Augmented Reality Visual Effects for Mitigating Anxiety of In-person Communication for Individuals with Social Anxiety Disorder

Juri Yoneyama[1,*], Yuichiro Fujimoto[1,*], Kosuke Okazaki[2], Taishi Sawabe[1], Masayuki Kanbara[1] and Hirokazu Kato[1]

[1]Nara Institute of Science and Technology, 8916-5, Takayama-cho, Ikoma, Nara, Japan
[2]Nara Medical University, 840 Shijo-Cho, Kashihara, Nara, Japan

Abstract

Individuals with social anxiety disorder (SAD) experience heightened anxiety in in-person communication, particularly when they perceive negative evaluations from their conversation partner based on eye gaze and facial expressions. To address this issue, this study developed an augmented reality (AR) system for in-person conversations, capable of concealing the presence and masking the facial emotions of the conversation partner. An online survey was conducted using Amazon Mechanical Turk (n = 130) to examine the visual effects that individuals with SAD find anxiety-inducing during in-person conversations. 91 out of them have a tendency for SAD. For the aspect of "ease to talk", individuals with SAD achieved significantly higher scores for the visual effects involving a smiley face or anime avatar overlaid onto the conversation partner, as opposed to the remaining visual effects ($p < .05$). Based on these findings, we developed two prototypes wherein an AR head-mounted display (HMD) overlays either a smiley face or an anime avatar onto the conversation partner during an in-person conversation. A small-scale user study was conducted to investigate participant's ability to interact with a conversation partner, using these visual effects, as well as their subjective experiences. The results were further analyzed through user interviews. The social anxiety levels of participants were measured using the Liebowitz Social Anxiety Scale. Participants with SAD exhibited a preference for communicating with an anime avatar rather than real individuals, while participants without SAD preferred real-person interactions. These findings inform future research, wherein the objective (heart rate, eye tracking, self-disclosure) and quantitative data (questionnaires) will be utilized to evaluate anxiety levels during in-person conversations.

Keywords

Social anxiety disorder, Augmented reality, Human-computer interaction, Avatar

1. Introduction

Social anxiety disorder (SAD), also known as social phobia, is one of the most prevalent mental disorders in the world [1]. SAD induces profound anxiety in response to social circumstances such as face-to-face communication and public speaking. Consequently, individuals with SAD often encounter a heightened likelihood of lower education, lower occupational status, and diminished income when compared to individuals without SAD [2].

Psychological and pharmacological interventions are employed to address SAD. However, the rate of seeking treatment is low because individuals with SAD perceive that social anxiety is intrinsic to their personality and resistant to change. Furthermore, barriers to seeking treatment involve an inadequate acknowledgement of the condition by healthcare professionals, apprehension regarding negative evaluation by healthcare providers, and a lack of awareness regarding the availability of effective therapeutic treatment across various regions [3].

A theoretical framework [4] describes the cognitive processes involved in the genesis of social anxiety, and clarifies the psychological mechanisms. Individuals with SAD tend to engage in a cognitive distortion wherein they incorrectly perceive others as evaluating them negatively, despite the absence of explicit negative evaluations. Additionally, in comparison to individuals without SAD, individuals with SAD exhibit a heightened tendency to selectively attend to threatening stimuli, particularly in social contexts. Within such social situations, cues indicative of potential negative evaluations from others, such as eye gaze and facial expressions, serve as salient threats. Consequently, individuals with SAD experience heightened anxiety when subjected to gaze from others, particularly when accompanied by neutral or negative facial expressions (e.g., anger, fear) [5]. Notably, individuals with SAD display a greater degree of comfort when engaging in text-based communication as opposed to face-to-face interactions. This indicates the significance

APMAR'23: The 15th Asia-Pacific Workshop on Mixed and Augmented Reality, Aug. 18-19, 2023, Taipei, Taiwan
*Corresponding author.
yoneyama.juri.yj5@is.naist.jp (J. Yoneyama); yfujimoto@is.naist.jp (Y. Fujimoto); psy-nara-ok@naramed-u.ac.jp (K. Okazaki); t.sawabe@is.naist.jp (T. Sawabe); kanbara@is.naist.jp (M. Kanbara); kato@is.naist.jp (H. Kato)
0009-0006-4942-0411 (J. Yoneyama); 0000-0002-8270-2609 (Y. Fujimoto); 0000-0001-9244-479X (T. Sawabe); 0000-0002-1877-996X (M. Kanbara); 0000-0003-3921-2871 (H. Kato)
© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-3467/paper7.pdf

of eye gaze and facial emotion as triggers for their social anxiety within social settings [6].

Although communication without visual cues such as text message promotes a relative reduction in anxiety experienced by individuals with SAD, numerous social circumstances need in-person interactions, such as seeking to counsel in a hospital or dealing with administrative tasks at a city hall. However, few research aims to aid individuals with SAD for an in-person conversation.

Augmented reality (AR) is a transformative technology that enables individuals to perceive the real world supplemented with virtual objects. AR has brought about enhanced convenience and futuristic experiences, ranging from visualizing virtual constructions within real-world environments to supporting surgeons through information navigation for their surgery [7]. However, the adoption of AR technology to support individuals with disabilities or impairments remains limited, for example, the development of an AR system that enables individuals with colour blindness to perceive real-time scenes through a filtered image [8].

To address this research gap, our study aims to mitigate the anxiety experienced by individuals with SAD during in-person conversations. We aim to accomplish this by designing an AR system capable of modifying or concealing the eye and facial features of others, which have been identified as triggers for social anxiety. We aspire to provide an approach for supporting individuals with SAD to reduce their anxiety about in-person conversation in their daily life.

2. Related works

The treatment and training for SAD have traditionally been administered by human professionals, specifically medical specialists. These days, a number of research has explored the utilization of computer technologies such as the Internet and virtual reality (VR).

Cognitive behaviour therapy (CBT) is an established psychological treatment employed in medical institutions for SAD and has demonstrated effectiveness [9]. Through structured exercises facilitated by a therapist, patients acquire coping skills and learn to modify their cognition, problematic emotions, and behaviours. Despite the overall efficacy of CBT, the utilization of this treatment remains limited. This is because individuals with SAD perceive their condition as their personality and apprehensions about negative evaluations from healthcare professionals prevent them from seeking counseling. Additionally, there exists a lack of information regarding the availability of effective treatment in various regions [3]. Internet-delivered self-help therapy, known as Internet-based cognitive behaviour therapy (ICBT) holds the potential for addressing the challenges. ICBT offers individuals with SAD the opportunity to receive treatment remotely without the need for in-person counselling. Andersson [10] devised ICBT by adapting established CBT protocols and incorporating group exposure sessions. Participants in their study engaged in therapy delivered via the Internet, consisting of learning materials and exercises. These exercises prompted participants to clarify their thoughts and respond to specific inquiries, facilitating therapist evaluation of material assimilation. Furthermore, participants received personalized feedback from therapists via email. CBT and ICBT are grounded in the theoretical understanding that anxiety in SAD derives from cognitive distortions. Their therapeutic objective is the reconstruction of their cognition. However, an alternative perspective suggests that anxiety in SAD can also arise from a lack of social interaction skills. According to this viewpoint, individuals with SAD can mitigate their symptoms by practising social skills. By addressing both cognitive distortions and a lack of social skills, comprehensive treatment approaches can offer a more effective approach to reducing social anxiety symptoms and enhancing overall social functioning [11].

Social skills training (SST) is a therapeutic approach that aims at enhancing social functioning by practising social skills. It has shown effectiveness in addressing various mental disabilities, including SAD, schizophrenia, and autism spectrum disorder. SST is typically conducted by psychological specialists and involves structured exercises that incorporate real-life social interactions, such as making eye contact, asking open-ended questions, and taking turns. Role-playing exercises are utilized in SST [12]. Virtual reality (VR) has been employed in the context of social skills training, specifically for individuals with schizophrenia [13]. These interventions utilize VR to create immersive environments for practising conversation skills, assertiveness, and emotional expression. By engaging in simulated social interactions within a controlled virtual setting, individuals can develop their social skills in a safe and supportive environment. However, it should be noted that individuals with SAD may possess social skills but struggle to demonstrate them due to anxiety or negative thoughts regarding their behaviour. It remains unclear whether SST is essential for every individual with SAD [11].

In order to avoid the challenges of in-person communication, text messages, including email and chat applications, are feasible alternatives. Pierce [6] found that individuals with SAD tend to prefer text messaging over face-to-face communication. Kang et al. [14] investigated the disclosure patterns of individuals with SAD in interactions with 3D avatars compared to real individuals through video calls. They found that individuals with SAD were more likely to disclose themselves to 3D avatars than to conversation partners during video calls. While Kang et al. focused on altering only the

appearance of the conversation partner, Ichino et al. [15] conducted a study exploring self-disclosure by modifying both the appearance of the subject and the conversation partner. They compared three conditions: avatars with high similarity to the participant in virtual reality (VR), avatars without similarity to the participant in VR, and unaltered video calls. The results demonstrated that participants exhibited the highest level of self-disclosure with avatars that lacked similarity to them in VR compared to the other conditions. Although this research did not specifically target individuals with SAD, it suggests that employing visual effects and mediums such as video or VR in face-to-face communication may potentially encourage individuals to engage more in conversation. Building upon the limitations of existing communication tools, a system which can support individuals with SAD for in-person interaction is essential.

3. Main idea

3.1. Causes of anxiety for in-person communication

Individuals with SAD often experience heightened anxiety in in-person conversation situations due to negative interpretations. Research [5] has demonstrated that factors such as eye gaze and facial emotion can trigger anxiety in individuals with SAD. Once they perceive a negative evaluation from the conversation partner, their attention becomes focused on their own physiological arousal, which may manifest as increased heart rate, sweating, or blushing. This heightened self-awareness intensifies their anxiety and leads to avoidance behaviours. Consequently, they feel anxiety about engaging in conversation.

3.2. Change the facial emotion or hide the presence of anxiety by AR

This study proposes an AR approach to address the causes of anxiety in in-person communication by manipulating the presence and facial emotions of the conversation partner. The method involves individuals with SAD wearing an AR head-mounted display (AR HMD) during in-person interactions. AR HMD rendered virtual objects onto the conversation partner (Fig. 1). To investigate the effectiveness of different visual effects in changing or hiding the conversation partner, an online survey was conducted using Amazon Mechanical Turk [16]. Considering that eye gaze and facial emotions are key triggers of anxiety, eight visual effects were prepared. The participants of the survey were presented with nine conditions, including eight variations of visual effects and one condition without any visual effect. The participants were asked to evaluate the level of anxiety they experienced in relation to the conversation partner and the ease of communication in each condition.

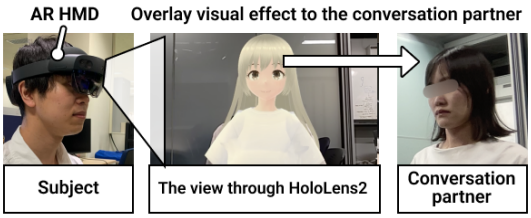


Figure 1: Use cases for our AR system during an in-person conversation. Individuals with SAD wear AR HMD and then talk with their conversation partner. In order to address privacy concerns for this paper, a blurred effect is implemented.

The proposed AR system is designed specifically for dyadic in-person conversation scenarios, where individuals with SAD can utilize the system during an in-person conversation. This approach is particularly relevant in situations where the conversation partner does not explicitly evaluate individuals with SAD negatively (e.g. asking for directions, self-introductions, and engaging in small talk with strangers). The current adaptation of AR HMDs into our daily lives is limited. However, recent advancements in technology have resulted in the development of smaller AR HMDs, such as the XReal Air [17], which is similar to sunglasses. This indicates the potential that AR HMDs will blend into our daily life.

3.3. Research aims

This study aims to reduce the anxiety experienced by individuals with SAD when engaging in in-person conversation by developing an AR system. As an initial phase, we aim to investigate the types of visual effects that can reduce anxiety levels in individuals with SAD. Furthermore, we seek to explore their ability to effectively communicate in situations where visual effects are overlaid onto their conversation partner. This is because overlaying of visual effects on conversation partners is not a common occurrence in real-life interactions and may elicit unfamiliar sensations or perceptions.

4. Pre-survey

In order to find out which specific visual effects are most helpful in supporting ease of conversation and reducing anxiety among individuals with SAD, we conducted an online survey via Amazon Mechanical Turk [18].

4.1. Visual effects

For our study, we prepared eight different visual effects that could be overlaid onto the appearance of the conversation partner (Fig. 2). These visual effects were specifically designed to address the sources of anxiety such as eye gaze and facial expressions experienced by individuals with SAD. The visual effects included: (a) Nothing overlaid (control condition): No visual effect applied, (b) Mosaic pattern eyes overlaid: The conversation partner's eyes are covered with a mosaic pattern, (c) White square overlaid: A white square is overlaid onto the conversation partner's eyes, (d) Anime eyes overlaid: The conversation partner's eyes are overlaid with anime-style eyes, (e) Smiley face overlaid: A smiley face is overlaid with the conversation partner's face, (f) Anime avatar face overlaid: The conversation partner's face is overlaid with an anime-style avatar face, (g) Realistic avatar face overlaid: The conversation partner's face is overlaid with a realistic avatar face, (h) Anime avatar overlaid: The entire body of the conversation partner is overlaid with an anime-style avatar, (i) Realistic avatar overlaid: The entire body of the conversation partner is overlaid with a realistic avatar. We chose a white square as one of the visual effects because it appears nearly opaque when viewed through devices such as Microsoft Hololens2 [19] therefore it can effectively cover the conversation partner's eyes. The use of a mosaic pattern is a common method of concealing the eyes. Additionally, research by Nowak and Raugh [20] has shown that anthropomorphic avatars are perceived as attractive and credible. Therefore, we incorporated anime-style eyes, faces, and full bodies into our visual effects options. Realistic avatars were also included as they are more lifelike than anime-style avatars but not as realistic as actual people. Kang et al. [14] demonstrated that individuals with SAD were more willing to disclose personal information to a realistic avatar compared to actual people during video calls. However, realistic avatars may trigger feelings of discomfort due to the possibility of encountering the uncanny valley phenomenon [21], where people experience unease when an avatar becomes almost indistinguishable from a real human but still falls short of complete realism.

4.2. Participants

We collected responses from a total of 130 participants through Amazon Mechanical Turk. Each participant received a compensation of $1 for completing all the survey questions. To ensure that our analysis focused specifically on individuals with SAD, we filtered the data using the Liebowitz Social Anxiety Scale (LSAS) [22]. Out of the 130 participants, 91 individuals (42 females, 49 males) scored above the threshold for SAD on the LSAS. The age range of these socially anxious participants was between 18 and 69 years. In terms of geographical distribution, 53 participants resided in North America, 34 in Asia, two in South America, and one each in Europe and Africa. Regarding the participants' self-identified ethnic groups, 46 described themselves as White, 34 as Asian or Pacific Islander, four as Latino or Hispanic, four as Black or African American, one as Native American or American Indian, and one participant chose "Other". One participant did not provide an answer to this question.

4.3. Questions

The survey comprised five sections. The first section is about demographic questions such as gender, age, self-identified ethnic groups, and place of residence. The second section is LSAS to filter the people with SAD. The third section is Gaze Anxiety Scale [23] to measure the level of gaze avoidance. The fourth and fifth sections are image evaluation and comparisons among nine conditions. We showed nine visual effects and the introduction "You are wearing AR glasses and try to talk with strangers for asking for directions. In the AR environment, you can overlay a visual effect on the stranger. Please answer the following questions per condition (Fig. 2)". The fourth section is image evaluation. Participants evaluated the nine visual effects individually. Using a seven-point Likert scale (ranging from 1: strongly disagree to 7: strongly agree), participants rated statements such as "I feel that this person would be easy to talk to" and "I feel the anxiety to talk to this person". The fifth section is image comparison. Participants chose the best visual effect from all conditions for specific questions like "Which person do you feel is the easiest to talk to?" and "Which person do you feel the most anxiety to talk to?" Participants selected their preferred visual effect based on their perception of ease or anxiety in communication.

4.4. Result

The statistical analysis was conducted using non-parametric tests due to the non-normal distribution of the data, as confirmed by the Shapiro-Wilk test. Friedman's test was used to determine if there were significant differences among the multiple related groups, and the results indicated a significant difference among the conditions ($p < .05$). To compare paired observations, the Wilcoxon signed-rank test was employed, with a significance level set at $p = .05$. The statistical analyses were performed using IBM SPSS Statistics software [25].

Regarding the image comparison section, the results for the question "Which person do you feel is the easiest to talk to?" showed that (e) was selected by more than 30 out of the 91 participants (Fig. 3). Additionally, for the question "Which person do you feel the most anxiety to

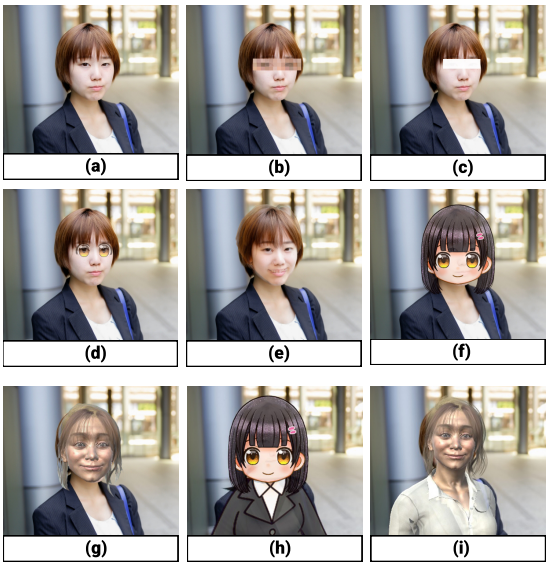


Figure 2: Nine conditions through HoloLens2. (a)nothing overlaid (control condition), (b) mosaic pattern eyes overlaid (c) white square overlaid (d) anime eyes overlaid, (e) smiley face overlaid, (f) anime avatar's face overlaid, (g) realistic avatar's face overlaid, (h) anime avatar's body overlaid, (i) realistic avatar's body overlaid. The image of the conversation partner is a royalty-free image from photo AC [24].

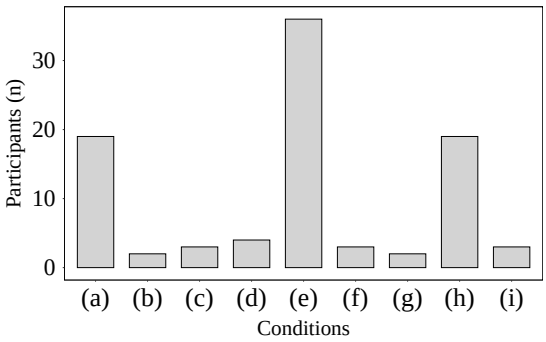


Figure 3: The number of participants who selected each condition for "Which person do you feel easiest to talk to".

talk to?", Figure 4 indicates that (a) was selected by more than 20 out of the 91 participants.

Table. 1 shows mean value for (e) is relatively higher than the rest of the visual effects for "I feel that this person would be easy to talk to". There is also a significant difference between (e) and the rest of the visual effects(Fig. 5)(Table. 2). Similarly, Table. 2 shows that (f) and (h) scored significantly higher than most of the rest of the visual effects.

Fig. 6 shows that (a) scored significantly higher than (e), (f) and (h) for "I feel the anxiety to talk to this person"

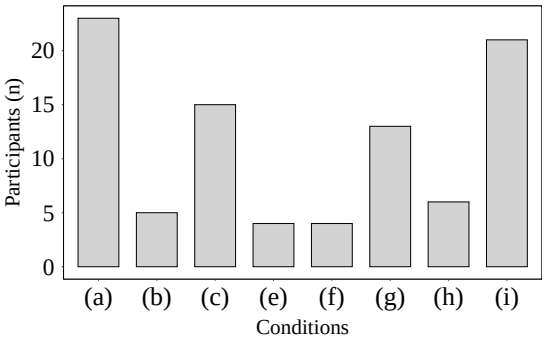


Figure 4: The number of participants who selected each condition for "Which person do you feel the anxiety to talk to is the most?"

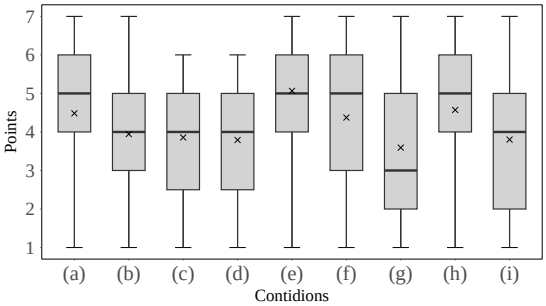


Figure 5: The points of the likert scale that participants scored to range from 1 to 7 for the question "I feel that this person would be easy to talk to" in the image comparison section.

(Table.3).

Table 1
Mean values and standard deviation of "I feel easy to talk to this person" and "I feel the anxiety to talk to this person" based on likert scale to a range from 1 to 7.

	Feel easy to talk	Feel anxiety to talk
	M (SD)	M (SD)
(a)	4.48 (1.486)	4.15 (1.653)
(b)	3.95 (1.493)	4.05 (1.566)
(c)	3.86 (1.488)	4.02 (1.535)
(d)	3.79 (1.567)	4.00 (1.606)
(e)	5.07 (1.467)	3.58 (1.713)
(f)	4.37 (1.644)	3.58 (1.720)
(g)	3.59 (1.619)	4.27 (1.745)
(h)	4.57 (1.674)	3.47 (1.715)
(i)	3.80 (1.701)	3.98 (1.713)

Table 2
p-values of Wilcoxon-signed ranks test for scores of "I feel easy to talk to this person" among the conditions. The red cell shows p less than .05.

	(b)		(c)		(d)		(e)		(f)		(g)		(h)		(i)	
	p	z	p	z	p	z	p	z	p	z	p	z	p	z	p	z
(a)	0.001*	-3.281	0.000*	-3.741	0.001*	-3.233	0.000*	3.483	0.817	-0.231	0.000*	-3.953	0.565	0.575	0.004*	-2.908
(b)			0.538	-0.616	0.334	-0.965	0.000*	5.648	0.044*	2.012	0.033*	-2.133	0.004*	2.917	0.478	-0.709
(c)					0.695	-0.392	0.000*	5.561	0.015*	2.438	0.148	-1.446	0.001*	3.181	0.822	-0.225
(d)							0.000*	5.970	0.005*	2.789	0.099	-1.652	0.001*	3.285	0.863	-0.173
(e)									0.002*	-3.078	0.000*	-6.398	0.024*	-2.253	0.000*	-5.701
(f)											0.000*	-4.044	0.259	1.130	0.004*	-2.844
(g)													0.000*	4.840	0.082	1.739
(h)															0.000*	-4.008

Table 3
p-values of Wilcoxon-signed ranks test for scores of "I feel the anxiety to talk to this person" among the conditions. The red cell shows p less than .05.

	(b)		(c)		(d)		(e)		(f)		(g)		(h)		(i)	
	p	z	p	z	p	z	p	z	p	z	p	z	p	z	p	z
(a)	0.696	-0.390	0.409	-0.826	0.316	-1.003	0.000*	-3.496	0.006*	-2.742	0.510	0.659	0.002*	-3.099	0.325	-0.985
(b)			0.502	-0.671	0.630	-0.482	0.005*	-2.789	0.011*	-2.554	0.274	1.094	0.002*	-3.024	0.496	-0.681
(c)					0.937	-0.079	0.010*	-2.567	0.010*	-2.591	0.135	1.494	0.003*	-2.977	0.889	-0.140
(d)							0.038*	-2.074	0.023*	-2.267	0.077	1.766	0.006*	-2.766	0.809	0.241
(e)									0.868	-0.166	0.001*	3.182	0.492	-0.687	0.050	1.959
(f)											0.000*	3.525	0.560	-0.584	0.028*	2.194
(g)													0.000*	-4.215	0.041*	-2.045
(h)															0.002*	3.144

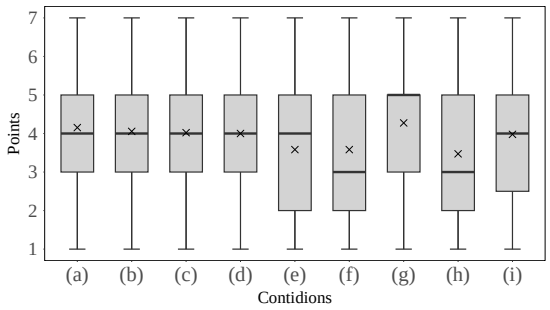


Figure 6: The points of likert scale that participants scored range from 1 to 7 for the question "I feel the anxiety to talk to this person." in the image comparison section.

4.5. Discussion

The findings of the study indicate that participants tend to feel more at ease and less anxious when interacting with a person who is smiling. This supports previous research that suggests individuals with SAD experience anxiety related to the facial emotions of others [5]. Additionally, participants reported feeling more comfortable talking to a person whose full body is covered by an anime avatar. This aligns with the notion that anthropomorphic avatars are perceived as attractive and credible [20]. Although the proportion of participants who chose the control condition is lower compared to the smiley image and full body covered by the anime avatar, it is still comparable to the full body covered by the anime avatar. Participants may have found it strange to interact with the conversation partner whose appearance was overlaid with visual effects, as this is not commonly experienced in real-world interactions. The visual effects were designed to cover the eyes, considering that individuals with SAD tend to avoid direct gaze. However, participants scored significantly lower for visual effects (b), (c), and (d) in terms of feeling that the person would be easy to talk to compared to (e). This suggests that simply hiding the eyes of a conversation partner may not necessarily make individuals with SAD less anxious in an in-person conversation. Furthermore, visual effects (g) and (i), which feature a realistic avatar only on the face, also scored significantly lower than (e). This could be attributed to the potential uncanny valley effect, where discomfort arises when an avatar appears too realistic [21].

5. Prototype

5.1. Design

Our AR system is designed for supporting dyadic in-person conversations for individuals with SAD. This system involves individuals with SAD wearing an AR HMD which visual effects are rendered onto the conversation partner(Fig. 1). Based on insights gained from a preliminary survey, we have developed a prototype that incorporates two distinct visual effects. Fig. 7 is the view through the AR HMD during the dyadic in-person

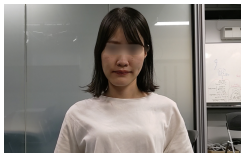
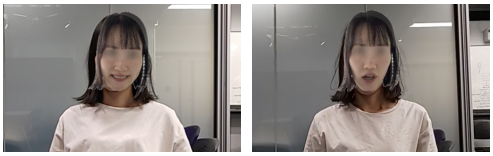


Figure 7: Control condition. In order to address privacy concerns for this paper, a blurred effect is implemented.



Normal Mouth opening

Figure 8: Smiley face overlaid condition. In order to address privacy concerns for this paper, a blurred effect is implemented.

Normal Blinking and mouth opening

Figure 9: Anime avatar overlaid condition. In order to address privacy concerns for this paper, a blurred effect is implemented.

conversation without any visual effect. The first effect involves displaying a 2D image of the conversation partner's smiley face, aimed at altering their facial emotion (smiley face condition)(Fig. 8). However, the static smiley face image may seem incongruous when the conversation partner is talking. To address this, 2D image of the conversation partner changes to an image where their mouth appears to open when they start talking and reverts back to the smiley face when they stop speaking (Fig. 8). In the second visual effect, we employ a 3D anime avatar that matches the gender of the conversation partner, effectively concealing their physical presence (anime avatar condition)(Fig. 9). Its mouth movements are synchronized with the ones of the conversation partner (Fig. 9). Additionally, based on the findings of Takashima et al. [26], which highlighted the influence of avatar blinking on subject impressions, we have also synchronized the blink rate of the avatar to create a more natural and friendly impression.

5.2. System description

The system was developed using Unity 2020.3.38f1 LTS [27]. The system works in Microsoft HoloLens2 [19]. Once it detected the face of the conversation partner, it overlaid the visual effect (i.e. smiley face condition or the anime avatar condition) to the conversation partner.

For the face, eye and mouth tracking of the conversation partner, HoloLens With Dlib FaceLandmark Detector Example [28] was adopted. Male and female anime avatars were created by VRoid Studio [29]. To equalize the condition between the male and female avatars, only the length of the hair and the physique are different. 2D images of the smiley face and the face with an opening mouse of the conversation partner were taken in advance and deployed in the AR environment.

The eye tracking feature is also implemented to get the gaze direction of the participant and whether the participant looks at the eyes or face of the conversation partner.

6. Pre-user study

The pre-user study aimed to assess the participant's ability to engage in natural communication with a conversation partner whose face was covered by a 2D image of a smiley face or whose full body was covered by an anime avatar. Additionally, the study compare the participants' experiences and perceptions across different visual effects to determine which ones were effective in reducing anxiety during the conversation.

6.1. Participants

The study involved five male students from the information science department (Nara Institute of Science and Technology), who voluntarily participated. All participants were Japanese and proficient in the Japanese language. Their ages ranged from 22 to 24 years, with a mean age of $\overline{x} = 22.6$ and a standard deviation of $SD = 0.89$. To measure their levels of social anxiety, the Japanese version of the Liebowitz Social Anxiety Scale (LSAS-J) [22] was administered, resulting in a mean score of $\overline{x} = 48.6$ and a standard deviation of $SD = 21.93$. The threshold for SAD on the LSAS-J is above 30 (Table. 4).

Table 4
Age and LSAS score of participants of user study

	n	M	SD	min	max
LSAS	5	48.60	21.93	23	77

6.2. Procedure

The user study began with a brief explanation of the task, during which participants were given the opportunity to ask any questions related to the study. Following this, participants completed the LSAS-J questionnaire. Once the questionnaire was completed, the experimental sessions commenced. Participants wore the HoloLens2 device and calibrated their eye movements. They then engaged in conversations with a conversation partner who was a well-trained student from the information science department (Nara Institute of Science and Technology). Although participants had some familiarity with the conversation partner, they were not extensively familiar with each other. The conversation partner maintained a neutral facial expression and engaged in passive, non-negative communication with the participants. The study consisted of three conditions: anime avatar overlaid, smiley face overlaid, and no overlay (control condition). The order of these conditions was randomly assigned to each participant. Three topics were provided for discussion, with each topic randomly assigned to one of the conditions. The topics included thoughts about trips abroad, high school life, and experiences at a hair salon. Upon completion of the user study, participants were asked to answer a final questionnaire to indicate which condition they felt easy or anxious. Additionally, they participated in interviews to provide further insights on the user study.

6.3. Result

In the user study, based on the LSAS-J scores (Table. 4), two participants exhibited a tendency for SAD, two participants showed symptoms of generalized SAD, and one participant did not exhibit symptoms of SAD.

6.3.1. Anime avatar condition

The results of the user study indicated that all participants reported an inability to see the eye gaze and facial expressions of the conversation partner. Participants with SAD preferred the conditions where the conversation partner was overlaid with an anime avatar. They found it more comfortable and less anxiety-inducing compared to the control condition. Participants without SAD or with lower social anxiety, however, preferred the control condition. They reported feeling more at ease and natural when communicating without the presence of an avatar overlay. Participants who preferred the anime avatar condition expressed that they did not feel strange or uncomfortable communicating with the avatar. They reported that the experience is similar to interacting with virtual characters such as Vtubers or anime characters with whom they were familiar.

6.3.2. Smiley face condition

Despite the participants' overall inability to perceive the eye gaze and facial emotions of the conversation partner, none of them expressed a preference for the smiley face condition compared to the other conditions. One participant specifically mentioned feeling strange about the fact that only the face of the conversation partner was represented by a 2D image. This suggests that the use of a smiley face overlay did not effectively meet the participants' expectations or provide a sense of comfort. Furthermore, some participants reported that the smiley face condition did not completely conceal the conversation partner's actual facial features. Participants without SAD or with lower social anxiety expressed a desire to see the actual facial emotions of the conversation partner. They believed that observing the genuine facial expressions would provide them with valuable cues about how the conversation partner was reacting to their interactions. One participant with generalized social anxiety reported feeling a sense of safety in the smiley face condition compared to the condition where nothing was overlaid onto the conversation partner's face.

6.3.3. Control condition

Participants with SAD exhibited a preference against the control condition. They expressed that being able to see the face of the conversation partner induced anxiety, as they were concerned about the partner's thoughts and judgments. Moreover, some participants reported heightened anxiety due to their attention being drawn to the eye gaze of the conversation partner. In contrast, participants without SAD or with lower levels of SAD preferred the control condition. They mentioned that the anime avatar and smiley face conditions, which obscured the facial emotion of the conversation partner, caused them to worry about the partner's reactions and responses.

6.4. Discussion

We hypothesized that participants with SAD would exhibit a preference for conditions where visual overlays are present, as these overlays can reduce sources of anxiety [5]. Consistent with our expectations, participants with SAD preferred the anime avatar condition, which effectively concealed the facial emotion of the conversation partner, compared to the control condition. This suggests that individuals with SAD are capable of engaging in conversation with a conversation partner whose presence is obscured by an anime avatar. Interestingly, despite their social anxiety, participants did not prefer the smiley face condition. This can be attributed to the perceived unnaturalness associated with the overlaid 2D image, which surpasses the anxiety deriving from the eye gaze and facial emotion of the conversation partner. As for only

one participant with generalized SAD, the absence of perceived eye gaze and facial emotions, combined with the presence of a smiling 2D image, contributed to a feeling of security.

Conversely, participants without SAD or with lower levels of SAD expressed a preference for the control condition. They experienced anxiety due to the inability to perceive the genuine emotions and reactions of the conversation partner. These findings highlight the divergent sources of anxiety experienced by individuals with SAD and those without SAD.

Participants were also asked about the individuals with whom they experienced anxiety while engaging in conversation and the underlying reasons for their anxiety. All participants reported feeling anxiety when talking with individuals in positions of high authority, such as professors, bosses, or seniors. The reason cited for this anxiety was the concern of being negatively evaluated by them. Participants expressed worries that their knowledge would be deemed insufficient, that they might discuss topics inaccurately, or that their conversations would be perceived as a waste of time by individuals in positions of high authority. Furthermore, some participants mentioned that their anxiety levels were mitigated when individuals in positions of high authority exhibited friendly and casual behaviour. These findings align with the social phobia model proposed by Rapee and Heimberg [4], which assumes that individuals with SAD hold a strong belief that they are being negatively evaluated by others, even in the absence of explicit evaluation. The anime avatar is lacking any cues suggesting high authority. Consequently, participants may not perceive interacting with individuals in the position of high authority, therefore our system may support them. Conversely, in the smiley face condition, participants may still recognize the conversation partner as a high-authority individual. Future investigations should consider the potential influence of vocal cues on participants' anxiety levels when talking with high-authority individuals.

Several individuals experiencing social anxiety have reported feelings of apprehension when engaging in conversations with unfamiliar individuals and those of the opposite gender. We employed an anime avatar characterized by a gender matching that of the conversation partner. While disparities in vocal expression persist, our AR system holds promise in reducing conversational anxiety when the anime avatar's gender aligns with that of the user employing our AR system.

7. Conclusion and future works

This study aimed to develop an AR system with the purpose of assisting individuals with SAD in mitigating anxiety during in-person communication. We conducted an online survey to investigate the specific visual effects that individuals with SAD perceive as reducing their anxiety. Our findings revealed that facial expressions and the visual appearance of these effects serve as cues for reducing anxiety when talking with people. Additionally, we conducted a preliminary user study to explore how individuals with SAD engage and experience communication when an anime avatar or a 2D image of a smiley face is overlaid onto their conversation partner. Qualitative research methods, including user interviews, were employed for evaluation. Participants indicated that communication with a conversation partner concealed by an anime avatar is possible. Facial expressions of the conversation partner play a significant role in influencing the anxiety levels of individuals with SAD, however, the smiley face condition did not yield positive outcomes. Therefore, further improvements are required in our future work. Additionally, the study identified contrasting experiences between individuals with SAD and those without SAD, which presents an intriguing finding deserving of further exploration.

This study is subject to certain limitations. Firstly, the sample size is small and we did not incorporate quantitative data, such as biological indicators such as heart rate and eye movement, and subjective questionnaires such as assessments of anxiety pertaining to verbal communication, in our evaluation. Secondly, participants were not encountering their conversation partners for the first time, which may have resulted in lower levels of anxiety during the conversation. Furthermore, the user study was exclusively conducted with male students, and the sample size was inadequate. We intend to incorporate quantitative measurements to assess anxiety levels during interactions with strangers.

In conclusion, the outcomes of this study bear significant implications for the design of technological interventions for SAD and the enhancement of individuals' quality of life. For instance, individuals with SAD can actively engage in education and broaden their career prospects.

8. Acknowledgments

This work was funded by JST CREST Grant Number JPMJCR19A, Japan.

We would like to acknowledge the guidance provided by Senior Lecturer Mitsuhiro Uratani from Nara Medical Universiy, Assistant Professor Hiroki Tanaka and Professor Satoshi Nakamura from Nara Institute of Science and Technology.

References

[1] D. J. Stein, C. C. Lim, A. M. Roest, P. De Jonge, S. Aguilar-Gaxiola, A. Al-Hamzawi, J. Alonso, C. Benjet, E. J. Bromet, R. Bruffaerts, et al., The cross-national epidemiology of social anxiety disorder: Data from the world mental health survey initiative, BMC medicine 15 (2017) 1–21.

[2] Y. Lecrubier, H. Wittchen, C. Faravelli, J. Bobes, A. Patel, M. Knapp, A european perspective on social anxiety disorder, European Psychiatry 15 (2000) 5–16. doi:10.1016/S0924-9338(00)00216-9.

[3] National Collaborating Centre for Mental Health (Great Britain), Social anxiety disorder: The nice guideline on recognition, assessment and treatment of social anxiety disorder, 2013.

[4] R. M. Rapee, R. G. Heimberg, A cognitive-behavioral model of anxiety in social phobia, Behaviour Research and Therapy 35 (1997) 741–756. URL: https://www.sciencedirect.com/science/article/pii/S0005796797000223. doi:https://doi.org/10.1016/S0005-7967(97)00022-3.

[5] L. Schulze, B. Renneberg, J. Lobmaier, Gaze perception in social anxiety and social anxiety disorder, Frontiers in Human Neuroscience 7 (2013). URL: https://www.frontiersin.org/articles/10.3389/fnhum.2013.00872. doi:10.3389/fnhum.2013.00872.

[6] T. Pierce, Social anxiety and technology: Face-to-face communication versus technological communication among teens, Computers in Human Behavior 25 (2009) 1367–1372. URL: https://www.sciencedirect.com/science/article/pii/S0747563209000971. doi:https://doi.org/10.1016/j.chb.2009.06.003.

[7] S. C.-Y. Yuen, G. Yaoyuneyong, E. Johnson, Augmented reality: An overview and five directions for ar in education, Journal on Educational Technology 4 (2011) 11.

[8] E. Tanuwidjaja, D. Huynh, K. Koa, C. Nguyen, C. Shao, P. Torbett, C. Emmenegger, N. Weibel, Chroma: A wearable augmented-reality solution for color blindness, in: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '14, Association for Computing Machinery, New York, NY, USA, 2014, p. 799–810. URL: https://doi.org/10.1145/2632048.2632091. doi:10.1145/2632048.2632091.

[9] A. P. Association, What is cognitive behavioral therapy?, 2017. https://www.apa.org/ptsd-guideline/patients-and-families/cognitive-behavioral.

[10] G. Andersson, P. Carlbring, A. Holmström, E. Sparthan, T. Furmark, E. Nilsson-Ihrfelt, M. Buhrman, L. Ekselius, Internet-based self-help with therapist feedback and in vivo group exposure for social phobia: a randomized controlled trial. journal of consulting and clinical psychology, 74(4), 677-686, Journal of consulting and clinical psychology 74 (2006) 677–86. doi:10.1037/0022-006X.74.4.677.

[11] T. L. Rodebaugh, R. M. Holaway, R. G. Heimberg, The treatment of social anxiety disorder, Clinical Psychology Review 24 (2004) 883–908. URL: https://www.sciencedirect.com/science/article/pii/S027273580400100X. doi:https://doi.org/10.1016/j.cpr.2004.07.007, social Phobia and Social Anxiety.

[12] A. . D. A. of America, Social skills training (sst): A tool for social anxiety disorder: In conversation with 2023 adaa annual conference presenters, 2017. https://adaa.org/learn-from-us/from-the-experts/blog-posts/conference-professional/social-skills-training-sst-tool.

[13] K.-M. Park, J. Ku, S.-H. Choi, H.-J. Jang, J.-Y. Park, S. I. Kim, J.-J. Kim, A virtual reality application in role-plays of social skills training for schizophrenia: A randomized, controlled trial, Psychiatry Research 189 (2011) 166–172. URL: https://www.sciencedirect.com/science/article/pii/S0165178111002848. doi:https://doi.org/10.1016/j.psychres.2011.04.003.

[14] S.-H. Kang, J. Gratch, Virtual humans elicit socially anxious interactants' verbal self-disclosure, Computer Animation and Virtual Worlds 21 (2010) 473–482. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/cav.345. doi:https://doi.org/10.1002/cav.345.

[15] J. Ichino, M. Ide, H. Yokoyama, H. Asano, H. Miyachi, D. Okabe, "i've talked without intending to": Self-disclosure and reciprocity via embodied avatar, Proc. ACM Hum.-Comput. Interact. 6 (2022). URL: https://doi.org/10.1145/3555583. doi:10.1145/3555583.

[16] Amazon mechanical turk, 2023. https://www.mturk.com/.

[17] XREAL, Xreal air, 2023. https://www.xreal.com/air/.

[18] C. B. Burnette, J. L. Luzier, B. L. Bennett, C. M. Weisenmuller, P. Kerr, S. Martin, J. Keener, L. Calderwood, Concerns and recommendations for using amazon mturk for eating disorder research, International Journal of Eating Disorders 55 (2022) 263–272. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/eat.23614. doi:https://doi.org/10.1002/eat.23614.

[19] Microsoft, Microsoft hololens2, 2023. https://www.microsoft.com/en-us/hololens.

[20] K. L. Nowak, C. Rauh, The influence of the avatar on online perceptions of anthropomorphism, androgyny, credibility, homophily, and attraction, Journal of Computer-Mediated Communication 11 (2005)

153–178. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1083-6101.2006.tb00308.x. doi:https://doi.org/10.1111/j.1083-6101.2006.tb00308.x.

[21] M. Mori, K. F. MacDorman, N. Kageki, The uncanny valley [from the field], IEEE Robotics Automation Magazine 19 (2012) 98–100. doi:10.1109/MRA.2012.2192811.

[22] R. Heimberg, K. Horner, H. Juster, S. Safren, E. Brown, F. Schneier, M. Liebowitz, Psychometric properties of the liebowitz social anxiety scale, Psychological medicine 29 (1999) 199–212. doi:10.1017/S0033291798007879.

[23] F. R. Schneier, T. L. Rodebaugh, C. Blanco, H. Lewin, M. R. Liebowitz, Fear and avoidance of eye contact in social anxiety disorder, Comprehensive Psychiatry 52 (2011) 81–87. URL: https://www.sciencedirect.com/science/article/pii/S0010440X10000398. doi:https://doi.org/10.1016/j.comppsych.2010.04.006.

[24] ACworks Co.,Ltd. , photo ac, 2023. https://en.photo-ac.com/.

[25] IBM, IBM SPSS Statistics, 2023. https://www.ibm.com/products/spss-statistics.

[26] K. Takashima, Y. Omori, Y. Yoshimoto, Y. Itoh, Y. Kitamura, F. Kishino, Effects of avatar's blinking animation on person impressions, GI '08, Canadian Information Processing Society, CAN, 2008, p. 169–176.

[27] Unity Software, Unity, 2023. https://unity.com/.

[28] EnoxSoftware, HoloLensWithDlibFaceLandmarkDetectorExample, 2023. https://github.com/EnoxSoftware/HoloLensWithDlibFaceLandmarkDetectorExample/blob/master/Assets/HoloLensWithDlibFaceLandmarkDetectorExample/HLARHeadExample/HLARHeadExample.cs.

[29] pixiv Inc., Vroid, 2023. https://vroid.com/en.