementarity principle [37, 26] posits that people prefer and seek others who have complementary interaction styles as this will result in a balance of power and less tension. It is unclear from these studies whether people will prefer or interact better with animated avatars that exhibit similar or complementary personalities. In this study, we explore the possibility of an interaction between participants' personalities and the perceived personality of an interactive animated avatar on the participant-avatar interaction.

INFORMATION PROCESSING MODELS

In our experiment, participants discussed a variety of items with a confederate-controlled avatar. The confederate attempted to persuade participants to change their opinions about the items. Research on information processing suggests that the way in which people process information can predict their persuasibility. The heuristic-systematic model [7] and the elaboration likelihood model [34] describe two ways in which people process persuasive information. Systematic or central route processing occurs when people focus on the actual information presented in a persuasive argument, whereas heuristic or peripheral route processing occurs when people use cues like source credibility and attitude judgments to form opinions. People tend to use systematic processing for important decisions that require more attention, and they tend to use heuristic processing for non-essential tasks or when they have time constraints. Researchers have shown that when people

use heuristic processing, they may ignore the actual information presented and instead rely on cues like source credibility, people's opinions, and length of the argument (for a review, see [12]). Because people rely on social judgments when they use heuristic processing, the facial motion magnitude of interactive animated avatars may affect their persuasiveness. For example, an avatar that has exaggerated motion and appears dominant may be more influential than an avatar with damped motion that appears submissive.

AN EXPERIMENT ON THE IMPACT OF AVATAR FACIAL MOVEMENT ON CONVERSATION

We present a study in which participants collaborated with a confederate-controlled avatar that exhibited damped, unaltered, or exaggerated facial motion. We measured the participants' personalities, their perceptions of the confederate and avatar, their conversational behaviors, and the confederate's influence. Participants completed the Desert Survival Problem [25], which has been used in prior work on interaction with agents (e.g., [23, 31]). Participants independently ranked a list of items in order of importance for survival. They then discussed their rankings with the confederate-controlled avatar, who argued for a permutation of the participants' initial rankings. After the discussion, participants re-ranked their items, allowing for measurement of the confederate's influence on participants' rankings. If a participant changed his or her rankings to align better with the confederate's rankings, then the confederate was considered to have influenced the participant. We compare how participants changed their rankings across the different avatar facial motion conditions to analyze the effect of facial motion magnitude on confederate influence.

HYPOTHESES

We expected facial motion magnitude to affect perceptions of the interactive animated avatar in much the same way as it did the perceptions of the non-interactive animated characters [20].

(H1) Facial motion magnitude will positively influence ratings of confederate extroversion and negatively influence ratings of confederate warmth and competence.

In an experiment using a non-interactive animated character to investigate perceptions of emotional expression [21], we found a negative relationship between facial motion magnitude and ratings of naturalness. For this study, we investigate whether this relationship exists with an interactive avatar. We examine naturalness using measures of humanness and eeriness.

(H2) Ratings of humanness will decrease and ratings of eeriness will increase as facial motion magnitude increases.

In addition to generalizing the perceptual results found in prior work [20, 21] to an interactive avatar, we also examined how to customize avatars to improve confederate persuasion. Based on the psychology literature, the confederate's influence should increase with perceptions of increased extroversion. Moreover, Burgoon and colleagues [6] found that greater facial expressiveness was related to increased persuasiveness. Therefore, we expect that increasing the avatar's facial motion will increase the confederate's persuasiveness.

(H3) The confederate's persuasiveness will be positively related to avatar facial motion magnitude.

Figure 1. Example of the AAM tracking a confederate and retargeting her motion to the corresponding character.

METHOD

We designed a between-subjects study in which each participant spoke with a confederate-controlled avatar that exhibited one of three different motion levels (damped, unaltered, exaggerated). The avatar exhibited perceptually equivalent damped (80%) and exaggerated (125%) motion, based on prior work [20]. The avatar was animated in realtime by tracking the face of a female confederate and retargeting her motion to the avatar. Participants communicated with the confederate's avatar using a custom audiovisual telecommunications system.

Realtime Tracking and Animation

To create an interactive animated avatar with realistic facial expressions, we used active appearance models [10, 11, 28], a computer vision algorithm, to track a confederate's facial motion as she spoke and used this motion to animate an avatar. Active appearance models (AAMs) can create mappings between the faces of people and characters. With AAMs, a person and character's face shape and appearance are modeled. Once the model of the person is learned, the AAM tracks his or her face in realtime and moves corresponding points in the character model (Figure 1).

We followed the procedure from Boker, Theobald, and colleagues [5, 38] to exaggerate and damp the facial motion of our avatar. We measured the distance that each AAM vertex moved from a neutral position. The confederate's neutral position consisted of a relaxed face with closed eyes and closed mouth. By selecting the specified neutral pose, the avatar always closed her eyes and mouth when the confederate did so. Given that each AAM vertex changes position during movement, we multiplied those changes in position by scale factors to exaggerate and damp the spatial movements across all features of the face. We exaggerated motion by multiplying the face shape variation by values greater than 1, and we damped motion by multiplying the face shape variation by values less than 1.

Audiovisual Telecommunications System

To facilitate interactive conversations between participants and the confederate, we built a custom-designed desktop audiovisual telecommunications system (Figure 2). The system

allows two people to converse with one another using natural audiovisual cues. Users see each other life-size from the shoulders up. Unlike traditional desktop conferencing systems, this system enables users to make eye contact with one another. The system uses a beam splitter, which acts as a one-way mirror, to hide a camera directly in front of users. The video of a user's partner is reflected off of the beam splitter so that when the user looks at his/her partner's eyes, the user is actually looking directly into the camera. When the confederate uses the system, we can use AAMs to track her motion, map the motion to her avatar, and send the animation to the participant so that it matches with the confederate's voice. Because the AAMs allow us to adjust facial motion, we can spatially exaggerate and damp the characters during conversation.

By using realtime tracking and animation, the participants could have conversations with an avatar that responded with humanlike mannerisms (e.g., blinking, smiling, mirroring, nonverbal feedback). Artists modeled the avatar after the confederate. During the experiment, participants saw the avatar animated with the confederate's facial motion while hearing the confederate's voice.

Task

Participants completed a modified version of the Desert Survival Problem (DSP) [25]. In the DSP, participants are given a survival scenario and must rank a list of items in order of their importance for survival. We modified the original DSP because prior research (e.g., [35]) and our own pilot testing indicated that the original task included outdated items and lasted more than 30 minutes. Following Rae and colleagues [35], we modified the task by exchanging obscure items for more familiar items, and we shortened the item list from twelve to nine (flashlight, pocketknife, raincoat, compass, first aid kit, water, book, vodka, mirror). We also shortened the scenario and emphasized that participants were supposed to match an expert ranking. Shortening the task was necessary because the confederate had to mentally keep track of what she was saying about each item while maintaining contingent discourse. On average, the conversations between participants and the confederate lasted 9.5 minutes.

An experimenter presented the scenario to participants and informed them that they would be discussing the items with a partner. Participants were also told that their partner would appear as an animated cartoon character while they appeared as themselves via video. We provide an excerpt of the scenario below:

> *"... You have just crash-landed in an unfamiliar desert. The helicopter you were in has completely burned ... you and the other passenger have escaped uninjured, but you were unable to contact anyone before the crash. ... Before everything caught fire, you managed to salvage nine items."*

After reading the scenario, participants ranked the items, conversed with the confederate, and then reranked their items. The confederate's rankings differed from the participants' rankings to ensure that the discussion involved why specific items should be ranked higher or lower. As in other research studies (e.g., [35]), the confederate in this study argued for ranking the items in an order that was a consistent permutation of the participants' initial rankings. We used the same permutation as Rae and colleagues [35] did throughout the study so that the confederate would always follow the same argument structure. We allowed the confederate to respond to participant comments to maintain the impression of the confederate as another participant, but all conversation and commenting was structured: the confederate provided two reasons for why each item should be ranked differently, and she only stated these reasons once in each conversation. For example, the two reasons for ranking the flashlight higher were that the flashlight could be used as a nighttime signaling device and that its parts could be used to start a fire. The confederate practiced her role with 26 participants during a pilot phase and prepared responses to common questions and comments. More importantly, the confederate was blind to avatar motion level so that her conversation would not be biased by this knowledge. After the discussion, participants reranked their items. We measured the confederate's persuasiveness by comparing participants' rankings before and after conversing with the confederate.

Materials

Participants completed all surveys, rankings, and questionnaires using a MacBookPro running OSX 10.6. We presented materials on a Dell UltraSharp 2408WFP monitor (22-in.$\times$15.6-in.), and participants used an external keyboard and mouse to complete the digital forms. When participants submitted their initial rankings, the form would automatically print a list of the participants' and confederate's ranks. The experimenter gave this list to the participants and confederate to reference during their conversation. When participants entered their final rankings, the form displayed their initial rankings and the confederate's initial ranking.

Participants

We recruited sixty-four native English-speaking adults (age range: 18-64 years, median age: 23 years, 36 females) using a university experiment scheduling website. We eliminated the data from five participants because the confederate could not maintain the conversation structure, or the participants were too familiar with the DSP. In total, we analyzed the data from 59 participants. The experiment lasted 20 to 40 minutes. We compensated participants for their time. The university's institutional review board approved this study.

Measures

The main independent variable was avatar facial motion magnitude. Participants saw the confederate with either damped, unaltered, or exaggerated facial motion. We also included participant characteristics as control variables. Participants completed a set of questionnaires that measured their age, gender, personality, and computer usage. The personality measure included scores for extroversion, agreeableness, conscientiousness, emotional stability, and openness. We calculated personality scores from participants' responses to the

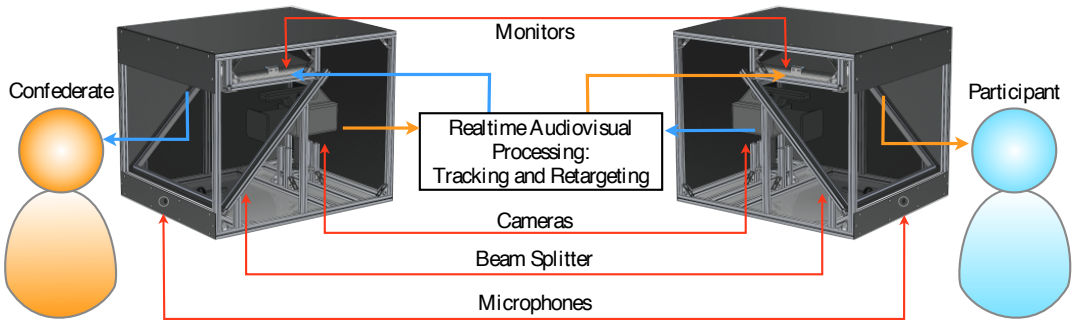


Figure 2. Diagram of telecommunications system.

Ten Item Personality Inventory (TIPI) [17], an experimentally validated measure of the Big-Five personality dimensions (for a review, see [29]). We categorized participants into low and high groups for each of the personality dimensions using Gosling and colleagues' [17] average scores for a sample of 1,813 participants. We used a modified version of Schroeders and Wilhelm's questionnaire [36] to measure participants' computer usage. We replaced their questions on chatting with four questions that asked about chatting with text, voice, video, and avatars. Participants answered how often they engaged in various computer activities. The majority of participants had never chatted with avatars. Because there were not enough participants with this experience to analyze, we did not include participants' experience with avatar chatting in the analysis. The participants had more varied experiences with videoconferencing, but their familiarity with videoconferencing did not affect the results; therefore, we did not include it in the analysis.

We recorded participants as they spoke with the confederate. Annotators labeled the videos with the conversation start and end times and the times when participants looked at the screen. We used these annotations to measure conversation length, percent of gaze (on-screen), and average length of gaze (on-screen). A primary annotator completed gaze annotations for all of the conversations, and a secondary annotator completed one third of each conversation so that we could calculate interrater reliability using Cohen's Kappa (κ). Because the annotators were determining the start and end times of the annotations in addition to labelling the annotations, we first calculated the percentage of aligned annotations. Aligned annotations overlapped by at least 50%. Alignment was 89%, and the annotators agreed on the labels for the aligned annotations with perfect reliability (Cohen's $\kappa = 1$).

To measure confederate influence, we calculated the correlation between the confederate's initial rankings and the participants' final rankings. We used a Spearman's rank correlation coefficient (ρ) to quantify the amount that participants changed their rankings to match the rankings of the confederate. Values closer to 1 indicate that participants changed their rankings to align more with the confederate's rankings. This measure has been used in previous studies to compare rankings from the Desert Survival Problem (e.g., [35]).

Measure	Items	Cronbach's α
Warmth	warm, good-natured, sincere, friendly, well-intentioned, trustworthy	.88
Competence	competent, skillful, capable, efficient, confident, intelligent	.88
Extroversion	extroverted, sociable, dramatic, inhibited	.84
Humanness	artificial – natural, synthetic – real, inanimate – living	.82
Eeriness	reassuring – eerie, numbing – freaky, ordinary – supernatural	.61

Table 1. Self-report measures.

Participants rated the confederate on scales of warmth, competence, and extroversion, and they rated the avatar on scales of humanness and eeriness (Table 1). We asked participants to rate, "*How ____ was your partner?*" on the warmth, competence, and extroversion items using five-point rating scales (i.e., *1=Not at all, 2=Slightly, 3=Moderately, 4=Very, 5=Extremely*). We used items from a questionnaire developed by Fiske and colleagues [13] to measure warmth and competence. To measure confederate extroversion, we used items from a previous study [20], but we removed items from the extroversion measure to improve reliability. We used semantic differential scales from the humanness and eeriness indices developed by Ho and MacDorman [18]. After rating the confederate and avatar, the participants could submit additional comments, which we used in a qualitative analysis.

Procedure

An experimenter told participants that they would be evaluating a new animation technique by conversing with another participant (the confederate), who would appear as an animated avatar. The experimenter also explained that participants would appear as themselves to the other participant. Participants completed an initial survey (see Measures), listened to the experimenter read the DSP scenario, and then ranked the items using a computer form. Afterwards, the experimenter gave each participant a paper copy with both the participant's and confederate's rankings. The experimenter then directed the participant to sit in front of the audiovisual telecommunications system and introduced the confederate to the participant. Next, the experimenter instructed the

participant and confederate to discuss their rankings, emphasizing that their goal was to match an expertly ranked list. The experimenter instructed the participant and confederate to discuss the items in the order that they appeared on the participant's list and to begin the conversation once she left the room. She told the participant that she would return in 15 minutes to transition to the next part of the study. However, the participant could get the experimenter if the task took less time. When the participants and confederate finished speaking, the experimenter led the participants back to the computer to submit their final rankings. Participants then completed more surveys to measure their perceptions of the confederate and avatar. The experimented ended when the experimenter thanked, debriefed, and paid the participants.

RESULTS

We conducted analyses of variance (ANOVAs) and post-hoc contrasts using joint F-tests to examine the effects of facial motion magnitude on participants' perceptions and behaviors. We included participants' personality traits and gender as control variables. We also conducted a qualitative analysis on the voluntary comments that 24 (41%) participants submitted.

Effects on Perceptions

We expected facial motion magnitude to influence participants' perceptions of the interactive avatar in ways similar to how facial motion magnitude influenced perceptions of non-interactive animated characters (H1 and H2). In our previous study [20], facial motion magnitude influenced ratings of extroversion, warmth, and competence. We replicated the previous effect of facial motion magnitude on extroversion in this study, $F(2, 48) = 3.28$, $p = .05$ (Figure 3(a)). The effect size ($\eta^2 = .10$) was medium to large according to Cohen's [8] conventions. Participants rated the confederate as less extroverted when she used the damped avatar compared to when she used the unaltered or exaggerated avatars, $F(1, 48) = 5.11$, $p = .03$, $\eta^2 = .08$ and $F(1, 48) = 4.42$, $p = .04$, $\eta^2 = .07$, respectively. Avatar facial motion magnitude did not affect perceptions of confederate warmth (Figure 3(b)) or competence (Figure 3(c)), $F(2, 48) = 1.45$, $p = .25$ and $F(2, 48) = 0.28$, $p = .75$, respectively. The effect sizes ($\eta^2 = .05$ and $\eta^2 = .01$, respectively) were small. These results partially support hypothesis H1: avatar facial motion magnitude related positively to perceptions of extroversion but had no influence on perceptions of warmth and competence.

We also hypothesized (H2) that participants would perceive exaggerated avatars as less human and more eerie than damped avatars. We found no effect of facial motion magnitude on participants' ratings of the avatar's humanness and eeriness, $F(2, 48) = 1.52$, $p = .23$ and $F(2, 48) = 0.30$, $p = .74$, respectively. The effect sizes ($\eta^2 = .05$ and $\eta^2 = .01$, respectively) were small. Although the humanness and eeriness ratings were not different across facial motion levels, eleven participants in the exaggerated condition made comments, and ten of these referred to discomfort with the avatar. Below are two comments from participants in the exaggerated condition:

> *"The voice of the character was very good, and her responses were natural and well-timed. It was weird at times because it was so well-timed and casual, while her animation still lacked. The animation exists in the same space in my head as The Polar Express — very human, but just slightly off enough to make me feel uncomfortable."*
>
> - Participant 7

> *"Felt a bit taken back at first impressions, but got used to the picture after sometime [sic]."*
>
> - Participant 35

Five participants in the damped condition made comments, four of which referred to the avatar's lifelikeness and realism. The fifth participant commented on the confederate's helpfulness. Here are two comments that indicate that participants found the avatar to be natural in the damped motion condition:

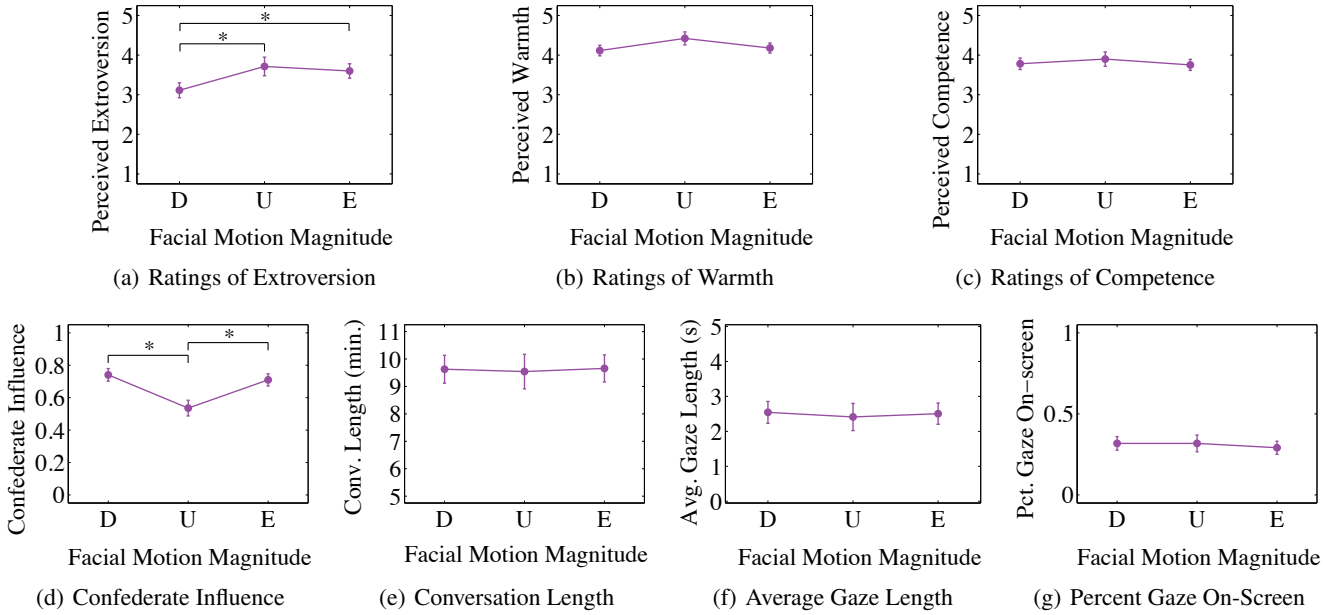
> *"Wow! The ability to capture the various extents of the smile was really impressive and really did convey a good sense of realism to the 'other person' I was speaking to."*
>
> - Participant 15

> *"I was surprised by the detail of the animation on my partner Rachel's face. I could see every little facial movement as she spoke, and it gave her a very life-like quality. It didn't seem like I was talking to an animated image, but a real person."*
>
> - Participant 46

From this partial qualitative analysis, we conclude that the avatar may have appeared less natural in the exaggerated condition than in the damped condition, mimicking the result from our previous study with non-interactive characters [21].

Effects on Behavior

We hypothesized that the confederate would be more persuasive when she appeared as the exaggerated avatar than when she appeared as the damped avatar (H3). We found a significant main effect with large effect size ($\eta^2 = .18$) of facial motion magnitude on confederate influence, $F(2, 48) = 7.61$, $p < .01$. Participants changed their rankings following the confederate's recommendations more when the confederate appeared as the damped and exaggerated avatars compared to when she appeared as the unaltered avatar, $F(1, 48) = 13.84$, $p < .001$ and $F(1, 48) = 10.29$, $p < .01$, respectively (Figure 3(d)). These effect sizes were large ($\eta^2 = .16$) and medium to large ($\eta^2 = .12$), respectively. This result partially supports the hypothesis (H3) because the exaggerated avatar more than the unaltered avatar persuaded participants to change their rankings. However, the damped avatar also was more successful than the unaltered avatar at persuading participants to change their rankings. We found no effect of facial motion magnitude on conversation length, average length of gaze, or percent of gaze on-screen (Figures 3(e)-3(g)).



(a) Ratings of Extroversion (b) Ratings of Warmth (c) Ratings of Competence

(d) Confederate Influence (e) Conversation Length (f) Average Gaze Length (g) Percent Gaze On-Screen

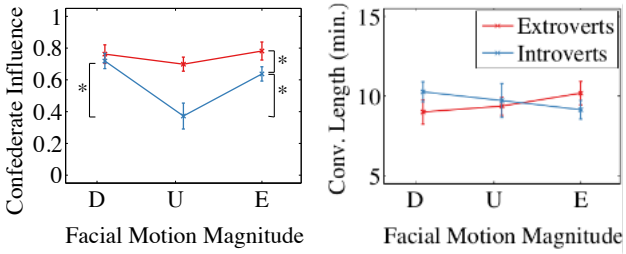
Figure 3. The graphs illustrate the influence of avatar facial motion magnitude (D=damped, U=unaltered, E=exaggerated) on participants' ratings and behavior. The asterisk indicates significance at $\alpha = .05$ or less.

Other Results

We explored the possibility that participants' personalities might moderate the effect of avatar facial motion magnitude on the confederate's influence. There are two competing theories regarding whether people with similar or complementary personalities work better together (see the Personality Effects on Interaction section). Therefore, we considered both of these theories in our exploratory analysis.

The statistical analysis, although not significant at the $\alpha = .05$ level, indicates a possible interaction of avatar facial motion and participant extroversion on confederate influence with medium effect size, $F(2, 48) = 2.93$, $p = .06$, $\eta^2 = .07$ (Figure 4(a)). The confederate influenced extroverted participants more than introverted participants across avatar motion levels. This result was supported by a significant main effect with large effect size of participant extroversion level on confederate influence, $F(1, 48) = 15.49$, $p < .001$, $\eta^2 = .18$. Extroverted participants may have been more persuasible because they also were more open to new experiences, $r(59) = 0.39$, $p < .01$. Researchers have found that open people are more persuasible than closed people [15, 33], and the data from this study aligns with this idea. We found a significant main effect with medium to large effect size of participants' openness on confederate influence, $F(1, 48) = 9.36$, $p < .01$, $\eta^2 = .11$. The participants' extroversion and openness may have trumped any similarity or complementary personality effects on confederate influence.

In contrast, introverted participants accounted for most of the variation in confederate influence. When the confederate appeared as the damped and exaggerated avatars, she persuaded introverted participants more than when she appeared as the unaltered avatar, $F(1, 48) = 14.74$, $p < .001$ and $F(1, 48) = 8.77$, $p < .01$, respectively. The effect sizes were large ($\eta^2 = .17$) and medium to large ($\eta^2 = .10$), respectively. The data suggest that introverts may have interacted better with the confederate when her extroversion level appeared as both similar and complementary to the participants' extroversion level. An alternative explanation for why introverted participants agreed more with the confederate when she appeared extroverted is that the exaggerated avatar may have intimidated participants. Both of these situations could induce similar changes in rankings.



(a) Confederate Influence (b) Conversation Length

Figure 4. Graphs illustrating the interaction of participants' extroversion and avatar facial motion magnitude on (a) confederate influence and (b) conversation length. The asterisk indicates significance at $\alpha = .05$ or less.

We also examined how long participants conversed with the avatar that matched or complemented their extroversion levels. The interaction of participant extroversion and avatar facial motion magnitude was not significant, $F(2, 48) = 1.55$, $p = .22$, $\eta^2 = .04$. Figure 4(b) illustrates a slight similarity-causes-attraction phenomenon. If participants had more extreme levels of extroversion, then their interactions with the avatars may have differed more.

Lastly, we explored whether participants would look more at the confederate when she appeared as the avatar that matched or complemented their extroversion levels. We found no significant interaction of participant extroversion and facial motion magnitude on participants' gaze behaviors. Motion did not influence introverted nor extroverted participants' percent of gaze on-screen, $F(2, 48) = 0.32$, $p = .73$, $\eta^2 = .01$, nor did it influence the average length of gaze on-screen, $F(2, 48) = 0.08$, $p = .92$,$\eta^2 = .003$. The qualitative analysis revealed that four participants purposefully avoided looking at the avatar. Three of these participants interacted with the exaggerated avatar and the other participant interacted with an unaltered avatar. Participants in the exaggerated condition made comments such as,

> *"I found it hard to look into the eyes of the animated character, even though her voice sounded natural."*
> - Participant 38

The participant in the unaltered condition noted difficulty making eye contact in daily life. Therefore it is unclear whether the participant's gaze avoidance was due to the avatar or to its social presence.

> *"It was hard for me to keep looking at it. Though, I have trouble keeping eye contact normally."*
> - Participant 39

Despite noting their discomfort with looking at the avatar, three of the four participants had higher than median percentages of on-screen gaze (Figure 5).

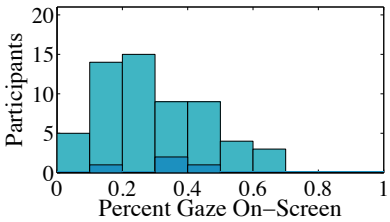


Figure 5. Distribution of participants' gaze by percentage on-screen. The darker bars represent the participants who wrote comments saying that they avoided looking at the avatar.

DISCUSSION

Prior work indicated that the facial motion of non-interactive characters could influence viewers' perceptions of the characters' personalities. However, interactivity provides more opportunities and sources of information that people can use to judge personality. It was unclear whether the perceptual effects of facial motion found using non-interactive characters would apply to interactive avatars. We examined the perceptions and behaviors of participants who spoke with an interactive animated avatar that was controlled in realtime by a confederate. The avatar's facial motion was exaggerated and damped to investigate the influence of avatar facial motion magnitude during conversation. Our results indicate that even with other strong personality cues, facial expression magnitude influences perceptions of an animated avatar. Avatar facial motion affected participants' perceptions of the confederate's extroversion, warmth, and competence. Participants found the confederate less extroverted when she appeared as the damped avatar compared to when she appeared as the unaltered or exaggerated avatars. Our qualitative analysis also indicated that participants may have found the damped avatar more realistic and lifelike than the exaggerated avatar. Several participants commented on feeling uncomfortable while interacting with the exaggerated avatar. These comments match our previous findings regarding participants' perceptions of non-interactive characters' naturalness [21].

In this study, we also examined the confederate's influence on participants by comparing the participants' rankings before and after their conversations with the confederate. We found that the participants who interacted with the damped or exaggerated avatars were more likely to change their rankings as suggested by the confederate than participants who interacted with the unaltered avatar. When participants saw the damped and exaggerated avatars, they may have relied more on heuristic processing. To use heuristic processing, participants would have spent more attention and cognitive resources on interpreting the avatar's facial motion and the confederate's manner of speaking than on the actual arguments.

In our exploratory analysis, we found that the behavior of the introverted participants was responsible for most of the variation in confederate influence. We suspect that the introverted participants responded well to the damped avatar because its extroversion level was more similar to their own, supporting the idea of similarity-causes-attraction. Alternatively, the exaggerated avatar may have intimidated introverted participants due to its perceived increase in extroversion level and dominance. It is unclear from these results which of the changed perceptions also changed the confederate's influence. Exaggerated and damped facial motion may have increased attention to the avatar, resulting in more confederate influence, or the combination of participants' personalities and avatar facial motion may have been responsible. Future work using participants with more extreme levels of extroversion would help answer this question. Differences in behavior may be more extreme between highly extroverted and introverted participants. Participants with extreme levels of extroversion also may behave more consistently than participants with mid-levels of extroversion.

From the quantitative analysis, we found that participants spent on average 30% of their conversation time looking at the avatar, and we saw no differences in participants' gaze behavior across facial motion conditions. Some participants noted their aversion to looking at the avatar, but these participants' behavior indicated that they looked at the avatar more than average. These participants may have consciously forced themselves to look at the avatar because they were aware of their discomfort. We did not include a video condition, so it is unclear whether participants looked less at the avatar than they would a person.

Our goal is to design interactive characters for therapy and education when video may not be appropriate. We found that some participants responded well to the avatar even though the experience may have been unsettling at first.

"It was only errie [sic] due to the life like emotions depicted by the Interface, other than that I want one :-)"
- Participant 10

"Great experience. I thought if more humans were like it I might interact socially more."
-Participant 48

It would be interesting to examine participants over time to see if they become more accustomed and receptive to interactive avatars. We did not measure participants' social anxiety, but it is also possible that the participants who responded well to the avatars have difficulty interacting with people. Future research should investigate the use of interactive avatars with people who have social and communicative difficulties.

Our method for changing the facial expressiveness of the avatar may have created an emotive mismatch between the avatar's facial and vocal expressions. We examined this type of emotive mismatch in a prior study [21], in which animations combined characters with different levels of facial expression with emotional speech of varying intensity. The mismatch affected naturalness ratings of the characters, but only if the characters were extremely emotional. We believe that the conversations between participants and the confederate in this study were less emotional than the animations in the prior work. Therefore, we do not believe that a mismatch in facial expression and speech occurred.

Additionally, our experiment used a single confederate and avatar, therefore we cannot generalize our findings to other confederates, face shapes, or genders. However, in our prior work using multiple confederates and actors [20, 22], we found that confederate characteristics did not interact with avatar facial expressiveness to influence perceptions.

Finally, there are limitations associated with the tracking and animation method that we used. Although the avatar displayed the facial motion of the confederate, the confederate's body motion was not tracked. The torsos of the animated characters moved rigidly with respect to a pivot located at the characters' mouths. Additionally, the characters always faced forward as they were created from 2D data, and they were unable to nod or shake their heads. It is possible that the lack of head nods and shakes may have influenced the participants' acceptance of our human-like avatar. One participant commented on the avatar body's rigid movement and lifelike face. The juxtaposition of realistic and unrealistic behaviors may have affected participants' feelings of discomfort, especially for the exaggerated character. It is unclear whether participants would be more comfortable with a less realistic looking character that exhibited similar behaviors. In the future, a similar interactive experiment comparing people's responses to cartoon and realistic characters would help answer this question. A comparison of different tracking and animation techniques would also be useful.

CONCLUSION

We determined that avatar facial motion magnitude positively affected participants' ratings of the confederate's extroversion. Facial motion also affected people's comfort with the avatar and their persuasibility. Participants changed their beliefs to align more with the confederate when she was represented by a damped or exaggerated avatar compared with an unaltered avatar. These results suggest that interactive avatars could be used to improve collaboration in education and therapy. Additionally, although some participants found it difficult to maintain eye contact with the avatar, others felt that conversing with the avatar was easier than conversing with other people. Our results suggest that interactive avatars may be suitable in therapy and counseling for adults with social and communicative difficulties.

ACKNOWLEDGMENTS

We thank our participants and collaborators, R. Browne, C. Evans, R. Backstrom, K. Miller, T. Simon Kreuz, P. Carr, I. Matthews, M. Mahler, C. Heiner, and M. Danver. This work was funded by Disney Research and award R03HD068816 from the Eunice Kennedy Shriver Institute of Child Health and Human Development of the National Institutes of Health.

REFERENCES

1. Ackerman, S. J., and Hilsenroth, M. J. A review of therapist characteristics and techniques positively impacting the therapeutic alliance. *Clin Psychol Rev 23*, 1 (2003), 1–33.
2. Andrist, S., Mutlu, B., and Gleicher, M. Conversational gaze aversion for virtual agents. In *Proc. IVA 2013*, Springer (2013), 249–262.
3. Bailenson, J. N., Yee, N., Blascovich, J., Beall, A. C., Lundblad, N., and Jin, M. The use of immersive virtual reality in the learning sciences: Digital transformations of teachers, students, and social context. *J Learn Sci 17*, 1 (2008), 102–141.
4. Bates, J. The role of emotion in believable agents. *Commun ACM 37*, 7 (1994), 122–125.
5. Boker, S. M., Cohn, J. F., Theobald, B.-J., Matthews, I., Brick, T. R., and Spies, J. R. Effects of damping head movement and facial expression in dyadic conversation using real-time facial expression tracking and synthesized avatars. *Phil Trans R Soc B 364*, 1535 (2009), 3485–3495.
6. Burgoon, J. K., Birk, T., and Pfau, M. Nonverbal behaviors, persuasion, and credibility. *Hum Commun Res 17*, 1 (1990), 140–169.
7. Chaiken, S. Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *J Pers Soc Psychol 39*, 5 (1980), 752.
8. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. Routledge, 1988.
9. Colvin, C. R., and Funder, D. C. Predicting personality and behavior: A boundary on the acquaintanceship effect. *J Pers Soc Psychol 60*, 6 (1991), 884–894.
10. Cootes, T. F., Edwards, G. J., and Taylor, C. J. Active appearance models. *IEEE TPAMI 23*, 6 (2001), 681–685.

11. Cootes, T. F., Wheeler, G. V., Walker, K. N., and Taylor, C. J. View-based active appearance models. *Image Vision Comput 20*, 9 (2002), 657–664.
12. Eagly, A. H., and Chaiken, S. *The Psychology of Attitudes.* Harcourt Brace Jovanovich College Publishers, 1993.
13. Fiske, S. T., Cuddy, A. J., Glick, P., and Xu, J. A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *J Pers Soc Psychol 82*, 6 (2002), 878–902.
14. Fogg, B. J. Computers as persuasive social actors. In *Persuasive Technology: Using Computers to Change What We Think and Do*, J. Grudin, J. Nielsen, and S. Card, Eds. Science & Technology Books, 2002, ch. 5, 89–120.
15. Gerber, A. S., Huber, G. A., Doherty, D., Dowling, C. M., and Panagopoulos, C. Big Five personality traits and responses to persuasive appeals: Results from voter turnout experiments. *Polit Behav 35*, 4 (2013), 687–728.
16. Gorla, N., and Lam, Y. W. Who should work with whom? Building effective software project teams. *Commun ACM 47*, 6 (2004), 79–82.
17. Gosling, S., Rentfrow, P., and Swann, W. A very brief measure of the Big-Five personality domains. *J Res Pers 37*, 6 (2003), 504–528.
18. Ho, C.-C., and MacDorman, K. F. Revisiting the uncanny valley theory: Developing and validating an alternative to the Godspeed indices. *Comput Hum Behav 26*, 6 (2010), 1508–1518.
19. Hodgkinson, G. The seduction of realism. In *Proc. SIGGRAPH ASIA 2009 Educators Program*, no. 4, ACM Press (2009), 1–4.
20. Hyde, J., Carter, E. J., Kiesler, S., and Hodgins, J. K. Perceptual effects of damped and exaggerated facial motion in animated characters. In *Proc. FG 2013*, IEEE Press (2013), 1–6.
21. Hyde, J., Carter, E. J., Kiesler, S., and Hodgins, J. K. Assessing naturalness and emotional intensity: A perceptual study of animated facial motion. In *Proc. SAP 2014*, ACM Press (2014), 15–22.
22. Hyde, J., Kiesler, S., Hodgins, J. K., and Carter, E. J. Conversing with children: Cartoon and video people elicit similar conversational behaviors. In *Proc. CHI 2014*, ACM Press (2014), 1787–1796.
23. Isbister, K., and Nass, C. Consistency of personality in interactive characters: verbal cues, non-verbal cues, and user characteristics. *Int J Hum-Comput St 53*, 2 (2000), 251–267.
24. Kandalaft, M. R., Didehbani, N., Krawczyk, D. C., Allen, T. T., and Chapman, S. B. Virtual reality social cognition training for young adults with high-functioning autism. *J Autism Dev Disord 43*, 1 (2013), 34–44.
25. Lafferty, J. C., and Eady, P. M. *The Desert Survival Problem*. Human Synergistics, Plymouth, Michigan, 1974.
26. Leary, T. *Interpersonal Diagnosis of Personality*. Ronald Press, 1957.
27. Letzring, T. D. The good judge of personality: Characteristics, behaviors, and observer accuracy. *J Res Pers 42*, 4 (2008), 914–932.
28. Matthews, I., and Baker, S. Active appearance models revisited. *Int J Comput Vision 60*, 2 (2004), 135–164.
29. McCrae, R. R., and Costa Jr, P. T. A Five-Factor theory of personality. In *Handbook of Personality: Theory and Research*, L. A. Pervin and O. P. John, Eds., 2 ed. Guilford Press, New York, NY, USA, 1999, ch. 5, 139–153.
30. Moreno, R., and Flowerday, T. Students' choice of animated pedagogical agents in science learning: A test of the similarity-attraction hypothesis on gender and ethnicity. *Contemp Educ Psychol 31*, 2 (2006), 186–207.
31. Nass, C., and Lee, K. M. Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. *J Exp Psychol-Appl 7*, 3 (2001), 171–181.
32. Nass, C., Moon, Y., Fogg, B., Reeves, B., and Dryer, D. C. Can computer personalities be human personalities? *Int J Hum-Comput St 43*, 2 (1995), 223–239.
33. Oreg, S., and Sverdlik, N. Source personality and persuasiveness: Big Five predispositions to being persuasive and the role of message involvement. *J Pers 82*, 3 (2014), 250–264.
34. Petty, R. E., and Cacioppo, J. T. *Attitudes and Persuasion: Classic and Contemporary Approaches*. Brown, Dubuque, IA, 1981.
35. Rae, I., Takayama, L., and Mutlu, B. One of the gang: Supporting in-group behavior for embodied mediated communication. In *Proc. CHI 2012*, ACM Press (2012), 3091–3100.
36. Schroeders, U., and Wilhelm, O. Computer usage questionnaire: Structure, correlates, and gender differences. *Comput Hum Behav 27*, 2 (2011), 899–904.
37. Sullivan, H. S. *The Interpersonal Theory of Psychiatry*. WW Norton & Co, 1953.
38. Theobald, B.-J., Matthews, I., Mangini, M., Spies, J. R., Brick, T. R., Cohn, J. F., and Boker, S. M. Mapping and manipulating facial expression. *Lang Speech 52*, 2/3 (2009), 369–386.
39. Thomas, F., and Johnston, O. *Disney Animation: The Illusion of Life*. Abbeville Press New York, 1981.
40. Vugt, H. C. V., Bailenson, J. N., Hoorn, J. F., and Konijn, E. A. Effects of facial similarity on user responses to embodied agents. *ACM TOCHI 17*, 2 (2008), 7:1–7:27.
41. Young, S., and Shaw, D. G. Profiles of effective college and university teachers. *J High Educ 70*, 6 (1999), 670–686.