

An Eye Gaze Model for Controlling the Display of Social Status in Believable Virtual Humans

Michael Nixon
*Institute of Communication, Culture,
Information and Technology
University of Toronto
Toronto, Canada
michael.nixon@utoronto.ca*

Steve DiPaola
*School of Interactive Arts and
Technology
Simon Fraser University
Vancouver, Canada
sdipaola@sfu.ca*

Ulysses Bernardet
*School of Engineering and
Applied Science
Aston University
Birmingham, United Kingdom
u.bernardet@aston.ac.uk*

Abstract—Designing highly believable characters remains a major concern within digital games. Matching a chosen personality and other dramatic qualities to displayed behavior is an important part of improving overall believability. Gaze is a critical component of social exchanges and serves to make characters engaging or aloof, as well as to establish character’s role in a conversation.

In this paper, we investigate the communication of status related social signals by means of a virtual human’s eye gaze. We constructed a cross-domain verbal-conceptual computational model of gaze for virtual humans to facilitate the display of social status. We describe the validation of the model’s parameters, including the length of eye contact and gazes, movement velocity, equilibrium response, and head and body posture. In a first set of studies, conducted on Amazon Mechanical Turk using pre-recorded video clips of animated characters, we found statistically significant differences in how the characters’ status was rated based on the variation in social status.

In a second step based on these empirical findings, we designed an interactive system that incorporates dynamic eye tracking and spoken dialog, along with real-time control of a virtual character. We evaluated the model using a presentational, interactive scenario of a simulated hiring interview. Corroborating our previous finding, the interactive study yielded significant differences in perception of status were found ($p = .046$). Thus, we believe status is an important aspect of dramatic believability, and accordingly, this paper presents our social eye gaze model for realistic procedurally animated characters and shows its efficacy.

Index Terms—procedural animation; believable characters; virtual human; gaze; social interaction; video games

I. INTRODUCTION

Designing good believable characters to inhabit virtual story worlds is an ongoing problem, and remains a major concern within digital games. Researchers have been investigating many different methods for creating believable expressive character that can dynamically adapt either at design-time to their authors’ needs, or during run-time as the situation they find themselves in evolves. The field of believable characters deals with the intersection of visual aesthetics and intelligent agent research. In general terms, intelligent agents are computer systems that possess autonomy, social ability, reactivity, and pro-activeness [1], and are conceptualized and

implemented using anthropomorphic terms, including beliefs, desires, intentions [2], and other types of cognitive models. Adding these models allows agents to participate in ongoing stories, as they possess enough cognitive depth so as to be interesting. Additionally, as the agents are visually represented as a human character, they have to meet a range of expectations for them to appear lifelike and interesting.

Believable character research focuses on providing embodied intelligent agents with overt human bodies and corresponding characteristics for various purposes including entertainment (e.g. *Façade* [3]), training (e.g. the Virtual Humans [4] project), and education (e.g. *Fear Not!* [5]).

To be believable, these characters’ bodies need the ability to communicate using both verbal and non-verbal channels simultaneously. The synchronization of these channels is vital to meeting appropriate expectations from human communication, reducing the potential Uncanny Valley effect and making the computer less of an “unfamiliar interlocutor” [6]. This means such improvements help in a wide range of use cases where appearing human is desirable, including games. Along with accompanying communication, it is important that characters can use non-verbal behavior to send social signals, which are “...complex aggregates of behavioral cues accounting for our attitudes” [7]. This allows them to participate in more dramatically interesting scenarios. Status is one of these important signals, defined as an embodied or felt sense of one’s social position relative to others.

We chose status as the aspect to investigate since it is considered very impactful in the performing arts, and thus will be valuable in expressive virtual performances, including those of NPCs in digital games. Status was first popularized in improv and interactive theater by Johnstone [8], and is now considered a core concept [9], useful for driving dramatic activity. This further builds on status as a signal sent by behaviour rather than overt symbols of e.g. wealth.

Gaze plays a central role in nonverbal communication; empirical research has found high rates of correlation between eye gaze and specific conversational actions [10]. Furthermore, these expectations persist even with androids [11] and seem to underlie the human ability to understand one another; violating them can lead to the Uncanny Valley [12] effect, where view-

ers find characters eerie or inhuman. Besides raw technical improvements to rendering and resolution, modeling agent behavior after expected human behavior provides another way of increasing their believability. As well as enhancement, gaze can function along with other body language to provide contrasting messages or subtle meanings to verbal phrases. We believe this can be synthesized with observations from performance art to allow us to incorporate powerful dramatic concepts that draw on embodied experience. Overall, concern for the gaze of characters, which impacts social signaling and conversational cues, implies a concern for the presentation of social behavior in agents and relationships that are portrayed between agents and between agents and users.

However, gaze has not been used often as a core game mechanic, apart from a few examples such as *L.A. Noire* [13] where the characters trying to deceive you may break eye contact. Typically, character gaze is functional, with manually-animated expressive qualities such in as *Civilization VI*'s interstitial video clips where a sovereign announces they're angry with you and give a unique flourish [14]. *Spore*'s [15] creatures that need to dynamically adapt to a variety of body arrangements are probably one of the most advanced examples of a game system that procedurally handles gaze; still, this is mainly functional to allow characters to survey their environment. Ultimately, since most of the animations within digital games are baked versions of animator-produced work, they are not very flexible or contextual. Therefore, to have the most potential as an effective game mechanic, gaze behavior should be proceduralized, with attention paid to its expressive and functional meanings.

To order to move in this direction by improving the quality of believable intelligent virtual agents' interactions within social settings, we examine how the important behavior of eye gaze can be improved. To do this, we created a social eye gaze model by synthesizing findings about gaze behavior and status-related behavior, drawn from both the psychological and performance art literature. These were developed into six conceptual parameters that operate on a spectrum. to assess the efficacy of the model's static parameters, we conducted an evaluation to allow us to assess it. After these studies validated the social gaze model, it was implemented in an interactive character system, which was then evaluated in a lab-based study.

II. RELATED WORK

Conversations between virtual humans and between users and such characters are an important way of promoting player engagement. On the level of physiological engagement, we know that eye gaze plays an important role in human conversations. Its role in avatar conversations has been reasonably studied, and computational models for controlling the flow of such conversations has been produced [16]. Due to its importance, and at times, central focus in animated scenes, its capabilities are well worth understanding.

Within conversations, gaze is key to regulating conversational flow, indicating attentiveness, and a variety of affective

cues [17]. Argyle and Cook found that people who exhibit high amounts of mutual gaze are perceived as competent, attentive, and powerful. They also describing findings that mutual gaze between adults in Western cultures lasts an average of 1-2s. This range has been found to carry over to virtual agents [18].

Gaze behavior is also coordinated with proximity to others. Equilibrium theory [19], where people reach an equilibrium based on their interpersonal comfort. If one of the two factors is varied, the other will be changed in response. The effect of mutual gaze affecting proximity has also been observed to be similar in virtual environments [20].

For virtual characters, mutual gaze has been found to build rapport with people [21] and increase positive perceptions of affiliation [22]. A functioning gaze model has been found to motivate interest compared to a fixed-gaze [23]. Moreover, [24] examined the contributions of avatars with differing gaze patterns for conversations between remote users by comparing audio-only, random gaze avatars, inferred gaze avatars, and full video feed variations. While the random gaze model did not improve users' experience, the inferred gaze variation, which used the audio stream to determine whether a user was speaking or not, significantly outperformed both the random gaze and audio-only variations. This indicates that avatar animation should reflect the ongoing conversation. This study is also relevant because it uses a dramatic scenario involving a negotiation between aggrieved parties to provide the participants with a meaningful task to undertake.

While many advances are being made in improving intelligent virtual agents, to date, relatively few systems have been designed to directly provide them with the ability to reveal characters inner states via their gaze and/or head behavior. Some of the systems devised so far include [18] who used an animated pair of eyes to display affective signals. Queiroz and colleagues [25] devised a parametric model for automatically generating emotionally expressive gaze behavior. Lance et al. have presented the Gaze Warping Transformation [26], which is a means of generating expressive head and torso movements during gaze shifts based on human motion data. Their later development on GWT lead to them proposing a model of realistic, emotionally expressive gaze [27].

Busso and colleagues [28] proposed a method to generate gaze that quantizes motion-captured head poses into a finite number of clusters, and builds a Hidden Markov Model for each of these clusters based on the prosodic features of the accompanying speech. [29] proposed a parametric gaze and head movement model that is linked with the emotional state machine developed at MIRALab [30], where mood is represented with Mehrabian's Pleasure-Arousal-Dominance (PAD) Temperament Model [31]. Arousal and dominance dimensions are used to drive the parameters of the gaze and head movement models. Custom MPEG-4 Facial Animation Parameter files are then played to provide appropriate character animation.

Additionally, gaze control has been implemented in agent frameworks at various levels. Autonomic behavior is the most likely to be provided, for example with the "Eyes Alive" model

[32] for saccades and Elckerlyc’s implementation of breathing and blinking [33]. Other nonverbal behaviour (NVB) tends to be generated manually by the content creator or automatically processing speech for meaningful content to cue off of.

An important precursor is the work done by Brenda Harger where she proposed using improvisational theater models to improve believable characters [34], [35]. In one study, she showed users simple animations of characters entering a room, while being able to vary the characters movement through adjusting a status parameter. By altering the parameter, participants could see the character performing the same action in different ways. Harger did not formalize a model, but this showed the potential effectiveness of demonstrating status.

Another important investigation into the development of a formalized model of NVB for portraying character attributes is [36]. They found that animators can meaningfully and reliably respond to prompts to produce character animation based on Keith Johnstone’s system of Fast-Food Stanislavski.

III. THE SOCIAL EYE GAZE MODEL

We produced a parametric model for controlling the eye-head gaze behavior of characters to portray differing levels of social status, which is shown in Table I. Cognitive models have been a mainstay in cognitive science, and can be specified at different granularities. This model is of the verbal-conceptual [37] variety, as we try to break down the components of believable gaze and assess their impact. It is a model for portraying status via gaze behavior at a medium-high level of detail, not at the minute or micro level of detail.

For example, there are existing statistical studies of saccade behavior, leading to the “Eyes Alive” model of saccades [38], which indicate how eye-only gaze changes correlate to either speaking or listening in a dyadic conversation. However, for portraying social status, behavior changes involving combinations of body parts should be considered. For this reason, the gaze model portrayed here is evaluated as a whole.

The essential components that affect the model include movement of the important qualifiers such as use of personal space and overt social signaling. The model is based on a number of psychological theories affecting the relationship between these qualifiers and the behavior performance traits that vary depending on personality type. After surveying the literature to compile the most relevant theories, we extracted the descriptions of NVB in dyadic conversations. We then grouped these into the aspects of gaze behavior that were affected. The resulting parameters are labeled following the naming convention for these phenomena established in [39].

The model contains six important components whose variation has been found to signify differences in social status. Table I provides a breakdown of the parameterized behavior that affects the general pattern described above along several continua. Movement toward an end of a continuum occurs simultaneously in the same direction along the other continua, although each parameter will have differently sized differences between gradations. *Length of eye contact* refers to the period of time during which mutual eye gaze is occurring, measured

in seconds. *Length of gazes* describes how long gazes toward the opposite person in the dyad last for, measured in seconds. *Movement velocity* refers to the speed of the eyes, head, and neck during shifts in gaze, measured in degrees per second. *Head posture* describes the overall inclination of the head. *Equilibrium response* indicates the immediate response made when the opposite person in the dyad moves into a closer personal reaction bubble. *Posture* indicates a general stance affecting the entire torso.

Table I
THE PARAMETERS OF THE SOCIAL EYE GAZE MODEL FOR PORTRAYING STATUS THROUGH VARIATIONS IN GAZE

Parameter	Low Status Behavior	High Status Behavior
Length of eye contact	Shorter	Longer
Length of gazes	Shorter	Longer
Movement velocity	Fast	Slow
Head posture	Bowed	Raised
Equilibrium response	Look away sooner	Stare longer
Posture	Lean away	Lean toward

IV. EVALUATION OF THE SOCIAL EYE GAZE MODEL PARAMETERS

We wanted to determine whether the listed parameters produce behaviour that functions as major drivers of audience perception. Furthermore, while we used the baseline established for humans in the psychological literature, we wanted to find workable ranges of values for these parameters. To do so, we first implemented a scenario within SmartBody that would facilitate the comparison of behavioral differences. The videos of these scenarios were then shown to participants using Amazon’s Mechanical Turk (MTurk) platform to compare the conditions and find effective values for the parameters. MTurk has been generally found to be a reliable mechanism for conducting experiments [40].

In the empirical evaluations, we test the hypothesis that an animated character whose gaze movements are faster, whose fixations are shorter, and whose perceived length of eye contact is shorter, will be perceived as lower status than one whose gaze movements are slower, whose fixations are longer, and whose perceived length of eye contact is longer. We do so by comparing the participants’ rating of the character’s status.

A. Method

To demonstrate the feasibility of our model, we used the SmartBody procedural animation framework [41] to animate a short scene from Keith Johnstone’s book on improvisational theater [8]. The scene for this study is set in an office environment with a corporate officer who fires an employee with a criminal record. We chose to use this scenario because it has been successfully used to evaluate the NVB of characters [36], [42], and it explores a situation with dramatic behaviour where social status plays a meaningful role. It provides a simple scene that constrains a number of factors including the length of the exchange, the number of actions that occur, and the personality traits and emotions on display. In addition, based on findings in psychological studies, 30 seconds of

exposure to people is long enough to make an impression about their personality [43], without tiring participants. This scenario was implemented using BML to generate the dialog, as well as a minimal set of accompanying gestures (e.g. head shakes for negative statements, deictic hand waves to refer to the other person) to avoid presenting unnaturally still avatars. These were designed to be emotionally neutral and were not varied between conditions.

1) *Experimental Design*: The scenario was produced to support two conditions, with a corporate officer displaying either high status or low status behavior as they fired an employee. This was achieved by varying only behavior associated with the following three model parameters: length of eye contact, length of fixations, and movement velocity. The scenario was implemented in the SmartBody environment follow the scenario’s script.

Four different evaluations were performed. For each of them, one human intelligence task (HIT) with 50 assignments was loaded onto MTurk; thus, 50 unique workers completed both trials, for a total of 200 participants. We restricted the participants to be those with US accounts. Overall, the average age of the 200 participants was 35.0 ($SD = 11.0$), with 57% reporting their gender as male; the remainder reported female.

In the first evaluation, we used audio recorded by local actors. The corporate officer was represented with a masculine character, and the employee by a feminine character.

In the second evaluation, we replaced the actors’ voices with TTS voices. The computer-generated voices were SAPI-compliant male and female voices from Cereproc.

In the third evaluation, we rotated the camera 180° from the first evaluation’s setup to focus on the employee character, whose NVB was also generated according to the social eye gaze model. The actors’ voices were retained.

To assess to which degree gender plays a role in the perception of status, we tested a forth factor, using a female character as the corporate officer. In this pair of videos, the camera framed the Corporate Officer character as in the first evaluation, while the scenario remained the same. The characters’ voices used the same TTS voices.

We used the Ten-Item Personality Inventory Measure (TIPI) [44] which is frequently used for measuring the “Big Five” personality traits (openness, conscientiousness, extraversion, agreeableness, and emotional stability). TIPI has been used successfully in assessing virtual characters before, e.g. [18].

2) *Procedure*: The experiment was registered on the MTurk website, and workers found it by browsing for HITs. They viewed the experiment “ad”, and this contained the task instructions and informed consent. If they accepted the HIT, they then saw the system capability verification screen. Next, they provided some basic demographic information (age, gender). Then, for each condition, they viewed the video of the scenario and then rated the character. After each video, participants completed the TIPI (10 questions), which uses a 7-point Likert scale response. They were then prompted: “Given that *social status* is a person’s standing or importance in relation to other people, please circle the character’s apparent social status”.

The status question was answered using a 5-point Likert scale, rated from 1 (Very Submissive) to 5 (Very Dominant).

The experiment was presented to workers on MTurk as an HTML webpage with task flow controlled by JavaScript code running locally in each worker’s web browser. Prior to the experiment workers verified their system functionality by listening to a brief audio track and transcribing the text. Trials were constructed by pairing videos of the scenarios in the different conditions of High and Low Status. The participant rated the character’s personality and status after each one. Randomizing the order of the videos counterbalanced the conditions.

B. Results

In our statistical analysis we compared the change in status rating across the different experimental conditions.

A univariate repeated measures analysis of variance was conducted to compare the effect of social gaze model condition on status. Results indicated there was a significant difference between the High Status condition ($M = 3.64$, $SD = 1.16$) and the Low Status condition ($M = 3.45$, $SD = 1.24$), $F(1, 196) = 5.93$, $p = .016$. No statistically significant interactions between the other conditions (character, voice) were present in this analysis. This finding supports our hypothesis and suggests that the differing gaze behavior produced by the model does have an effect on the perceived social status of animated characters.

V. INTERACTIVE EVALUATING THE DYNAMIC SOCIAL GAZE MODEL

This section presents a study that investigates how to send social signals related to status by varying a virtual human’s eye gaze and how people evaluate virtual humans engaging them in social behavior. We integrated the social eye gaze model into a behavior control module for a character animated with the SmartBody procedural animation system. By combining this with a chatbot-driven dialog system, we created a job interview simulator with two different variations on the character’s eye gaze. Participants then practiced interviewing the characters while wearing an eye tracker, which allowed the character’s behavior to vary based on mutual eye gaze. Again, they reported that the varying gaze model changed their impression of the character’s social status, providing further support for our social gaze model.

A. Method

1) *System Design*: The system, as shown in Figure 1, facilitates gathering input from a participant, and simultaneously controlling a SmartBody embodied conversational agent (ECA) based on that input. The two inputs are eye tracking data, which is monitored for mutual gaze with the ECA, and voice input, which is used to generate dialog responses. All scripts and libraries were programmed in Python.

The main sensor used is an eye tracking headset from Pupil Labs [45], which coordinates a gaze camera and a world camera to determine where the participant is looking. We used

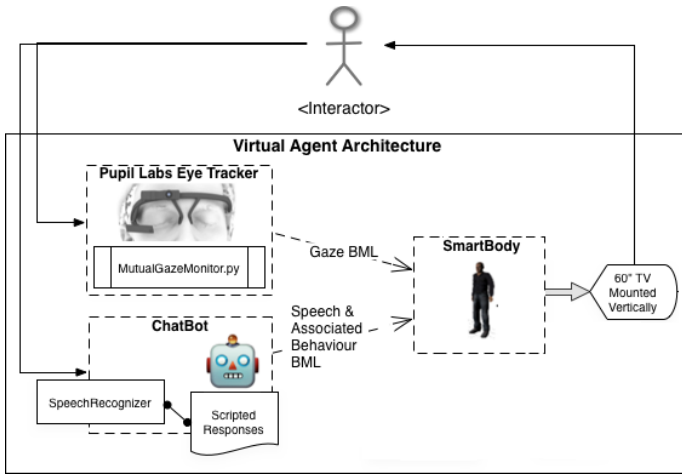


Figure 1. Diagram of the virtual agent architecture for eye tracking study.

the world camera’s medium setting of 1280x720 to capture 60fps. The gaze camera captured 640x480 at 120fps, using IR illumination. Specifically, we defined a virtual region on the screen surrounding the character’s eyes. These “virtual surfaces” can be seen in Figure 2, registered relative to fiducial markers. Doing so allowed Pupil Capture to record and stream information about the participants’ gazes on the specified surface, generating surface visibility reports and gaze counts. This was found to be a very useful way to stream real-time information about gaze on a user-defined area of interest within an arbitrary software application.

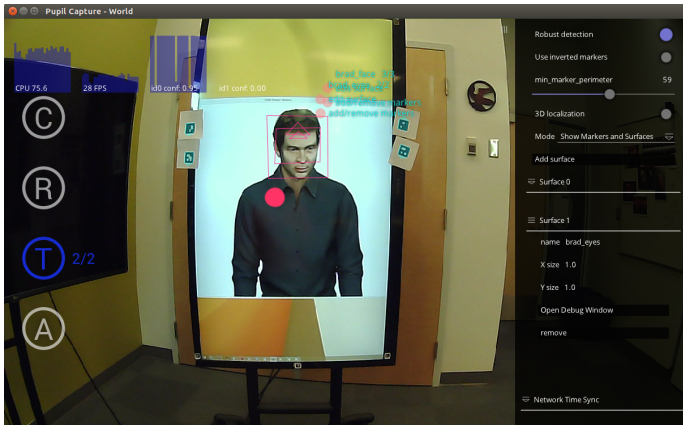


Figure 2. Sample eye tracking surfaces as seen in Pupil Capture.

To provide dialog for the system, we used a custom chat bot [46], [47], which uses the Google Speech API to obtain audio from a microphone, and then relies on a probabilistic match system to provide the best response to a given prompt [48]. We provided the bot with dialog relating to a hiring scenario. The chat bot uses StompSender, which is a Python class that can send control messages to SmartBody via Apache ActiveMQ using the STOMP protocol. The chat bot produced vocal output using computer-generated voices (TTS) that were SAPI-compliant male and female voices. TTS voices were

used since being able to rely on procedural speech is important to the overall design goals of flexibility and portability in a virtual human system.

Finally, the mutual gaze behavior is managed by custom Python scripts. Whenever contact is made with the virtual character’s eye region, the script determines how long the character will sustain mutual gaze for. If the contact is maintained, then eventually the script will direct the character to avert their eye-head gaze. We implemented two different versions of this behavior, to correspond to the experimental conditions. In the High Status condition, the wait time is based on a random number chosen from a Gaussian distribution with $M = 6.0s$ and $SD = 1.0s$. In the Low Status condition, the wait time is based on a random number chosen from a Gaussian distribution with $M = 4.0s$ and $SD = 0.5s$. These numbers were chosen to produce a range of times similar to those in the psychological literature and tweaked with adjustments based on early pilot studies. The other aspects of the gaze behavior were timed to provide the maximum realistic separation, based on the spectrum of parameters in the social eye gaze model.

2) *Procedure:* All experimental sessions were conducted individually. The experimenter greeted each participant, and explained the purpose of the study and the consent form. Once the participant signed the consent form, the participant completed a demographic data form and reviewed a set of instructions for the interaction. To interact with the system, the participant sat at a desk and wore the eye tracking headset. Then, a brief calibration routine was conducted to ensure the eye tracking was functioning properly. The participant then interacted with the character in a simulated hiring scenario, by asking the virtual character questions. On the desk was placed a sheet of potential questions, a microphone, and a mouse that could be clicked to allow the system to receive voice input.

After each condition, just as in the previous step, the avatar’s personality was assessed using the Ten-Item Personality Inventory (TIPI), and a rating of the character’s apparent social status on a Likert scale from 1 (Very Submissive) to 5 (Very Dominant). The participants were also asked to describe their main reason for the status rating. They were also asked to rate the character’s competence similarly. They then repeated the process with the other condition. The study took about 30 minutes to complete.

The participants were recruited from undergraduate and graduate students. The undergraduate participants received nominal credit from their course instructors for their participation. The remainder of the participants were volunteers. 35 participants were recruited for this study. Typically, eye tracking usability-style studies recruit between 6–30 participants, although oversampling is recommended due to the potential for technical issues [49]. However, one participant reported that they had difficulty paying attention to the task, and subsequent review of their eye tracking data showed they only looked at the animated character’s head region seven times during a session, which was well below the average number of such gazes across participants ($M=2440$). As perceiving the character’s activity was the major focus of the study, and a



Figure 3. The visual portrayal of the two versions of the character. (left) Brian; (right) Brad.

prerequisite for making judgments about it, this participant’s data was therefore excluded from the subsequent descriptions and analyses.

Of the 34 participants, 24 of them were aged 20–24. 18 were undergrad students. 11 indicated their gender was male, and 24 indicated their gender was female. Regarding their experience with video games, 19 of them chose responses that indicated they didn’t play games or did occasionally, but wouldn’t identify with the term gamer. 15 of them indicated they were either casual or hardcore gamers. The average number of times they interviewed someone else in a work setting previously was 2 ($SD = 3$), while the average number of times they were interviewed for work was 6 ($SD = 6$). Most had not interviewed someone else, but had been interviewed for work purposes. The observations of this study are therefore expected to generalize to young adults with some higher education, and minor work and video game experience.

The participants engaged in a scripted interaction with our virtual human character. They were given a list of questions they could ask in a simulated hiring interview. There were two versions of the character, with its gaze behavior adapted to represent a person of either high or low status. Participants interacted with both versions of the characters, in a within-subjects design. The variations of the character were assigned a differently colored shirt and name (“Brian” wore a dark gray shirt and “Brad” wore a blue shirt, as shown in Figure 3), and the study was counter-balanced by presenting the two conditions in a random order.

B. Results

The statistical analysis of the study, wherein participants interacted with an interactive character compared the effects of the two different conditions. Order of first condition did not have a statistically significant effect on the results; neither did participant gender and age range, which was the case even when assessing virtual characters and gender stereotyping [50].

We performed a paired-sample t -test comparing the social status rating results for the character in both conditions. There is a significant difference in the scores for the High Status condition ($M = 3.18$, $SD = 0.72$) and the Low Status condition

($M = 2.79$, $SD = 0.98$); $t(33) = 2.03$, $p = .046$. These results suggest that the differing gaze behavior really does have an effect on the perceived social status of interactive animated characters. Cohen’s $d = 0.45$, which is in the range of a Medium effect. This finding supports our hypothesis.

We performed a paired-sample t -test comparing the competence rating results for the character in both conditions. There was no significant difference in the scores for the High Status condition ($M = 3.85$, $SD = 0.93$) and the Low Status condition ($M = 3.74$, $SD = 1.11$); $t(33) = 2.03$, $p = .54$.

Paired-sample t -tests comparing the five personality factors from each condition found no significant difference. This implies that the use of the model did not alter the perceived personality of the character in this evaluation.

VI. DISCUSSION

A. The Parameterized Gaze Model

This study determined that a specific combination of gaze behaviors could be used to provide virtual humans with the ability to communicate social status. Implementing the social gaze model for these static videos allowed us to develop the necessary BML commands and timing values that could be used when the model is implemented within an interactive system, and so this effectively allowed us to prototype the required behavior for the interactive study in Section V. Finding empirical support for the effects of the social gaze model’s parameters also gave us confidence it would be effective and worthwhile evaluating further.

While participants mentioned receiving an emotional impression of the character in the open-ended comments, such as appearing anxious and uncertain, the assessment of social status was still effective. A few comments did make direct reference to the characters’ eye movements. The fact that they brought up emotional impressions rather than the gaze behavior actually helps show that social status is perceived without being overwhelming. Overall, participants reported that the characters’ roles made the biggest impression on their rating of social status, with NVB as the second-most frequent factor. This shows the importance of the narrative context for indicating to players why character actions occur.

B. Interactive Study

In this study, participants interacted with an ECA that used the social eye gaze model’s fundamental values to control its dynamic behavior, while adding in the parameter of length of eye contact through its ability to respond to mutual eye gaze. Since this study showed additional support for our hypothesis, it serves as additional validation of the model. Importantly, it shows support for the novel element added in this study: varying response according to mutual gaze.

This finding is in contrast to at least one previous finding, wherein a character implemented in Elckerlyc was used to mediate communication between two people [51]. There, participants saw each other as an avatar while conversing, and its gaze was delayed to see if it thereby mimicked high status behavior and portrayed dominance. The researchers weren’t

able to support that hypothesis. In that case, participants knew they were communicating to another person through the avatar, so it is possible that participants cognitively assessed it differently than a communicating character, or that the selected delay didn't operationalize into a signal of social status. Additionally, a study of gaze in interpersonal interviews found that gaze aversions had strong negative connotations such as perceptions of lowered credibility and attraction [52], whereas nearly constant levels of high gaze are not significantly more favorable than normal levels. Thus, it may be that gaze behavior sends signals of low status more readily than the inverse. This was supported in the comments about the Low Status condition in our study. Participants associated the NVB they saw with low status, while the high status behaviour was often viewed in neutral terms, and their comments switched to focus on the content in that condition.

The assessment of competence in this study was an attempt to determine if there was any connection between the gaze model and participations' perception of this characteristic, following prior findings that interviewees using normative to nearly constant gaze patterns were seen as more competent [52]. We did not find that, and indeed most participants commented on the chatbot's dialogue options and ability to respond to questions when asked about their reasoning. It seems that task performance is a priority in this context.

It is also worth noting that this study examined a different scenario than those in Section IV. Those videos examined a dramatic scenario taken from improv acting related to a firing incident. In this case, participants participated in a practice interview session. However, it seemed important to evaluate the gaze model parameters in an interactive setting. The hiring scenario is a more constrained and formulaic experience. Still, training simulators represent an important venue for ECAs [53], even in sophisticated contexts such as training psychologists to do interviews [54]. Still, both settings represent a range of experience that is important to equip virtual humans for.

VII. CONCLUSION

In this paper we propose a verbal-conceptual procedural model that provides a correspondence between the dramatic quality of social status and six parameters affecting eye gaze behavior. The studies documented in this paper provide evidence for the efficacy of its parameters, through interaction with both static and interactive representations of virtual humans.

These studies show the importance of gaze to the social interaction of interactive virtual agents, particularly during conversations. Consisting of coordinated motion between the torso, head, and eyes, gaze provides a broad range of abilities including general vision, task-monitoring, emotional display, and even subtler social signals in terms of the social status of participants. Status was identified as a valuable realistic and dramatic component of communication, worth investigating for its contribution to believability. As we create intelligent

virtual agents that are intended to be dramatically believable, the concern for meeting viewer expectations surrounding communication as well as providing the ability for subtler contrasting messages leads us to consider ways of improving the gaze capability of virtual avatars.

As mentioned, [55] proposes a means of guiding participants by creating status vacuums, based on the principles of interactive theater. In that case, the believability level of the agent is important for building user empathy and interest in the dramatic outcome. If social status is a reliable way of changing participant behavior, then this model of believable social gaze could help improve the portrayal of characters involved in similar kinds of scenarios, as well as contribute to a new way of directing player actions during dramatic scenes.

Social status is a powerful concept that has the potential to produce believable behavior in virtual humans. In improv theater, status refers to power differences in the relationship between two characters; additionally, the most relevant differences are seen in actions taken, rather than overt differences such as apparent wealth and rank. Since Johnstone believes that the most interesting scenes arise out of status changes, it is essential that it is possible to portray characters of different status and have means for changing status. Our social eye gaze model will enable virtual humans with a reliable way to communicate status and thus play a part in diverse dramatic situations. This could increase the number of contexts where they can be used, including more genres of digital games, especially those incorporating social simulations to provide lifelike activity.

ACKNOWLEDGMENT

We would like to thank our reviewers for their very helpful feedback.

REFERENCES

- [1] M. Wooldridge and N. R. Jennings, "Intelligent agents: Theory and practice," *Knowledge Engineering Review*, vol. 10, no. 2, pp. 115–152, 1995.
- [2] A. S. Rao and M. P. Georgeff, "Modeling rational agents within a bdi-architecture," *Readings in agents*, pp. 317–328, 1997.
- [3] M. Mateas and A. Stern, *Procedural authorship: A case-study of the interactive drama Faade*. IT University of Copenhagen, Dec 2005.
- [4] W. R. Swartout, J. Gratch, R. W. Hill Jr, E. Hovy, S. Marsella, J. Rickel, and D. Traum, "Toward virtual humans," *AI Magazine*, vol. 27, no. 2, p. 96, 2006.
- [5] R. Aylett, S. Louchart, J. Dias, A. Paiva, and M. Vala, *Fearnot!: an experiment in emergent narrative*, ser. Lecture Notes in Computer Science. Springer-Verlag, 2005, vol. 3661/2005, pp. 305–316.
- [6] J. Cassell, H. Vilhjmsson, and T. Bickmore, *BEAT: the Behavior Expression Animation Toolkit*. ACM, 2001, pp. 477–486.
- [7] A. Vinciarelli, M. Pantic, H. Bourlard, and A. Pentland, *Social signals, their function, and automatic analysis: a survey*, ser. ICMI 08. ACM, 2008, pp. 61–68, aCM ID: 1452405.
- [8] K. Johnstone, *Impro: Improvisation and the Theatre*. Methuen Drama, 1981.
- [9] V. Spolin, *Improvisation for the theater: A handbook of teaching and directing techniques*. Northwestern University Press, 1983.
- [10] A. Kendon, "Some functions of gaze direction in social interaction," *Acta Psychologica*, vol. 26, no. 1, pp. 22–63, 1967.
- [11] M. Shimada, T. Minato, S. Itakura, and H. Ishiguro, "Evaluation of Android Using Unconscious Recognition," in *2006 6th IEEE-RAS International Conference on Humanoid Robots*. IEEE, dec 2006, pp. 157–162. [Online]. Available: <http://ieeexplore.ieee.org/document/4115595/>

- [12] M. Mori, "The uncanny valley," *Energy*, vol. 7, no. 4, pp. 33–35, 1970.
- [13] Team Bondi, "L.A. Noire," 2011.
- [14] Firaxis Games, "Civilization VI," 2016.
- [15] Maxis, "Spore," 2008.
- [16] C. Peters, C. Pelachaud, E. Bevacqua, M. Mancini, and I. Poggi, "A Model of Attention and Interest Using Gaze Behavior," in *Intelligent Virtual Agents*, ser. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, Sep. 2005, pp. 229–240. [Online]. Available: https://link.springer.com.proxy.lib.sfu.ca/chapter/10.1007/11550617_20
- [17] M. Argyle and M. Cook, *Gaze and mutual gaze*. Cambridge University Press, 1976.
- [18] A. Fukayama, T. Ohno, N. Mukawa, M. Sawaki, and N. Hagita, *Messages Embedded in Gaze of Interface Agents — Impression Management with Agent's Gaze*, ser. CHI '02. New York, NY, USA: ACM, 2002, pp. 41–48. [Online]. Available: <http://doi.acm.org/10.1145/503376.503385>
- [19] M. Argyle and J. Dean, "Eye contact, distance, and affiliation," *Sociometry*, vol. 28, pp. 289–304, 1965.
- [20] J. Bailenson, J. Blascovich, A. Beall, and J. Loomis, "Interpersonal distance in immersive virtual environments," *Personality and Social Psychology Bulletin*, vol. 29, no. 7, p. 819, 2003.
- [21] N. Wang and J. Gratch, *Dont just stare at me!* ACM, 2010, pp. 1241–1250. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1753513>
- [22] S. Andrist, T. Pejisa, B. Mutlu, and M. Gleicher, *Designing effective gaze mechanisms for virtual agents*. ACM, 2012, pp. 705–714. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2207777>
- [23] A. Colburn, M. F. Cohen, and S. Drucker, "The role of eye gaze in avatar mediated conversational interfaces," *Sketches and Applications, Siggraph00*, 2000. [Online]. Available: <https://pdfs.semanticscholar.org/2e9a/7e7654d7a1f9c92598f2cc8711dc19add3c.pdf>
- [24] M. Garau, M. Slater, S. Bee, and M. A. Sasse, *The impact of eye gaze on communication using humanoid avatars*. ACM, 2001, pp. 309–316. [Online]. Available: <http://dl.acm.org/citation.cfm?id=365121>
- [25] R. B. Queiroz, L. M. Barros, and S. R. Musse, "Providing expressive gaze to virtual animated characters in interactive applications," *Computers in Entertainment (CIE)*, vol. 6, no. 3, p. 41, 2008.
- [26] B. Lance and S. C. Marsella, "Emotionally expressive head and body movement during gaze shifts," in *Intelligent Virtual Agents*, C. Pelachaud, J.-C. Martin, E. André, G. Chollet, K. Karpouzis, and D. Pelé, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 72–85.
- [27] B. J. Lance and S. C. Marsella, "A model of gaze for the purpose of emotional expression in virtual embodied agents," in *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 1*, ser. AAMAS '08. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems, 2008, pp. 199–206. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1402383.1402415>
- [28] C. Busso, Z. Deng, U. Neumann, and S. Narayanan, *Learning expressive human-like head motion sequences from speech*. Springer, 2008, pp. 113–131. [Online]. Available: https://link.springer.com/content/pdf/10.1007/978-1-84628-907-1_6.pdf
- [29] C. Cig, Z. Kasap, A. Egges, and N. Magnenat-Thalmann, *Realistic emotional gaze and head behavior generation based on arousal and dominance factors*, ser. Lecture Notes in Computer Science. Springer-Verlag, 2010, pp. 278–289. [Online]. Available: <http://link.springer.com/content/pdf/10.1007/978-3-642-16958-8.pdf#page=289>
- [30] Z. Kasap, M. B. Moussa, P. Chaudhuri, and N. Magnenat-Thalmann, "Making them remember emotional virtual characters with memory," *IEEE Computer Graphics and Applications*, vol. 29, no. 2, pp. 20–29, 2009.
- [31] A. Mehrabian, "Analysis of the big-five personality factors in terms of the pad temperament model," *Australian Journal of Psychology*, vol. 48, no. 2, pp. 86–92, 1996.
- [32] S. Kopp, B. Krenn, S. Marsella, A. Marshall, C. Pelachaud, H. Pirker, K. Thirsson, and H. Vilhjlmsson, *Towards a common framework for multimodal generation: The behavior markup language*, ser. LNCS (LNAI). Springer, 2006, vol. 4133, pp. 205–217.
- [33] H. Van Welbergen, D. Reidsma, Z. Ruttkay, and J. Zwiers, "Elckerlyc-a bml realizer for continuous, multimodal interaction with a virtual human," *Journal on Multimodal User Interfaces*, vol. 3, no. 4, pp. 271–284, 2010.
- [34] B. Harger, *Workshop on Improvisational Acting*, 2007.
- [35] —, *Entertaining AI: Using Rules from Improvisational Acting to Create Unscripted Emotional Impact*, 2008.
- [36] M. Seif El-Nasr and H. Wei, *Exploring Non-verbal Behavior Models for Believable Characters*. Springer, Nov 2008, vol. 5334, pp. 71–82.
- [37] R. Sun, *Introduction to computational cognitive modeling*. Cambridge University Press, 2008, pp. 3–19. [Online]. Available: http://cindy.informatik.uni-bremen.de/cosy/teaching/CM_2011/intro/sun_08.pdf
- [38] S. P. Lee, J. B. Badler, and N. I. Badler, *Eyes alive*. ACM Press, 2002, vol. 21, pp. 637–644. [Online]. Available: <http://portal.acm.org/citation.cfm?doid=566570.566629>
- [39] B. J. Lance and S. C. Marsella, "Glances, glares, and glowering: how should a virtual human express emotion through gaze?" *Autonomous Agents and Multi-Agent Systems*, vol. 20, no. 1, pp. 50–69, May 2009.
- [40] M. J. Crump, J. V. McDonnell, and T. M. Gureckis, "Evaluating amazons mechanical turk as a tool for experimental behavioral research," *PLoS one*, vol. 8, no. 3, p. e57410, 2013.
- [41] M. Thiebaux, S. Marsella, A. N. Marshall, and M. Kallmann, "Smartbody: Behavior realization for embodied conversational agents," in *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 1*, ser. AAMAS '08. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems, 2008, pp. 151–158. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1402383.1402409>
- [42] M. Nixon, "Enhancing believability: Evaluating the application of de-larties aesthetic system to the design of virtual humans," Master's thesis, Simon Fraser University, Oct 2009.
- [43] N. Ambady and R. Rosenthal, "Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis," *Psychological Bulletin*, vol. 111, no. 2, pp. 256–274, 1992.
- [44] S. D. Gosling, P. J. Rentfrow, and W. B. Swann, "A very brief measure of the big-five personality domains," *Journal of Research in Personality*, vol. 37, no. 6, pp. 504–528, 2003.
- [45] M. Kassner, W. Patera, and A. Bulling, *Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction*. ACM, 2014, pp. 1151–1160. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2641695>
- [46] S. S. Feldman, O. N. Yalcin, and S. DiPaola, "Engagement with artificial intelligence through natural interaction models," in *Proceedings of the Conference on Electronic Visualisation and the Arts*, ser. EVA '17. Swindon, UK: BCS Learning & Development Ltd., 2017, pp. 296–303. [Online]. Available: <https://doi.org/10.14236/ewic/EVA2017.60>
- [47] O. N. Yalcin and S. DiPaola, "Virtual AI agent for Adults with Mental Health Issues," in *Stanford Workshop on VR and Behavioral Change, Poster*, Palo Alto, California, 2017.
- [48] A. Leuski, R. Patel, D. Traum, and B. Kennedy, *Building effective question answering characters*. Association for Computational Linguistics, 2009, pp. 18–27. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1654600>
- [49] J. H. Goldberg and A. M. Wichansky, "Eye tracking in usability evaluation: a practitioners guide," in *The Mind's Eyes: Cognitive and Applied Aspects of Eye Movements*, J. Hyn, R. Radach, and H. Deubel, Eds. Oxford: Elsevier Science, 2002.
- [50] J. Fox and J. N. Bailenson, "Virtual virgins and vamps: The effects of exposure to female characters sexualized appearance and gaze in an immersive virtual environment," *Sex roles*, vol. 61, no. 3–4, pp. 147–157, 2009.
- [51] B. Gennep, "In the eye of the wizard: effects of (mutual) gaze on an avatar mediated conversation," Ph.D. dissertation, University of Twente, 2013. [Online]. Available: <http://essay.utwente.nl/64986/>
- [52] J. K. Burgoon, D. A. Coker, and R. A. Coker, "Communicative effects of gaze behavior," *Human Communication Research*, vol. 12, no. 4, pp. 495–524, 1986.
- [53] J. Rickel, *Intelligent virtual agents for education and training: Opportunities and challenges*. Springer, 2001, pp. 15–22. [Online]. Available: <http://link.springer.com/content/pdf/10.1007/3-540-44812-8.pdf#page=24>
- [54] P. Kenny, A. A. Rizzo, T. D. Parsons, J. Gratch, and W. Swartout, "A virtual human agent for training novice therapist clinical interviewing skills," *Annual Review of CyberTherapy and Telemedicine*, vol. 5, pp. 77–83, 2007.
- [55] J. Zhu, K. Ingraham, and J. Moshell, *Back-Leading through Character Status in Interactive Storytelling*, ser. Lecture Notes in Computer Science. Springer, 2011, vol. 7069/2011, pp. 31–36.