How Anxiety State and Acceptance of an Embodied Agent Affect User Gaze Patterns*

Nermin Shaltout[1,2,*,†], Diego Vilela Monteiro[3], Monica Perusquía-Hernández[2], Jason Orlosky[1,4] and Kiyoshi Kiyokawa[2]

[1]Osaka University, 1-32 Machikaneyama, Toyonaka, Osaka 560-0043, Japan
[2]Nara Institute of Science and Technology, 8916-5, Takayama, Ikoma, Nara 630-0192, Japan
[3]École Supérieure d'Informatique Electronique Automatique, 38 Rue des Docteurs Calmette et Guérin, 53000 Laval, France
[4]Augusta University, School of Computer and Cyber Sciences 100 Grace Hopper Ln, Augusta, GA 30901, USA

Abstract
In virtual reality (VR), the interactions of users with embodied agents when the users are anxious or when they do not accept an agent are not yet completely understood. Gaze can be indicative of the user's anxiety and acceptability of an embodied agent. An agent's expressions or actions can, in turn, be used to accommodate the user's anxiety. Previous work on social anxiety disorder (SAD) found evidence of avoidance or hyper-vigilant gaze patterns in relation to agents or people the participants were gazing at. Thus, we investigated if there are specific gaze patterns for normal individuals experiencing anxiety in the moment when gazing at an embodied agent. We focused mostly on avoidant gaze patterns. Based on evidence of gaze patterns in SAD and autism, we designed an experiment where normative individuals interact with an agent showing a neutral, happy and angry expressions. We aim to examine if normal anxious participants have similar gaze patterns or avoidance patterns to those with SAD. We also investigated if the user's acceptability or preference of the virtual agent's display of emotions had an effect on the avoidance via eye gaze. In particular, we investigated the user's eye patterns in relations to the agent's eyes, face or body to see if there were similarities to people with SAD. Using correlation analysis, we found a significant positive correlation between the acceptability of the participant to the virtual agent's expression and their fixation on the agent's eyes. We also found a significant correlation between fixations on the agent's body and how anxious the participant was at the experiment's start. Later, these results can be used to find a link between acceptability, anxiety and SAD.

Keywords
Virtual Reality, Embodied Agent, Eye Gaze, Anxiety

1. Introduction

In the field of virtual reality (VR), embodied agents are commonplace as non-player characters (NPCs) or as other users (avatars) in-game. Thus, determining how individuals react to virtual agents is an important topic in the field [1]. The adaptation of embodied agent or avatar facial expressions can influence user behavior [2]. In particular, for those who might use the agents for learning [3], social support, or feedback [4]. Previous studies assessed the gaze patterns of individuals in social situations to understand psychological and emotional patterns. These studies are used to better understand and train people with disorders such as high social anxiety and were usually conducted using still photographs [5]. Individuals with Social Anxiety Disorder (SAD) react differently to facial displays of emotion. This happens in VR too. It happens independently of the avatar fidelity [6]. Little is known of how normative individuals' anxiety (aka those without SAD) affects gaze behaviour with respect to facial displays of emotion. The exposure of individuals to virtual situations has also risen with the rise of new platforms like VR, and increased by the advent of covid-19. Studying the effects of anxiety on virtual embodied agents, is thus important, for SAD as well as for anxious normative individuals.

VR offers the possibility of presenting dynamic facial stimuli with a wealth of parameters, leading to detailed descriptions of the facial movements required to convey a socio-affective message accurately [7]. Furthermore, the advent of biosensors allows real-time reproduction of facial expressions from other users [8]. Moreover, with a recent increased interest of the general public in the metaverse after the 2019 pandemic. We believe it is timely to study the effects of VR agents' facial expressions on the user gaze.

Thus, this study aims to explore different gaze parameters and their effectiveness to determine how comfortable the user is with the agent as it presents different facial

APMAR'22: Asia-Pacific Workshop on Mixed and Augmented Reality, Dec. 02-03, 2022, Yokohama, Japan
*Corresponding author.
nermeena@gmail.com (N. Shaltout); diego.vilelamonteiro@esie.fr (D. V. Monteiro); m.perusquia@is.naist.jp (M. Perusquía-Hernández); jorlosky@augusta.edu (J. Orlosky); kiyo@is.naist.jp (K. Kiyokawa)
0000-0002-1570-3652 (D. V. Monteiro); 0000-0002-0486-1743 (M. Perusquía-Hernández); 0000-0002-0538-6630 (J. Orlosky); 0000-0003-2260-1707 (K. Kiyokawa)

expressions; which might be an alternate method for measuring the reaction towards VR agents. The design of the study was inspired by previous works conducted on individuals with SAD. Individuals with SAD usually do not deal well with emotions presented on the face. They tend to avoid gaze when faced with emotional people or their representations. The avoidance might increase with negative emotions. They especially avoid looking at the eyes of individuals displaying emotions [9]. Thus, we analyzed the effect of user's anxiety and acceptance when confronted with an embodied agent showing different emotions on the user's eye gaze patterns.

We hypothesize that we can use gaze location on the agent to measure the degree of comfort towards an agent's facial expression or degree of user anxiety. To this aim, eye gaze was measured using the VIVE Pro Eye tracker while participants looked at a VR agent with varying expressions. The main contributions of this paper are:

- Analyzing the correlation between general anxiety of a normative user and their gaze patterns on the embodied agent.
- Analyzing the correlation between the user acceptance to the embodied agent's emotional display and their gaze patterns on the embodied agent.
- Comparing the findings to those found in SAD using similar studies.

2. Prior Work and Hypotheses

2.1. Gaze Analysis Studies Related to Social Anxiety

Eye metrics are promising tools to assess attitudes towards virtual agents. The main inspiration for this study stems from gaze analysis of individuals with HSA (High Social Anxiety) towards the facial expressions of other individuals in social situations. Previous gaze studies showed that individuals with HSA averted their gaze when shown photos of individuals expressing positive or negative emotions [5, 10]. In such studies, static photos of people presenting happy, sad, and neutral facial expressions were commonly shown while gaze directions and fixations were measured. Thus, we adapted our study to find a relation between gaze direction on the embodied agent displaying emotions and the user's acceptance of the emotional display.

Wieckowski et al. explored variability in bias toward social stimuli in the form of vigilant attention and avoidant attention using eye gaze techniques instead of a traditional probe task technique often used to study attention bias in anxious youth with clinical SAD [11]. The visual dot probe task involves allowing the users to select between two agent pairs (e.g., angry/neutral, happy/neutral) using their eye gaze. Participants show both avoidance and hyper-vigilance according to the age group, the agent display, and the passage of time during the trial. The bias is measured with the duration of fixations towards angry faces and towards more pleasant faces such as neutral or happy faces. The fixation duration of neutral faces is subtracted from angry faces to create a negative or positive bias. Based on the above studies we hypothesize that individuals with SAD might have two factors that affect their gaze patterns. The level of anxiety of the person when they are looking at the face, and whether or not the individuals accept the facial expression of the embodied agent. We would like to also observe if this affects normative individuals and if it mimics those with SAD.

2.2. Hypotheses

Our hypotheses are as follows.

H1 The agent's facial expressions have an effect on the participant's self-reported arousal and valence.

H2 The user's acceptance or preference of the agent's emotional display can be observed in the eye fixation patterns on the agent.

H3 The overall anxiety state of the participants could alter the eye fixation patterns on the agent.

For **H1**, the Self Assessment Manniken (SAM) was used to assess if the facial expressions had an effect on the participant's affect, to see if the avatar's expressions affected the user. Regarding **H2**, while participants were answering questionnaires in 3.6 about the different emotional display of the agent, not all of them accepted the agent's emotional display in the same manner. E.g. while some people highly disliked the happy face, other people were comfortable with it. We assessed if there is a pattern between acceptability of the emotional display and the number of fixations on the agent. For **H3**, we assessed if there is a relation between the anxiety of the participant and their fixation behaviors on the agent's different body parts. We expected the more anxious participants to be avoidant of the agent's face and eyes. The anxiety state in this case is the user's default state before and during the experiment.

3. Experiment

3.1. Participants

A total of 21 student volunteers in their early twenties participated in the study; 10 Japanese, 1 Kenyan, 1 German, 2 Nepali, 1 Colombian, 4 Chinese, 1 Thai, and 1 Malaysian. No participant tried our system before. The participants were asked to wear glasses if their vision was not good. Sources of error were accounted for by removing three participants in which there was missing data (e.g., the VIVE Pro Eye tracking was disabled accidentally for one of the faces). Three participants that were extremely fatigued were excluded using a fatigue score in the pre-questionnare. After the exclusions, the number of participants was 15. The experiment was approved by the ethics committee of our institution.

3.2. Experiment Design

We tested the participant's eye gaze patterns when presented with different facial expressions from a humanoid agent in VR. There were three conditions corresponding to three facial displays expressed by the virtual agent: a happy, a sad, and an angry facial expression. The conditions were presented in a within subjects design, i.e., each participant saw all three faces. We chose to enable the agent to have only facial expressions; to avoid confounding factors caused by other agent behaviors.

3.3. Procedure

The participants were exposed to each of the agent's facial expression one minute at a time. There were three runs total, one for each facial expression. The participants were seated in front of the agent as to be the same height as the agent and faced the agent head-on without an angle. Before every run, the agent was adjusted to be the same height as the participant. The participants were seated throughout the course of the experiment and encouraged to use only gaze and head movements. A full agent was used so the participants could freely choose whether or not to gaze at the agent's face, body, or outside of the agent completely. The participants answered questionnaires pre- and post-experiment and after each avatar was displayed. Details are mentioned in the measurements section.

3.4. Stimuli

The agent was designed using Ready Player Me [12], which is a tool that converts a photograph of a person into an avatar with similar facial features. It is used for players to make agents of themselves in-game. It is currently most popular on platforms such as VR chatting programs. Ready Player Me is also equipped with the ability to map the user's emotions onto the agent via eye and facial tracking. The readyplayer.me avatar was based on FACs and is usually embodied by people in VR but we controlled it using the animation module of Unity to introduce more control over the experiment

The possibility that there is an uncanny valley in the virtual agent's appearance is higher with agents that are hyper realistic [13]. Thus we used a semi realistic avatar.

Figure 1: A diagram showing the resulting avatar (right) created when using an average face (left) on the Ready Player Me avatar creator (https://bit.ly/34C2G).

We used an average Asian face to accommodate the majority Asian demographic involved in the experiment. Figure 1 shows the resulting avatar when inputting an average Asian face to the Ready Player Me interface. The Ready Player Me was used because its low-poly characteristics make it more likely to be used by people in current virtual chats and metaverse settings.

Though readyplayer.me characters are usually used as avatars in VRchat, we use it in this case to animate the agent, as if it's an example user in VR. To animate the happy, angry, and neutral emotions, the Facial Action Coding System (FACS) was used [14]. The FACS presents action units (AUs) used for coding facial movements without making inferences about the underlying emotions. It is a popular tool in emotion studies to either create faces with a certain expression or to interpret a facial expression. The FACS is now incorporated in most VR chat avatars to enable the avatars to express emotions by encoding certain AU movements. Ready Player Me avatars come equipped with most of the values available in the FACS. We focused on prototypical AUs according to the Basic Emotion Theory [15] to animate the avatar, together with guidelines described in Farnsworth's visual FACS guide [16]. For instance, to animate a happy face, we used AU 6 (cheek raiser, Fig. 2) and AU 12 (lip corner puller) with values of 1.0. The agent's default face with some minor adjustments was used to represent the neutral face, Because the Ready Player Me avatars are designed to look slightly happy by default, we adjusted brow lowered AU 4 and lip corner depressor AU 15 to bring back the avatar to its normal state. The facial expressions are animated using blend shapes. There are three separate faces shown to the same participant. We refer to them as three different trials with questionnaires in between. All facial expressions start from the neutral expression. It took one second for the facial animation to reach their maximum intensity. Then the avatar's expression remains for the duration of the trial. The animation stays constant for one minute per trial. We measure the participant's reaction within one minute for the purpose of measuring fixations. We do not consider the initial animation to have a detrimental effect on the gaze patterns or number of fixations thus we did not account for the baseline when observing gaze patterns.

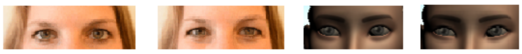

Figure 2: A diagram showing action unit 6, cheek raiser (left two images, permission was taken from imotions to use the image). The rightmost two images show the AU6 applied on the Ready Player Me avatar, used in the experiment as an agent (https://bit.ly/3bKN9a).

To reduce the chances of a participant experiencing the uncanny valley, the avatar's blinking was animated. The frequency of the blinks was randomized, between 0.5 and 4 seconds. The agent's gaze was fixed during the run of this experiment. The agent was also given a breathing animation using Maximo to give it a more realistic feel [17]. No other interactions other than varying facial expressions were added to the agent. In this experiment we focused more on the relationship between user's anxiety and user's gaze patterns rather than design a complex interaction system. Thus a simple design was used. Future work will feature a more interactive avatar.

3.5. Calibration

To ensure that eye and facial feature tracking worked correctly, we used a cube display to calibrate and confirm participant's gaze prior the experiment's start. The agent's height was adjusted to be the same as the participant's height in every trial. The agent's face was only revealed once the VIVE Pro Eye was calibrated, as shown in Fig. 3. We calibrated each participant's eye gaze pre-experiment using the VIVE's internal calibration software before running the experiment. The participant was then positioned to see the front view of the agent. However, the participant was free to move his/her head or gaze freely and look either at the avatar's face or body for the duration of the one-minute trial. No background objects were visible so that the participant would focus on the agent. Participants were seated during the experiment and received no instructions regarding as to where to look when facing the agent and were left to interact naturally with the agent using eye movements and head movements once the experiment started. The participant's location did not change.

3.6. Measurements

Eye Metrics We implemented a gaze ray to determine the intersection of the participant's gaze with the avatar's face. To determine the number of collisions between the gaze ray and parts of the face, we added colliders to the face, eyes, and body (Fig. 5). When the ray did not collide with the avatar, it was recorded as 'other.' We defined that a constant gaze on the same body part for

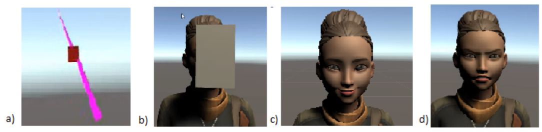

Figure 3: (a) Cube used for calibrating the eye gaze pre-experiment (b) height adjusting the participant while obstructing the avatar's face (c), (d) examples of the neutral face and angry faces used for the experiment. The gaze ray is only shown for illustrative purposes but is omitted in the actual run of the experiment. The Action Units (AU)s for both happy and angry faces were amped to 1 for both angry and happy faces to get the maximum effect.

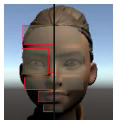

Figure 4: Figure showing the face mask (highlighted on the left) used to detect collisions/fixations

150 ms constituted a "fixation." The number of fixations on each body part was counted within one minute of the participant looking at the avatar. The fixations were collected for each facial expression; Neutral, Angry, and Happy. The process was repeated per participant. By analyzing the number of fixations on each body part, we expected to find a pattern related to the participant's current state, the participant's preference of the face, and patterns related to the facial expression that the avatar displayed to the participant.

Additionally, we added a face mask not visible to the user to roughly measure the number of fixations on the face. We counted the number of fixations surpassing 150 ms on the colliders added to this face mask, and divided the number of fixations into upper face region and lower face region. Any collider above the lower eye was considered upper face while colliders below the face were considered lower face. If the upper face collisions are higher in the results, then the user accepts the avatar, otherwise if the lower face collisions are higher the user rejects the avatar according to source.

Questionnaires The participants answered questionnaires at different points during the experiment. Sometimes a questionnaire like SAM was repeated more than once so were questionnaires about the acceptability of the avatar. To identify each of which point of the experiment the participant answered the questionnaires we assigned codes as follows:

B : **Before** seeing any avatar.

AH : **After** seeing the **happy** facial expression agent.

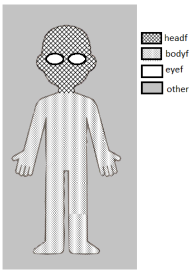

Figure 5: Figure showing how the number of fixations is measured using colliders placed on the avatar. (a) headf: represents the number of times the participant fixated on the avatar's head. (b)bodyf: the number of times the participant fixated on the avatar's body, (c) eyef: the number of times the participant fixated on the avatar's eyes, and (d) other: the number of fixations outside of the avatar. Collisions on the avatar were detected using a convex collider respectively.

AA : **After** seeing the **angry** facial expression agent.

AN : **After** seeing the **neutral** facial expression agent.

All three avatars were shown to each participant in a counterbalanced order. The following permutations were used: (AA, AH, AN); (AA, AN, AH); (AH, AN, AA); (AH, AN, AA); (AN, AA, AH); (AN, AH, AA)

Pre-questionnaire Before the experiments, participants reported their demographics and their current fatigue and anxiety. They also reported their general anxiety level on a 9-point Likert scale.

Self Assessment Manikin (SAM) A SAM questionnaire with 9 point likert scale is used to measure valence,arousal and dominance. (Fig. 6) shows a sample of the valence and arousal questionnaires used. SAM was given to the participant before the experiment and after every face. SAM was used to measure how the avatar affected the participant. The SAM questionnaire was presented prior to and after looking at the avatar.

Acceptability Questionnaire This was a more detailed questionnaire about the avatar. The questionnaires were given a code AH, AA, or AN, after each different avatar with different facial expression. It asked how people felt about the avatar. It consisted of several questions answered on a 9-point Likert scale. The scores are on a scale from 1 to 9 to match the same format of the SAM. From this questionnaire, we only used two questions for further analyses:

Q1 : I felt comforted by the avatar.

Q2 : I felt disturbed by the avatar respectively.

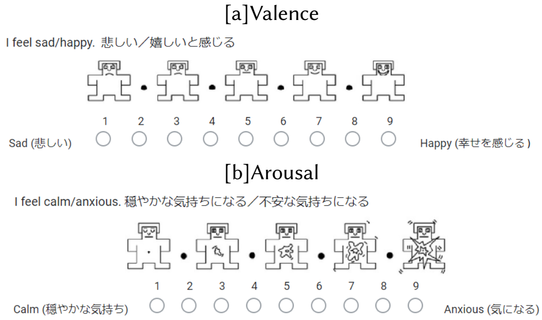

Figure 6: The [a]Valence and [b]Arousal questionnaire from SAM.

The questionnaire about the avatar was used to measure the participant's **acceptance** aka contextual comfort of the virtual agent's expression after each trial. The questionnaires were answered thrice per participant after each face. From the answers of the acceptability questionnaire, we created a variable known as **acceptability** by taking as follows:

$$Acceptability : Score(Q1) - Score(Q2)$$

A positive or zero value indicated that the agent was accepted, while a negative value indicated that the agent was rejected by the user. We calculated the acceptability per face per participant. This score was then used to find a correlation between the acceptability of the agent. Each facial expression was compared to the participant's eye on fixations different locations of the embodied agent (head, body, eye, etc).

The acceptability questionnaire was conducted three times per participant, after the participant viewed every expression of the agent for one minute. For instance, after the participant views the happy expression for one minute, the experiment stops and the participant answers the SAM and the acceptability questionnaires. This procedure is then repeated with the other two expressions. The questionnaires are not validated which might be a limitation of the study.

4. Analysis and Results

Despite being given no instructions on where to direct their gaze, the majority of participants gazed at the agent's face or body. However, the gaze patterns changed according to the facial expression shown, the participant's acceptability score, and the anxiety score, as detailed below. Deviations away from the face changed according to the acceptability and anxiety scores.

SPSS was used to analyze the data. To test **H1**, we measured if the agent's facial expression had an effect on the participant's affect as follows. The SAM questionnaire was taken four times and symbolized with experiment codes described in the Questionnaire portion of Sec. 3.6 as follows: Before the experiment (B), after the neutral agent(AN), after the happy agent (AH), after the angry agent (AA) for each participant.

A Shapiro-Wilk test showed a significant departure from normality W(98) = 0.94 , $p < 0.01$; W(98) = 0.94, $p < 0.01$; W(98) = 0.95, $p < 0.01$ for valence, arousal and dominance respectively. We thus ran the Kruskal-Wallis as a non-parametric test to compare if there were significant differences between conditions B, AH, AA, and AN. A Kruskal-Wallis rank-sum test carried out on arousal valence and dominance showed that there was a statistical significance in valence ($\chi^2(3) = 21.48$, $p < 0.001$) and arousal ($\chi^2(3) = 8.867$, $p < 0.05$) between the different embodied agent expressions when compared to the baseline. The mean rank valence scores of 41.07, 45.16 , 28.67, and 62.5 for B, AN, AH, and AA, respectively. For arousal, the mean rank scores were 39.98, 41.61, 57.6, 37.84 for B, AN, AH, and AA, respectively. The results for dominance showed no statistical significance.

Pairwise Wilcoxon rank sum tests with Benjamini-Hochberg (BH) p-value adjustment were carried out for the SAM valence and arousal scores since they showed significance. Experiment codes (B, AN, AH, AA) were used to represent the experiment stage as detailed in 3.6. The following score pairs were compared to each other: (B, AN), (B, AH), (B, AA) representing the comparison between the baseline state and the state of the participant after viewing agents of different emotions. This is to test if the different emotional agents had a significant effect on the affect of the participant. The difference between the scores of the following face pairs were compared to each other: (AN, AH), (AN, AA), (AH, AA).

The significant results for valence and arousal are summarized using starred brackets in Fig. 7. Significant results were as follows: The valence was larger after introducing the happy agent (AH), compared to the baseline (B): ($p < 0.05$, Cohen's $d = 0.25$). The valence was conversely lower than the baseline after introducing the angry agent (AA): ($p < 0.01$, Cohen's $d = 0.25$). There was no significant differences in valence when comparing the score of the baseline condition to introducing the neutral agent (AH). The valence score was also lower after viewing the angry face as compared to after viewing the neutral face ($p < 0.05$, Cohen's $d = 0.25$). Conversely, the valence score was higher after the happy face as compared to the neutral face ($p < 0.05$, Cohen's $d = 0.25$). The difference in valency values between the AH and AA pair was significant ($p < 0.01$, Cohen's $d = 0.25$).

The results for arousal were as follows: the difference between arousal values for the following score pairs with the following experiment codes showing statistical significance: (B, AH) with ($p < 0.05$, Cohen's $d = 0.25$); (AA, AH) with ($p < 0.01$, Cohen's $d = 0.25$) and (AN, AH) with ($p < 0.01$, Cohen's $d = 0.25$); $N = 15$ for all the test concerning valence and arousal

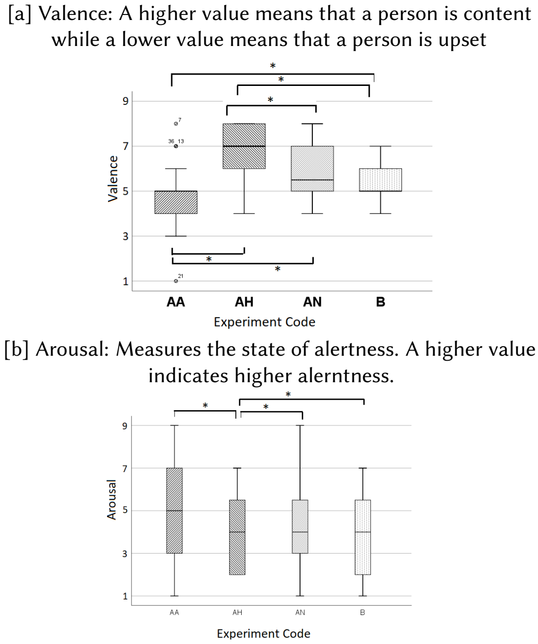

Figure 7: Boxplot showing [a] Valence and [b] Arousal values entered by participants after viewing each facial expression (AA, AH, AN) compared to the Baseline, B. The starred brackets show statistically significant pairings using the Wilcoxon rank sum tests.

We tested **H2** and **H3** and see if anxiety and acceptability affect where the participant looks most at the avatar and whether they are avoiding the agent's face or eyes as a result. We compared those results to previous findings for socially anxious individuals. Kendall's τb correlation was used to run the correlation tests since the sample size is small, and the normality tests showed that the sample did not follow a normal distribution as before.

To test **H2**, when running the tests on all the participants, we found no correlation between the acceptability score and number of fixations on all portions of the agent. We ran another Kendall's τb correlation using only anx-

Table 1
Correlation values between acceptability score and fixations on different portions of the agent for N = 13 (anxious participants) (*: $p < 0.05$, **: $p < 0.01$),

	headf	bodyf	eyesf	other
bodyf	.04			
eyesf	.51*	−.22		
other	.59**	.21	.16	
Acceptability	.31	.03	**.51***	.22

ious participants. The results are summarized in Table 1. A positive correlation between the participant's acceptability score and the number of fixations on the eye were found which was statistically significant ($\tau b = 0.510$, $p < 0.05$). The acceptability score was standardized between -1 and 1.

To investigate **H3**, a Kendall's τb correlation was run to determine the relationship between the participant's anxiety score and amount of fixations on a certain portion of the avatar, as shown in Table 2, with $N = 45$, regardless of the face used. There was a strong, positive correlation between the participant's anxiety score of and the number of fixations on the agent body per minute, which was statistically significant ($\tau b = 0.30$, $p < 0.001$). Additionally, there was also a strong negative correlation between the participant's ($\tau b = -0.22$, $p < 0.01$). The results are summarized in Table 2.

A Kendall's τ-b correlation was run to determine the relationship between the anxiety score of the participant and the number of fixations in a certain section of the agent, among 15 participants, per facial expressions. We observed if there are any patterns for specific facial expressions. We found no significant correlations between the anxiety score and the body fixations when presenting the participants with the happy face. However we found a strong correlation between the number of body fixations (bodyf) and the anxiety score when participants were presented with the neutral face, which was statistically significant ($\tau b = 0.411$, $p < 0.05$). There was also a strong correlation between the number of fixations detected outside the agent and the anxiety score when participants viewed the angry face, there was also a significant correlation between the number of fixations outside the agent when participants were viewing the angry face, which was statistically significant ($\tau b = 0.420$, $p < 0.05$). The values were Benjamini-Hochberg corrected. The results are summarized in the scatter plot detailed in Fig. 8. We ran two additional Kendall's τ-b correlation tests to see if there was a correlation between anxiety score and if the participant gazed at the upper or lower part of the avatar's face, first regardless of the face with $N = 45$, then per facial expression (angry, neutral, happy) with $N = 15$. For the results, regardless of the face presented there was strong negative correlation between the anxiety score and the upper face with statistical significance ($\tau b = -0.263$, $p < 0.05$). For the result per facial expression, only the happy expression showed a strong negative correlation between the anxiety and the upper face, with statistical significance ($\tau b = -0.408$, $p < 0.05$).

[a] Angry Expression

[b] Neutral Expression

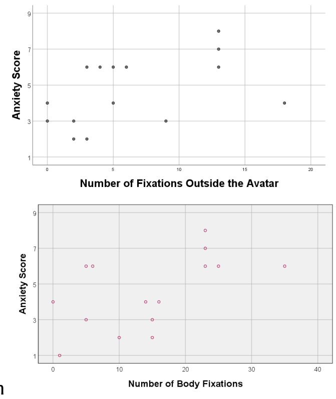

Figure 8: Scatter plot illustrating the correlation between the anxiety score of the participant and the number of fixations detected on outside of the agent's face, when observing the avatar with the (a) **angry expression** as opposed to number of fixations on the agent's body, when observing the agent with the (b) **neutral expression** over one minute.

Table 2
Correlation values between anxiety score and fixations on different portions of the agent for N = 45. (*: $p < 0.05$, **: $p < 0.01$)

	headf	bodyf	eyef	other
Anxiety Score	-0.12	**.30****	-.22*	0.2
headf		-.32**	.56**	-.0.3
bodyf			-.50**	.24*
eyef				.27*

5. Discussion

In **H1** we suggested that the agent affects the user valence and arousal. The results in Fig. 7 show that the agent has a significant effect on the valence and arousal of the user according to the emotion displayed. This shows that the model of the avatar actually works to affect the participant. The differences between valence and arousal of the participant when viewing the neutral agent compared to the baseline were not statistically significant. This is expected because there was no emotion conveyed with the neutral facial expression. The angry and happy facial expressions affected the valence and arousal of the participants significantly compared to the baseline. This showed that there the facial expressions of the agents affected the participants' valence and arousal. The angry facial expression also induced a lower valence score while the happy facial expression induced a higher valence. This confirmed that the facial expressions of the avatar were perceived correctly and supports **H1**.

In **H2**, we predicted that user's acceptability of the agent emotions affects gaze patterns on the agent. SAD individuals do not want to confront the emotions of others, thus exhibiting behaviors such as avoiding the face or being hyper-vigilant to cope with emotional display.

In addition to the anxiety score, we measured if the participant's acceptance or rejection of the agent's emotions played a factor in avoidant gaze patterns. We created the acceptability score as measure of the individual participant's preference to the agent's facial expressions were variant. Individuals with SAD are also anxious at the time they do not accept facial expressions. Thus we measured if there's a correlation between the acceptability score for anxious participants and number of fixations on specific parts of the agent. A positive correlation was found between fixations on the agent's eye and the acceptability score in anxious participants. This shows that even anxious participants are more likely to gaze at the agent's eyes even if they accept the agent's facial expression. Conversely, the opposite is true when the person avoids the gaze of the agent, if they find the agent's emotions unacceptable. This matches the literature for SAD individuals avoiding eye gaze for displays of affection and supports **H2**.

The acceptability score can be used to change how agent's faces are studied. It can be used in the future to study the root cause of SAD or even cultural differences in social norms when both accepting and reacting to varying facial expressions. The acceptability measure can be used as an extra factor in future SAD studies.

For **H3**, we hypothesized that user's overall anxiety affects gaze patterns on the agent. In this study, we took the number of fixations on each part of the agent as an indication of acceptance or avoidance of the facial expressions of the avatar. On one hand, an increased number of fixations on the eyes or head suggest that the participant accepts the avatar. This is consistent to previous studies in [5, 10]. On the other hand, an increased number of fixations on the body or outside of the avatar indicates that the user is avoiding the agent.

We also explored the relationship between agent avoidance and the participants' anxiety level. The initial anxiety score represents an approximation of the participants' overall state before facing the agent. We found significant correlations between the user's initial anxiety and body fixations, invariant of facial expression as shown in Table 2. When analyzing the correlations, on per facial expression basis, there was a strong correlation between the anxiety score and the number of fixations on the body of the agent or outside of the body when presented with the neutral and angry expressions respectively. When viewing the angry expression, the fixations were completely outside the body showing that the participant is more likely to avoid the agent completely the more negative the expression is.

We also found a strong negative correlations between the participants' anxiety score and the number of fixations on the agent's eyes, invariant of facial expression. This indicates that the participant avoids the agent's eyes and is more likely to look at the agent's body or outside of the agent, when anxious, matching the pattern of those with SAD as per [5, 10] and supports **H3**.

When analyzing the correlation between anxiety score and user's fixations on the upper and lower part for all the faces, we found a strong negative correlation between the upper face and the anxiety score. This indicated that the anxious users avoided the agent's eyes as cited in [5, 10]. When analyzing the correlations per facial expression, we only found a strong negative correlation between user fixations on the upper face and the user's anxiety score, with the happy agent. This is probably because the participants were avoiding the face otherwise for the neutral and angry agents. This indicates that anxious users are more likely to look at the agent's face, if it has positive affect, despite avoiding the eyes.

6. Conclusions

The study observes the correlation between anxiety of normative users, their acceptability of the agent's emotions and the respective gaze patterns on an agent with varying emotions. Our results suggest that individuals with anxiety in the moment have similar gaze patterns to those with SAD. The similarities between normative individuals facing anxiety in the moment, their gaze patterns and their acceptability of the agent, can unlock a better understanding of the way SAD individuals operate. The techniques used for SAD individuals can also be used to accommodate normative anxious individuals facing social situations in VR.

In this model the participant is allowed to look away from the face to other sections of the embodied agent including the body, which emulates an actual social situation in VR. Whether the users gazed at the upper or lower parts of the agent's face, were also analyzed. The more negative the expression, the further the anxious participant strayed away from agent's eye, then the face and entire body respectively. These findings are an important indication in designing future systems. E.g. Nonverbal agents with positive affect might be a better choice for an anxious normative individual as they were still more likely to gaze at the face regardless of their anxiety. The studies also show that users are more likely to gaze at the agent's eyes if they accept the agent's emotional display.

A Metaverse VR avatar and VIVE tracker were used in the experiment. The technique can be easily applied to a more ecologically valid setting to find avoidance patterns of anxious users in real-time. This is useful to adjust a non-verbal agent's expression to accommodate for the user's anxiety and facial preference. Other sensors can also be added to find stronger patterns e.g heart sensor or facial tracker. The studies suffer some limitations due to limited participant count and non-validated questionnaires.

References

[1] D. Monteiro, H.-N. Liang, J. Wang, L. Wang, X. Wang, Y. Yue, Evaluating the effects of a cartoon-like character with emotions on users' behaviour within virtual reality environments, in: 2018 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR), IEEE Computer Society, 2018, pp. 229–236.

[2] E. Hasegawa, N. Isoyama, D. V. Monteiro, N. Sakata, K. Kiyokawa, The effects of speed-modulated visual stimuli seen through smart glasses on work efficiency after viewing, Sensors 22 (2022) 2272.

[3] H. Lee, H. Kim, D. V. Monteiro, Y. Goh, D. Han, H.-N. Liang, H. S. Yang, J. Jung, Annotation vs. virtual tutor: Comparative analysis on the effectiveness of visual instructions in immersive virtual reality, in: 2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), IEEE, 2019, pp. 318–327.

[4] D. Monteiro, H.-N. Liang, H. Li, Y. Fu, X. Wang, Evaluating the need and effect of an audience in a virtual reality presentation training tool, in: International Conference on Computer Animation and Social Agents, Springer, 2020, pp. 62–70.

[5] K. Roelofs, P. Putman, S. Schouten, W.-G. Lange, I. Volman, M. Rinck, Gaze direction differentially affects avoidance tendencies to happy and angry faces in socially anxious individuals, Behaviour research and therapy 48 (2010) 290–294.

[6] J. H. Kwon, C. Alan, S. Czanner, G. Czanner, J. Powell, A study of visual perception: Social anxiety and virtual realism, in: Proceedings of the 25th Spring Conference on Computer Graphics, 2009, pp. 167–172.

[7] R. E. Jack, P. G. Schyns, The Human Face as a Dynamic Tool for Social Communication, Current Biology 25 (2015) R621–R634. URL: https://www.sciencedirect.com/science/article/pii/S0960982215006557. doi:10.1016/j.cub.2015.05.052.

[8] H.-S. Cha, S.-J. Choi, C.-H. Im, Real-time recognition of facial expressions using facial electromyograms recorded around the eyes for social virtual reality applications, IEEE Access PP (2020) 1–1. doi:10.1109/ACCESS.2020.2983608.

[9] L. A. Rutter, D. J. Norton, T. A. Brown, Visual attention toward emotional stimuli: Anxiety symptoms correspond to distinct gaze patterns, Plos one 16 (2021) e0250176.

[10] M. Garner, K. Mogg, B. P. Bradley, Orienting and maintenance of gaze to facial expressions in social anxiety., Journal of Abnormal Psychology 115 (2006) 760–770. doi:10.1037/0021-843x.115.4.760.

[11] A. T. Wieckowski, N. N. Capriola-Hall, R. Elias, T. H. Ollendick, S. W. White, Variability of attention bias in socially anxious adolescents: differences in fixation duration toward adult and adolescent face stimuli, Cognition and emotion 33 (2019) 825–831.

[12] Wolf3D, Cross-game avatar platform for the metaverse, in: https://readyplayer.me/, accessed January 14th, 2022, 2022.

[13] M. Shin, S. J. Kim, F. Biocca, The uncanny valley: No need for any further judgments when an avatar looks eerie, Computers in Human Behavior 94 (2019) 100–109.

[14] P. Ekman, W. P. Friesen, Measuring facial movement with the Facial Action Coding System, in: P. Ekman (Ed.), Emotion in the human face, second edi ed., Cambridge University Press, 1982, pp. 178–211.

[15] P. Ekman, E. Rosenberg, What the face reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS), second edition ed., Oxford University Press, 2005.

[16] B. Farnsworth, Facial action coding system (facs) - a visual guidebook, 2021. URL: https://imotions.com/blog/facial-action-coding-system/.

[17] Mixamo, 2022. URL: https://www.mixamo.com/#/.