Avatar-Based Feedback in Job Interview Training Impacts Action Identities and Anxiety

Sarinasadat Hosseini, Jingyu Quan, Xiaoqi Deng, Yoshihiro Miyake, *Member, IEEE*, and Takayuki Nozawa, *Member, IEEE*

***Abstract*—This study examined the use of avatars to provide feedback to influence action identities, anxiety, mood, and performance during job interview training. We recruited 36 university students for the experiment and divided them into two groups. The first group received avatar-based feedback whereas the other group received self-feedback after the first interview session. Results showed that the avatar-based feedback group experienced significantly higher levels of action identities, reduced anxiety, and happier mood in the second interview after the feedback session. Additionally, compared to the self-feedback group, the avatar-based feedback group rated their performance better in the second interview. Furthermore, the effect of avatar feedback on mood and performance varied depending on the participants' initial anxiety levels. For those with low initial anxiety, avatar feedback resulted in higher levels of action identification, a marginally significant increase in positive mood, higher pitch, better word usage, and better self-assessment. In contrast, for those with high initial anxiety, avatar feedback did not cause any significant changes in the action identities or moods, however, it reduced anxiety and the use of weak words. We believe that our findings have implications not only for improving job interview performance but also for designing future communication coaching systems.**

***Index Terms*—Emotion in human-computer interaction, Peripheral measures, Social agents/robotics, Video analysis.**

I. Introduction

JOB interviews are a critical stage in the employee selection process and have been widely studied by scholars [1], [2], [3], [4]. Undoubtedly, some see it as a challenging experience that necessitates time and effort to master nonverbal communication and efficient question-answering. Mastering job interviews requires more than one skill, the most important being overcoming anxiety. Experiencing social anxiety is common when anticipating or experiencing evaluation through social contact [5], [6]. Job interviews [7], [8] and interview practice sessions are both extremely evaluative and can lead to negative thoughts and anxiety [1], [9].

Commonly practiced strategies to overcome job interview anxiety includes performing mock interview sessions, recording oneself answering questions and relistening to the recording, and watching other applicants' interview sessions. Additionally, there is a growing trend of using programmed virtual agents to improve social skills and emotional intelligence [10]. Previous studies used computer simulations [11], virtual reality (VR) training systems [12], [13], or computer-generated characters [14] to reduce anxiety or improve job interview performance. For instance, performing mock interviews with agents and receiving feedback—such as audio, video recordings, or data based on verbal and nonverbal behaviors—demonstrated the ability to improve one's performance during employment interviews [1], [10], [15]. While these systems have the potential to help overcome interview anxiety, they still face several challenges, such as the reluctance of many individuals to review recordings of their performance [1].

How individuals see their actions can impact their emotions and behavior. Previous research shows that analyzing feelings surrounding a negative experience from the third-person perspective can reduce emotional reactivity and increase adaptive behavior [16]. Similarly, adopting an objective view when considering negative experiences can help people focus less on emotionally charged details, thereby enabling them to gain insights from the experience without avoiding it, and protecting them from future negative effects [17]. Therefore, objective perspectives can help people recount their experiences less and reconstrue more [17]. Actions can be identified in different ways, ranging from low-level to high-level [18]. Studies have demonstrated that adopting a third-person perspective has a bidirectional causal relationship to high-level action identification [19]. In addition, non-depressed adults tend to have low-level action identifications when experiencing an increase in sad moods or a decrease in happiness [20].

Based on these observations, this study developed a hypothesis that helping job seekers objectively evaluate their performance during a feedback session could help them analyze their behavior with less emotional attachment, potentially leading to better overall performance during job interviews. To achieve this, we proposed an avatar system with a dissimilar appearance. We suggest this system because previous studies demonstrated

Manuscript received 18 April 2023; revised 11 October 2023; accepted 4 February 2024. Date of publication 8 February 2024; date of current version 2 September 2024. This work was supported by KAKENHI under Grants JP20H03553 and JP21K19787 from JSPS/MEXT, Japan, and in part by the Center of Innovation Program under Grant JPMJCE1309 from Japan Science and Technology Agency (JST), Japan. Recommended for acceptance by M.T. Sezgin. *(Corresponding author: Takayuki Nozawa.)*

Sarinasadat Hosseini is with Panasonic Corporation, Osaka 571-8506, Japan (e-mail: hosseini.sarinasadat@gmail.com).

Jingyu Quan, Xiaoqi Deng, and Yoshihiro Miyake are with the Tokyo Institute of Technology, Tokyo 152-8550, Japan (e-mail: quan.j.aa@m.titech.ac.jp; deng.x.ac@m.titech.ac.jp; miyake@c.titech.ac.jp).

Takayuki Nozawa is with the University of Toyama, Toyama 930-8555, Japan (e-mail: nozawa@eng.u-toyama.ac.jp).

This article has supplementary downloadable material available at https://doi.org/10.1109/TAFFC.2024.3363835, provided by the authors.

Digital Object Identifier 10.1109/TAFFC.2024.3363835

that avatars provide people with a different persona to relate to and induce behavioral changes [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33]. Therefore, an avatar may allow applicants to view their actions from a third-person perspective, thereby mitigating the impact of emotional focus and anxiety by creating a sense of psychological distance from the self. In the following section, we present additional information on anxiety and action identity [18], [34], [35], [36], automated job interview training, and social anxiety therapy systems before introducing our novel system and hypothesis regarding the potential outcomes that could be achieved by using it.

II. Related Works

A. Anxiety and Action Identities

Anxiety is related to the apprehensive anticipation of a threat in the future. Although anxiety is a normal feeling, it has been recognized as a disorder in psychiatric classifications since the 20th century [37]. Due to the evaluative and competitive nature of the job application process, anxiety is a normal feeling in job applicants [38]. Several factors can cause anxiety during job interviews, which can lead to low score in job interviews. For example, the interviewer is often a stranger, and the circumstances of the interview are not always within the applicants' control [39], [40], A variety of ways exist to treat anxiety [41]. Visual feedback, pioneered by [42] and later improved by incorporating cognitive preparation for the task [43], is one of the most practical methods to handle anxiety. Although visual feedback has yielded promising results in speech preparation and patients with distorted self-perception [44], [45], it did not significantly reduce social anxiety [45]. Therefore, it is necessary to improve the video-feedback method to reduce anxiety in social situations.

Visual feedback may not be a useful tool for preparing for job interviews owing to several reasons. For instance, anxious people tend to have a more negative self-image [43] and may become even more anxious once they receive negative feedback. Moreover, when faced with subjective feedback and a sense of failure during an interview practice session, even individuals with lower to moderate levels of anxiety may move to lower levels of action identities [18], [34], [35], and become overly fixated on small details, such as their posture or tone of voice, rather than the interview objectives.

According to the Action Identification Theory [34], actions can be identified at various levels, ranging from low-level identities that consider details of how an action is performed to high-level identities that consider the purpose or effect of an action. In general, when an action is difficult or complex, people develop lower-level action identities or agency to focus more on details or the means of action [18], [34], rather than understating goals and objectives. This makes individuals with lower action identities more sensitive to contextual cues, more impulsive, and less self-motivated compared to people with higher levels of action identities [18]. Therefore, while video feedback may help individuals reflect on their behavior at low-level identifies, it may not help them adapt to stressful circumstances that trigger significant anxiety, such as job interviews.

B. Automated Job Interview Training and Social Anxiety Therapy Systems

The development of artificial intelligence (AI)-based interview recruiting systems has advanced significantly in recent years [46], [47]. Furthermore, the number of interview training systems developed has also increased, with many of them focusing on visualizing nonverbal behavioral data during or after the interview session [1], [10], [15], [48], [49], [50], [51]. This can be attributed to the increased availability of virtual reality (VR) and conversational agents (CA). We will further review previous literature in sequence.

Virtual reality exposure treatment (VRET) is less expensive and more accessible than therapists. According to poll results, most people prefer VRET to regular therapy [52], [53]. Owing to several reasons, avatars in VRET might be beneficial in reducing anxiety, such as feeling less self-conscious when interacting with others anonymously. Therefore, VRET has been widely used to treat social phobia [54], [55], performance anxiety in musicians [56], and even public speaking [55], [57]. Although patients frequently choose the first-person view in VRET, some effects of avatars in the third person while using doppelgangers for male participants or dissimilar avatars relative to the assigned similar avatar to reduce public speaking anxiety have been recorded [58], [59].

Earlier research has shown that the type of virtual body and its appearance can cause perceptual and behavioral changes [21], [22], [23], [24], [25], [26]. Initially referred to as the "Proteus effect" [23], Yee et al. stated that avatar size and attractiveness can influence social interactions and the user's behavior. Several studies have reported the impact of different features of avatars, such as body size [27], age [28], gender [29], or skin color [24], on users' attitudes. Some of these behavioral changes are reported to last even for a week [30]. These behavioral changes apply to cognitive processes as well. For example, a previous study [31] found that using an avatar depicting Dr. Freud improved participants' moods and helped change their way of thinking compared to when the therapist was a replica of themselves. Some studies suggested that the appearance of an avatar can change learner motivation and purpose [32], or embodying Einstein could help individuals perform better in cognitive activities [60]. Therefore, it is not surprising that previous studies [33], [59] demonstrated the usefulness of impersonating someone else while using VR.

Similar to VRET, CA systems such as smart speakers and personal assistants are gaining wide popularity. CA systems might soon be compared to human therapists in terms of speaking anxiety treatment owing to their accessibility and capacity to replicate human speech. CA systems can potentially conduct training sessions on social skills [61], particularly for those afraid of public speaking [62]. Although anxiety reduction using VRET or public speaking training with CA systems has been extensively researched in previous studies, the possibility of using avatars to alter one's action identities or lessen anxiety

following a training session with CA systems opens a unique path for investigation.

C. Hypothesis

This study investigates the effect of evaluating one's own behavior in a mock job interview session through a dissimilar avatar on action identities, anxiety levels, and interview performance compared to a camera-recorded video of oneself. Furthermore, this study aims to determine if changes in behavior differ based on participants' initial anxiety levels.

First, we hypothesize that evaluating behavior through a dissimilar avatar, compared to the camera-recorded video of oneself, will cause a shift in action identities to higher levels or anxiety reduction during the feedback session, resulting in better performance in the second interview session.

Second, we hypothesize that people with higher anxiety levels would evaluate their performance in a worse manner. We anticipate that if the feedback is in a camera-recorded format and relates to the self, they will experience increased anxiety owing to their negative self-image or inability to interpret the visual feedback objectively. Conversely, regardless of the feedback mode, we do not expect a significant change in anxiety levels for those with low to moderate initial anxiety levels considering they do not struggle with anxiety. However, because action identity level decline is possible due to negative feedback related to self in the camera-recorded format, we may observe worse performance or even the emergence of anxiety in participants with low-anxiety who receive such negative feedback.

III. Interview and Feedback System

This study investigates how an avatar feedback system can benefit job applicants by modifying their action identities or moderating their anxiety during a feedback session. To achieve this, we conducted a study wherein university students used our system to practice interviewing skills. In the following sections, we explain how we developed automated agents that function as interviewers and avatars in our system.

A. Avatars Used in the Feedback Session

We utilized MetaHuman Creator [63] to create avatars in three-dimensions and exported them to Unreal Engine 5 using the Quixel Bridge Plugin. We created one male and one female persona and matched them to each participant's self-reported gender, as shown in Fig. 1. We used Live Link Face (Version 1.2.1) for Unreal Engine to collect and map the facial motions of participants for the feedback session. Details on how to animate avatars using Live Link face can be found in [64].

B. Interviewer Agents Used in Interview Session

We utilized MetaHuman Creator [63] along with Unreal Engine 5 to create one female and one male agent as interviewers, as shown in Fig. 2. We chose the artificial agents' gender to match the self-reported gender of the subjects to prevent any potential unwanted feelings of intimacy with the allocated agents based on the avatars' gender or vice versa, assuming such a connection

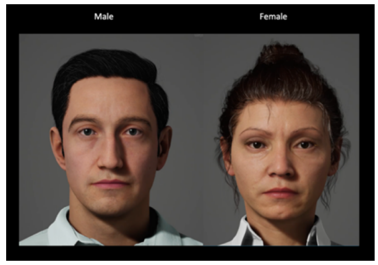

Fig. 1. Avatars used during the feedback session.

Fig. 2. Agents used as interviewers during the interview sessions.

may result in a higher/weaker agent likeability or attachment to the avatar.

The agents' voices were pre-generated using a free online text-to-speech generator (FROMTEXTTOSPEECH.COM). During the interview sessions, the automated agents introduced themselves to the participants and asked five typical job interview-related questions [1] (Appendix 1, available online). Questions were given to participants beforehand to standardize the difficulty level of questions between the first and second interviews and allow subjective comparison of candidates' conduct and performance between the two sessions. Although providing questions in advance might reduce the potential anxiety and unpredictability for interviewees, we should note that common questions are frequently asked in real job interviews. Accordingly, applicants' anxiety is often caused by meeting a stranger or feeling judged rather than by the unpredictability of the interview questions.

To stimulate anxiety associated with encountering a stranger, we carefully designed the appearance and behavior of our interviewers. The interviewers were intentionally modeled as middle-aged professionals, which is commonly observed in real-world interviews. Their speech and facial expressions were scripted to convey a neutral tone and incorporated gestures such as nodding, which have been observed in empirical studies of job interview interactions. Further details regarding the design of the agents are provided in the following.

To animate agents, experimenters used Live Link Face to record their facial movements with neutral expressions and create facial movements for artificial agents based on the content of the questions or nonverbal behaviors during a listening mode. These movements were mapped onto the agents using Unreal Engine 5, and a video was created featuring the agents' facial movements. Video frames were then trimmed and turned into GIF files to program the agents' movements in JavaScript.

Nonverbal behaviors such as blinking, nodding, frowning, and smiling were added to the agents during the listening mode to make them more human-like.

We developed four versions of combined head nods and blinks so that the agent could produce any of them within 4 s on average during the listening mode. We used such timings based on previous studies, which showed that counselors nod their heads every 4.12 s on average during interviews [1] and blink every 2–10 s. Frowns were another nonverbal action designed to occur when participants remained silent or spoke in a quiet voice for more than 3–5 s to urge communication. To detect vocal activity and silence, we utilized VAD.js [65]. Smile was the fourth nonverbal activity designed to occur randomly between head nods and blinks. Notably, although agents' nonverbal behaviors were automated, the experimenter played the Wizard of Oz role and controlled the questions' timing.

IV. Experiment

A. Participants

We recruited 36 international students (18 males and 18 females, aged 20–37) for this experiment using posters and social media. The students were of various nationalities: 26 Asian, 8 White/Caucasian, and 2 African. Participants were randomly assigned to either the avatar-based feedback or self-feedback groups (eighteen in avatar-based feedback and eighteen in self-feedback). The Human Subjects Research Ethics Review Committee of the Tokyo Institute of Technology approved this study procedure, and participants were paid 3000 Yen as a compensation for their time and effort. To determine the sample size, we used G∗Power software, considering a significant level α (Type I error rate) of 0.05, aiming statistical power $1-\beta$ of 0.8, and strong effect size (Cohen's d) of 0.85. G∗Power estimated for a minimum of 18 subjects per group to detect significant behavioral differences.

B. Procedure

The experiment was conducted in a laboratory. Participants were asked to sit in front of a monitor displaying the experiment procedures and the artificial agent. First, the experimenter provided an explanatory video and asked participants to sign a consent form, followed by a pre-questionnaire using the Measure of Anxiety in Selection Interview (MASI) [66], Self-Assessment Manikin (SAM), an affective rating system developed by Lang in 1980 [67], and the Behavior Identification Form [34] to measure their initial anxiety levels, emotional states, and action identities before the interview. After completing the pre-questionnaire, the experimenter set up the measurement equipment and launched the artificial agent. When the agent began speaking, the experimenter moved out of the way to create a one-on-one space for the interview. As shown in Fig. 3, the experiment comprised three sessions: two interviews with the agent and a feedback session in which participants watched a video of their performance (self-feedback or avatar-based feedback) from the first interview session.

Depending on the randomly assigned condition, we divided the participants into two groups (self-feedback and avatar-based feedback). The self-feedback group watched a camera-recorded video of their behavior, while the avatar-based feedback group watched a video wherein an avatar replayed their behavior from the first interview session. We used a video camera (FDR-AX40; Sony) to record the participants' video and audio during all three sessions. There was no time limit for the interviews, but on average, participants took approximately 4.5 min to complete each interview session. Examples of avatar-based feedback and self-feedback are shown in Fig. 4. How our system works can be observed in Appendix 2, available online.

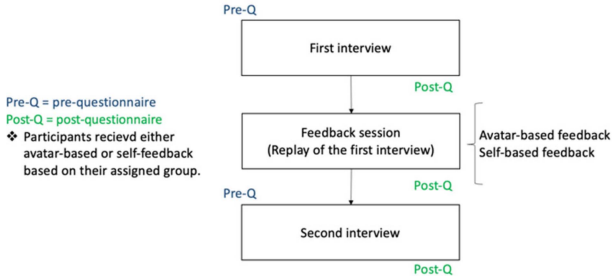

Fig. 3. Flowchart of the experimental procedure.

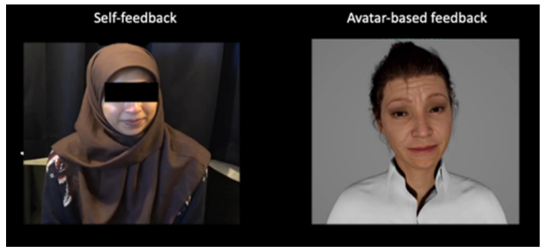

Fig. 4. Self-feedback on the left and avatar-based feedback on the right.

We instructed participants to complete another pre-questionnaire before the second interview session, including SAM and the Behavior Identification Form. Furthermore, at the end of all three sessions, including the two interview sessions and the feedback session, we asked participants to complete a post-questionnaire to assess their emotional state, action identities, and self-rating for job interview performance.

Notably, we asked all participants to read the questions aloud before the experiment; for the avatar-based feedback, we replaced the agents' voices with the applicants' voices. As a result, while the avatar was ready immediately after the interview, we needed a few minutes to modify the feedback video and dub agents' audio with applicants. Therefore, we added a 30 min break between the first interview and the feedback section. Although dubbing was only required for the avatar-based condition, we applied this break time to both conditions to ensure equal conditions for all participants. Furthermore, both groups had a typical 6 min break following the feedback section.

C. Measures

1) *Measuring Initial Job Interview Anxiety:* We expected that individuals with different anxiety levels would perceive feedback differently and behave differently upon receiving it. Therefore, we used the average responses from all scales

in the MASI questionnaire to measure initial job interview anxiety. We set the threshold as 16.8 (median value of anxiety scores) and divided the participants into two equal-sized groups of 18: high and low to moderate anxiety. There were exactly 9 participants with high anxiety and 9 participants with low to moderate anxiety in both avatar and control groups. Furthermore, no significant difference was found in comparing the avatar and control group based on their initial levels of anxiety ($t(34) = 0.35$, $p = 0.72$).

2) *Action Identities Assessment:* We measured action identities before and after each interview session and immediately after the feedback session using the Behavior Identification Form to detect a possible difference between the two groups in action identity change upon receiving feedback [34]. We mainly focused on action identity changes during the feedback session, as previous studies show that people move to lower-level identities upon receiving negative feedback on their performance [34], [35].
3) *Mood Assessment:* We examined the mood before and after each interview session and immediately after the feedback session using the SAM questionnaire. Participants rated their feelings in three dimensions of valence, arousal, and dominance on a nine-level scale from -4 to +4.
4) *Interview Performance Assessment:* Interview performance was assessed after each interview and feedback session. Participants rated their overall performance during the interview on a scale of 1 to 7. We asked participants to self-evaluate themselves because traditional blind scoring of evaluators to grade interview performance could be invalid owing to the varied expertise of the participants. Additionally, we assessed how participants felt about the system and agent, but we did not publish the results due to a lack of relevance to the objective of this study.
5) *Real-time Measurements of Anxiety:* As receiving self-related feedback can link to anxiety, we measured real-time anxiety during both interview sessions and the feedback session using a Shimmer3 GSR+ device, which is a highly extensible wireless sensor platform that can be utilized for biomedical research applications [68]. We attached the device to the right hand of the participants and advised them to avoid moving their hands during the measurements. Using the device, we could assess Skin Conductance Response (SCR), whose variations in the phasic component are known as Galvanic Skin Response (GSR) peaks and can provide information about emotional arousal to stimuli and stress [69]. To assess those peaks, we used NeuroKit2, version 0.2.1, an open-source Python package for neurophysiological signal processing [70]. GSR data were collected at a sampling rate of 128 Hz. The Python code used for the preprocessing and peak detection is provided in Appendix 3, available online.
6) *Behavior Assessment:* Previous studies demonstrated the link between verbal and non-verbal behaviors and interview performance. However, considering the short time of interviews, interviewers primarily focus on nonverbal cues to address applicants' personalities, and the interview outcome is affected more by nonverbal cues than verbal cues [71]. For instance, non-verbal behaviors such as smile information and occurrence [72], direct gaze [73], high vocal pitch, or moderate speech rate [1] are reported to influence the interviewee's evaluation, leading to positive impressions. Along with such non-verbal behaviors, speech fluency, typically measured through word usage, is one of the principal variables for an interview outcome [74].

   In this experiment, we focused on smile ratio and vocal pitch as non-verbal behaviors. We assessed the occurrence of a smile per frame using OpenFace [75], which can record Action Unit (AU) occurrences per frame from videos. Based on Paul Ekman's group, AU 12 defines smiles [76]; therefore, for each participant, we summed up the number of frames with AU12 and divided it by the duration of the session to determine the smile ratio of that session. To assess the pitch activity per second, we first used Parselmouth, a Python interface for Praat [77], and averaged the pitch scores. Furthermore, similar to [1], we measured weak word usage as verbal behavior during job interviews. While other aspects of speech content could be used to evaluate their speech, we utilized the ratio of weak words or fillers to speech length to measure fluency. Excessive use of fillers has been linked to reduced effectiveness [78], and considering fluency's importance in job interviews, this ratio is a more appropriate metric than merely counting weak words or considering the speech duration when evaluating word fluency. To assess the weak word ratio, we transcribed interview sessions and automatically identified filler words (e.g., "like," "basically," "umm," and "totally") using the NLTK library [79]. Finally, to calculate the weak word ratio, we divided the number of weak words used in a session by the session duration.

V. Results

A. Action Identities Assessment

The t-test was used to assess changes in action identities between the avatar-based feedback (avatar) and self-feedback (control) groups before and after the feedback session. The results showed that after the feedback session, individuals in the avatar group generally thought at a higher level ($M = 0.44$, $SD = 1.34$) than those in the control group ($M = -0.50$, $SD = 1.42$; $p = 0.04$; Fig. 5(a)). Furthermore, a significant difference was observed in action identity after receiving feedback between the avatar and control groups for individuals with low initial levels of anxiety ($t(16) = 2.49$, $p = 0.02$); however, no difference was observed for those with high initial levels of anxiety ($t(16) = 0.43$, $p = 0.67$; Fig. 5(b)). The results are summarized in Appendix 4, available online.

B. Mood Assessment

To determine the effect of the feedback mode on mood, we studied the self-reported mood reports of the two groups

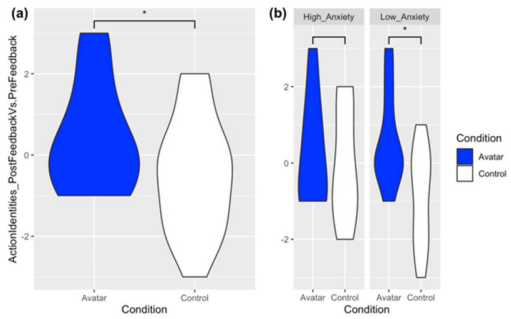

Fig. 5. Effects of feedback mode on action identities during the feedback session. (a) All participants and (b) low-anxiety vs. high-anxiety.

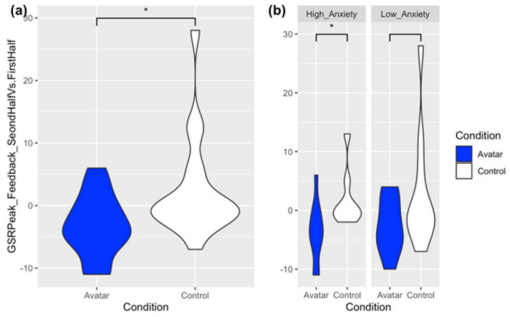

Fig. 7. Effect of feedback mode on GSR peak changes during the feedback session. (a) All participants and (b) low-anxiety vs. high-anxiety.

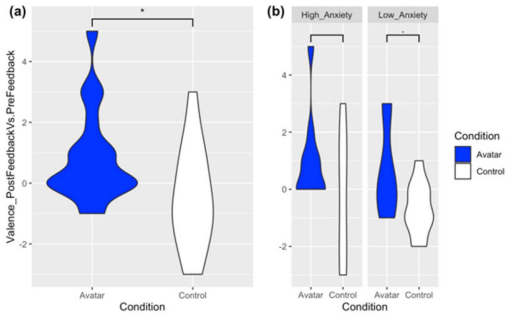

Fig. 6. Effects of feedback mode on valence during the feedback session. (a) All participants and (b) low-anxiety vs. high-anxiety.

before and after the feedback session. The t-test results showed that valence in the avatar group (M = 0.83, SD = 1.54) was significantly modified toward happiness when compared to the control group (M = -0.5, SD = 1.65; p = 0.01; Fig. 6(a)). Because the Shapiro-Wilk normality tests revealed that the arousal (W(36) = 0.90, p = 0.004) and dominance (W(36) = 0.88, p = 0.001) data distributions were significantly different from a normal distribution, Wilcoxon rank-sum tests were used, and no significant difference was found in arousal (Z = 0.93, p = 0.35) or dominance (Z = 1.49, p = 0.13) between the two groups.

Then, we studied the difference between the two groups in self-reported mood changes between the second and first interview sessions. We used the Wilcoxon rank-sum test for valence because the valence data significantly differed from a normal distribution (W(36) = 0.94, p = 0.04), and the two groups had no significant difference (Z = 0.36, p = 0.71). When we used the t-test to compare arousal and dominance, no difference was found in arousal (t(34) = 1.46, p = 0.15) or dominance (t(34) = 0.48, p = 0.63) between the two groups.

Finally, we observed the influence of the feedback mode on the mood in the two groups for individuals with initial low and high anxiety. The results showed that the avatar feedback, except for just a marginally significant impact on leading valence to happiness for low-anxious individuals during the feedback session (Z = 1.78, p = 0.07; Fig. 6(b)), had no significant impact in other categories; the results are summarized in Appendix 4, available online.

C. Real-time Anxiety Assessment

To analyze the potential influence of the feedback mode on real-time anxiety, measured by GSR peaks, we calculated the ratio of GSR peaks to the session duration. First, we compared the number of peaks between the first and second halves of the feedback session. The Shapiro-Wilk test indicated that the distribution of peak changes significantly differed from a normal distribution (W(36) = 0.84, p = 0.0001). Therefore, we used the Wilcoxon rank-sum test, and results showed that participants in the avatar group had fewer peaks during the second half of the feedback session (M = -3.22, SD = 4.86) compared to those in the control group (M = 2.50, SD = 8.13; p = 0.01; Fig. 7(a)). However, the Wilcoxon rank-sum test did not reveal a significant difference in the total GSR peak ratio between the two groups during the feedback session (Z = 1.08, p = 0.18).

Next, we checked the anxiety peak ratio changes between the two interview sessions. Because the distribution of GSR peak ratio changes was significantly different from a normal distribution (W(36) = 0.92, p = 0.010), we used the Wilcoxon rank-sum test, and no significant difference (Z = 1.11, p = 0.26) was observed between the two groups.

Additionally, we compared the changes in the GSR peaks during the feedback session between the avatar and control groups for individuals with high and low anxiety levels. The t-test results revealed that avatar feedback significantly reduced the GSR peaks (t(16) = 2.32, p = 0.03; Fig. 7(b)) for highly anxious participants during feedback but not for low anxiety participants (Z = 1.2, p = 0.23). Furthermore, no significant impact of the feedback mode on GSR change was observed between the two interview sessions for either low (Z = 1.07, p = 0.28) or high anxiety (t(16) = 0.03, p = 0.97) individuals. The results are summarized in Appendix 4, available online.

D. Interview Performance Assessment

First, we performed a t-test on the self-reported scores after the first interview session to see if participants with different levels of anxiety rated their performance more negatively or positively depending on their initial anxiety level. The results showed that highly anxious individuals (M = 3.44, SD = 1.08) rated their performance worse than the low-anxious ones (M = 4.64, SD = 1.12; p = 0.004).

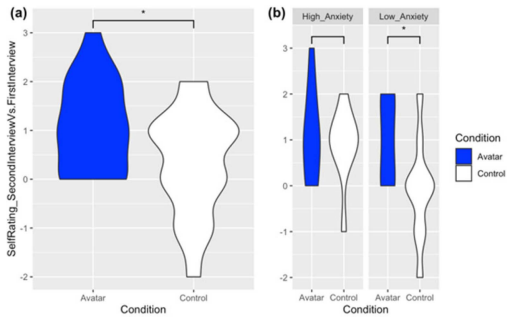

Fig. 8. Effects of feedback mode on changes in self-rating on job interview performance between two interview sessions. (a) All participants (b) low-anxiety vs. high-anxiety.

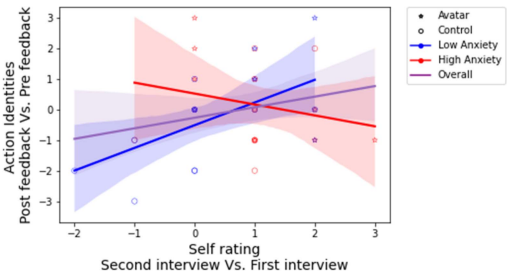

Fig. 9. Correlation between self-rating performance changes in two interview sessions and changes in the action identities level during the feedback session.

Next, to test the possible effect of feedback mode on the overall performance, we acquired self-reported scores after both interview sessions and calculated the difference between those two ratings for each participant. Since the Shapiro-Wilk normality test illustrated that the data distribution significantly differed from a normal distribution ($W(36) = 0.93$, $p = 0.01$), we used the Wilcoxon rank-sum test. The results showed that participants in the avatar group rated their performance marginally better in the second interview than the first interview ($M = 1.06$, $SD = 0.94$) compared with participants in the control group ($M = 0.33$, $SD = 1.08$; $p = 0.08$; Fig. 8(a)). Furthermore, an improvement was observed in the self-reported scores in the second interview for participants who were initially low-anxious and assigned to the avatar group ($Z = 1.95$, $p = 0.05$). No difference was observed in participants with high anxiety levels regardless of their assigned group ($Z = 0.18$, $p = 0.85$; Fig. 8(b)). The results are summarized in Appendix 4, available online.

Additionally, we checked the correlation between the self-rating performance score changes in two interview sessions with action identities, valence, and anxiety level changes during the feedback session using linear regression.

Regarding the correlation between changes in the self-rating performance and action identities alteration during the feedback session, although we did not observe any overall significant correlation between these two variables ($t(34) = 1.53$, $p = 0.13$), a significant correlation was observed for individuals with low initial anxiety levels ($t(16) = 2.81$, $p = 0.01$; Fig. 9), with an R2 of 0.33. No such correlation was found for participants classified as highly anxious ($t(16) = -0.93$, $p = 0.36$).

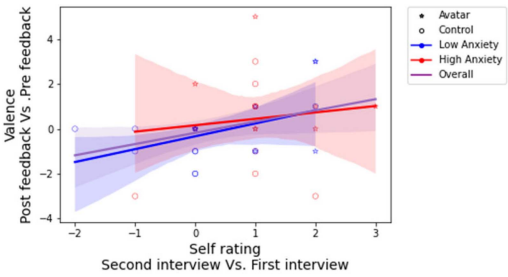

Fig. 10. Correlation between self-rating performance changes in two interview sessions and changes in mood valence during the feedback session.

Considering the correlation between changes in mood valence and self-rating performance change, we observed an overall marginally significant correlation between two variables ($t(34) = 1.90$, $p = 0.06$); such correlation was observed in individuals with low initial anxiety levels ($t(16) = 2.10$, $p = 0.05$; Fig. 10), with an R2 of 0.21, than in highly anxious individuals ($t(16) = 0.52$, $p = 0.6$).

Finally, regarding the correlation between self-rating performance improvement and GSR peaks increase during the feedback session, an overall significant negative correlation ($t(34) = -2.23$, $p = 0.03$), with an R2 of 0.12 was observed. There was no significant relationship for individuals with low anxiety ($t(16) = -1.44$, $p = 0.16$) or those with high anxiety ($t(16) = -1.46$, $p = 0.16$).

E. Non-Verbal Behavior Assessment

To test the effect of the feedback mode on non-verbal behavior, we studied the changes in the smile ratio and vocal pitch during the two interview sessions.

Assessing the normality of smile data using the Shapiro-Wilk test, we found that data were not normally distributed ($W(36) = 0.75$, $p < 0.0001$), and we used the Wilcoxon rank-sum test. Results showed no significant difference in the smile ratio change for either of the groups between interview sessions ($Z = 0.70$, $p = 0.48$).

Next, to study the impact of the feedback mode on the average pitch change between the two interview sessions, we used the Wilcoxon rank-sum test because the data was not normally distributed ($W(36) = 0.87$, $p < 0.0001$). The results showed that the average pitch increased significantly for the avatar group during the second interview session compared to that for the control group ($M = 6.63$, $SD = 20.38$ vs. $M = -7.90$, $SD = 14.98$; $p = 0.01$; Fig. 11(a)). Such increments in pitch for individuals assigned to the avatar group were significantly observed for the ones with initially low anxiety ($Z = 2.04$, $p = 0.041$; Fig. 11(b)) and not those with high levels of anxiety ($t(16) = 1.23$, $p = 0.23$). The results are summarized in Appendix 4, available online.

Finally, we investigated the relationship between pitch change, action identities, and GSR peak changes during the feedback session. A positive correlation was observed between average pitch change and action identification change ($t(34) = 2.07$, $p = 0.04$), with an R2 of 0.11. However, such correlation only existed for individuals with initially low levels of anxiety

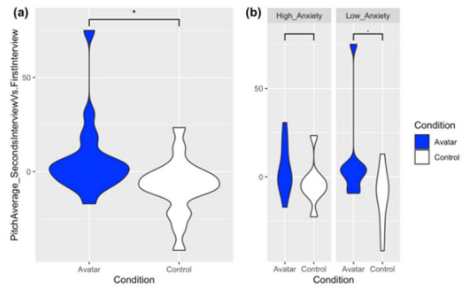

Fig. 11. Effects of feedback mode on the average pitch change between two interview sessions. (a) All participants and (b) low-anxiety vs. high-anxiety.

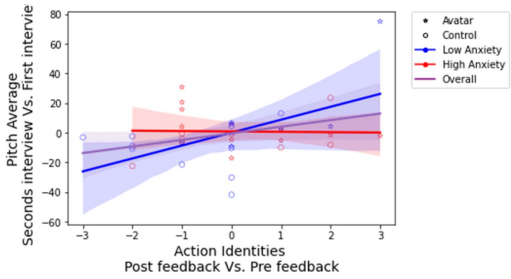

Fig. 12. Correlation between pitch change between the two interview sessions and action identities change during the feedback session.

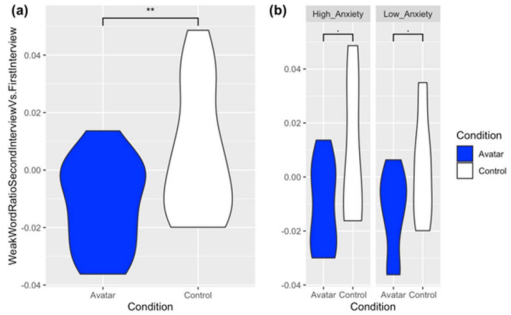

Fig. 13. Effects of feedback mode on weak word usage difference between two interview sessions. (a) All participants and (b) low-anxiety vs. high-anxiety.

(t(16) = 2.60, p = 0.01, with an R2 of 0.29) and not for highly anxious ones (t(16) = -0.10, p = 0.91; Fig. 12). There was no correlation between the average pitch change and GSR peaks change during the feedback time (t(34) = -1.33, p = 0.19).

F. Verbal Behavior Assessment

We focused on weak word usage to test the effect of feedback mode on verbal behavior as speech fluency. Therefore, we calculated the change in weak word usage ratio for each participant in the second interview from the first interview. The results showed that the ratio of weak word usage significantly decreased in the second interview session for participants in the avatar group (M = -0.01, SD = 0.02) than that for participants in the control group (M = 0.01, SD = 0.02; p = 0.007; Fig. 13(a)). Additionally, only a marginally significant effect was observed for avatar feedback mode on reduction of weak word usage for individuals with low (t(16) = 2.04, p = 0.05) and high (t(16) = 1.96, p = 0.06) initial levels of anxiety (Fig. 13(b)). The results are summarized in Appendix 4, available online.

Finally, although no relationship was observed between the change in weak word usage between interview sessions and action identity change during the feedback session (t(34) = -1.11, p = 0.27), a significant positive correlation was observed between weak word usage and GSR peak changes between the interview and feedback sessions, respectively (t(34) = 2.18, p = 0.03), with an R2 of 0.12. There was no significant difference in the relationship between the individuals with low anxiety (t(16) = 1.59, p = 0.13) or those with high anxiety (t(16) = 1.63, p = 0.12).

VI. Discussion

While good performance in a job interview is critical [80], more than 90% of Americans claim that they were nervous about certain components of the job interview [81]. Anxiety can impair cognitive function, social interactions, and others' impressions and judgments [6]. Therefore, its effect on interview performance is one of the most researched issues [6], [9], [82], [83]. Although video feedback is one of the most commonly used techniques for reducing anxiety, a previous study [84] indicated that when socially anxious individuals view video feedback, they tend to re-experience some of their original feelings and rate their behavior more negatively. Therefore, they suggested cognitive preparation before viewing the video, such as if it was of a stranger.

We hypothesized that providing dissimilar avatar-based feedback would assist interviewees in associating information obtained during feedback to the avatar and objectively judging themselves as if the person being evaluated in the video is a stranger without any direct instructions. While we anticipated that avatar feedback would positively affect action identities, anxiety management, and interview performance, we did not expect it to have the same effects on mood and behavior in people with varying anxiety levels. In the following sections, we first discuss the overall effect of avatar feedback on the perception and mood during the feedback session and behavior and performance during the feedback session and subsequent interview session. Then, we analyze how avatar feedback affected mood, behavior, and performance based on the participants' initial anxiety levels.

A. General Discussion

Focusing on the general effects of the avatar-based feedback mode, our findings showed that higher-level action identities and pleasant moods increased significantly following the feedback session, only for the avatar-based feedback group. Furthermore, anxiety peaks decreased during the feedback session in the second half, only for the avatar group. Regarding performance change, individuals in the avatar-based feedback group exhibited higher pitch on average, used fewer weak words, and judged themselves marginally better in the second interview than participants in the self-feedback group. However, we could not detect any difference in smile ratios, which could have been due to this

study's short interview sessions, small sample size, or because the interviewer was a programmed agent rather than an actual human.

Behavior changes in participants who received the avatar-based feedback can be explained by an increase in anxiety in the control group during the feedback session and the fact that socially anxious individuals have a tendency against self-disclosure [85], which could jeopardize a successful job interview. Another possible explanation for these overall findings is that the avatar group initially felt more anxious due to the disparity between their recorded behavior and the avatar's appearance. Therefore, compared to the control group, the decrease in anxiety peaks could be attributed to the higher anxiety value in the first half rather than the lower value in the second half. Conversely, our findings in the following section revealed a negative relationship between anxiety change during the feedback period and the self-rating performance change from the first to the second interview session. This demonstrates the functional contribution of avatar feedback, such as subjects in the avatar group becoming more confident and excited during the second interview by anxiety reduction during the feedback period. Consequently, these findings may support our hypothesis that an avatar with a dissimilar appearance generally affects action identities, induce a positive mood, and reduces anxiety during job interview practice sessions, which can result in better performance and self-rating during a subsequence interview session.

We can expand our theory based on findings from previous studies, which have demonstrated that people with higher construal levels can think of more benefits for a project and are more likely to continue as a result of negative feedback [86]. In contrast, people move to lower-level identities when the task is difficult, or complex, or when their performance is disrupted [34]. Furthermore, using real-time camera filters can decrease the sense of embodiment compared to no filter [87]. Similarly, as evidenced by previous research [60], [87], [88], [89], digital representations can impact sensory perception, behavior, and cognitive abilities. Therefore, it is reasonable to expect that feedback received in avatar mode can affect action perception and subsequent performance.

B. Influence of Feedback on Action Identities, Mood, Behavior, and Performance: A Comparison of Participants With Different Initial Anxiety Levels

For participants with low initial anxiety levels, avatar feedback resulted in higher levels of action identification, and a modest increase in positive mood; however, no decrease in anxiety was observed during the feedback session compared to the control group. Participants also rated their performance marginally higher than the control group during the second interview, considering higher-level action identities during the feedback session positively correlated with changes in the self-reported performance evaluation during the second interview session. Additionally, their performance improved regarding vocal pitch, which was also positively correlated with higher levels of action identification. However, the situation was different for highly anxious individuals. First, we observed that individuals with higher anxiety levels tend to perceive their performance during the first interview session more negatively. It was found that neither the action identity nor the mood changed during the feedback session; however, they experienced fewer GSR peaks. Moreover, although highly anxious individuals who received avatar feedback did not rate their performance higher during the second interview session than the control group, they used slightly fewer weak words during the second interview session.

These differences in performance can be attributed to the opposing views of the two groups on job interviews. Highly anxious individuals generally rated their first interview performance as worse; therefore, it can be assumed that the job interview was a relatively difficult task for them. Vallacher et al. (1992), argued that optimal performance can be achieved by processing easy tasks at higher-level identities and difficult tasks at lower-level action identities [36]. Consequently, because individuals with low initial levels of anxiety did not have to deal with anxiety and perceived the interview as a relatively simple task, identifying actions at a higher level may have aided their interview performance. However, this was not the case for highly anxious individuals. Thus, perceiving actions at a higher level helped less anxious individuals in the avatar group to view their performance objectively and be more resistant to negative feedback [34]. This resulted in a happier mood during the feedback session, which in turn led to better self-performance and evaluation during the second interview session. Yet, while the dissimilar avatar decreased anxiety during the feedback session, it failed to affect action identification for highly anxious individuals.

Another explanation for our results is that receiving the avatar feedback did not necessarily result in higher action identities, but the video feedback caused lower action identities, especially for low-anxious participants, resulting in a more analytical and less positive mindset, lower self-rating of performance, and lower pitch in the second interview. This effect was more prominent for low-anxiety subjects, as high-anxiety subjects had more room for improvement even with the video feedback. This can be corroborated by a previous study that argued that participants who highly underrated themselves in an initial speech tended to benefit more from video feedback [90].

C. Limitations

While we primarily discussed our findings based on action identities and anxiety reduction, we should consider three alternative explanations. First is the presence of an avatar mapped to participants' actual facial expressions, which could have resulted in a positive mood, a pleasant atmosphere, and mood modulation. The second possibility is role reversal and taking turns. The positive effects of role reversal on empathy, stress control, and communication are well established [91], [92], [93]. Owing to the closeness of avatars' appearance with agents, participants in the avatar condition could have felt some level of role reversal

during the second interview session, resulting in closeness and empathy with the interviewer, and better performance. Third is the novelty effect, based on which subjects could respond stronger to treatment due to increased interest in the used novel technology. Notably, while the majority of the participants were already familiar with avatars, the human-like avatar mimicking their behavior in a job interview setting was a new experience for them. This may have briefly altered their behavior and outcome. Future research can investigate comparable effects that occur in people who are accustomed to using avatars. Another limitation is the focus on avatars to create a third-person perspective without exploring other AI video face-swapping technologies such as Deepfacelab [94]. This decision was made due to the time-consuming data collection and potential ethical concerns surrounding deepfake technologies at the study's inception. Therefore, we chose avatars considering they offered a quicker and more accessible option, thereby avoiding biases associated with recognizable individuals. Another point to consider is that the categorization of high/low anxiety in this study is merely relative within our tested population, with a notable deviation in the median MASI score (16.8) compared to Julie McCarthy's original study [66], which reported a mean score of 14.75. Finally, it should be noted that we did not consider the impact of gender, age, familiarity with technology, or cultural differences, which should be investigated in future studies.

VII. Conclusion

This study attempted to identify whether dissimilar avatars could modify action identification or anxiety during feedback sessions to improve performance during job interviews. Additionally, we explored the potential use of dissimilar avatars to assist participants in connecting feedback they received to an external self and conducting an unbiased evaluation of themselves. Finally, we pointed out the potential for avatars to affect people's moods, behavior, and performances during job interviews.

The results of this study indicate that compared to video feedback, avatar-based feedback can positively impact action identities, mood, and anxiety management during feedback sessions, thereby improving people's performance during job interviews. In particular, we found that higher levels of action identities and positive mood significantly increased, while anxiety peaks decreased during the feedback session in the second half for the avatar group. Additionally, we observed a negative relationship between the anxiety change during the feedback period and the self-rated performance change from the first to the second interview, which indicated that reducing anxiety during the feedback period can lead to greater confidence and enthusiasm during a subsequent interview session. Furthermore, participants in the avatar group had a higher pitch on average, used fewer weak words, and rated themselves marginally better than participants in the self-feedback group in the second interview.

However, it is essential to note that the effects of avatar feedback on mood and performance varied depending on the participants' initial anxiety levels. For those with low initial anxiety, avatar-based feedback resulted in higher-level action identities and higher vocal pitch than the control group, with only a marginal improvement in positive mood and word usage. On the other hand, for participants with high initial anxiety, avatar-based feedback did not significantly change the mood but decreased GSR peaks and marginally improved performance in terms of word usage. These findings indicate that avatar-based feedback may be more effective than video feedback for individuals with low initial anxiety levels because they can think more objectively and positively about the received avatar-based feedback without overthinking their actions. However, for those with high initial anxiety, avatar-based feedback may help reduce anxiety and improve performance without significantly changing subjective feelings about the performance.

In conclusion, this study supports using avatar-based feedback as a promising tool for managing anxiety and improving performance during job interviews. Future research could further explore the mechanisms behind these effects and examine the potential benefits of avatar-based feedback in other contexts where anxiety may be a factor. Although this study focused on job interviews, we believe our findings can be used to develop automated coaching to help people deal with anxiety in tasks that involve stress, require communication, or self-evaluation. Furthermore, we believe that the insights from this study have broader value for those working to advance emotional measuring methods in various real-world applications. Finally, by understanding the factors that influence anxiety and emotional responses through avatar-based interventions, future studies can develop more sophisticated tools for a wide range of real-life circumstances. Such advancements may encompass areas such as communication skills training, public speaking coaching, and social anxiety interventions.

References

[1] Hoque, M. C., J.-C. Martin, B. Mutlu, and R. W. Picard, "Mach: My automated conversation coach," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2013, pp. 697–706.

[2] M. A. Campion, J. E. Campion, and J. P. Hudson Jr, "Structured interviewing: A note on incremental validity and alternative question types.," *J. Appl. Psychol.*, vol. 79, no. 6, 1994, Art. no. 998.

[3] W. H. Wiesner and S. F. Cronshaw, "A meta-analytic investigation of the impact of interview format and degree of structure on the validity of the employment interview," *J. Occup. Psychol.*, vol. 61, no. 4, pp. 275–290, 1988.

[4] W. Rahman, S. Mahbub, A. Salekin, M. K. Hasan, and E. Hoque, "HirePreter: A framework for providing fine-grained interpretation for automated job interview analysis," in *Proc. 9th Int. Conf. Affect. Comput. Intell. Interact. Workshops Demos*, 2021, pp. 1–5.

[5] J. M. Cheek and S. R. Briggs, "Shyness as a personality trait," *Shyness Embarrassment: Perspectives Social Psychol.*, pp. 315–337, 1990. [Online]. Available: https://psycnet.apa.org/record/1990-98384-011

[6] D. M. Powell, D. J. Stanley, and K. N. Brown, "Meta-analysis of the relation between interview anxiety and interview performance," *Can. J. Behav. Sci./Revue Canadienne Des Sci. Du Comportement*, vol. 50, no. 4, 2018, Art. no. 195.

[7] R. G. Heimberg, K. E. Keller, and T. A. Peca-Baker, "Cognitive assessment of social-evaluative anxiety in the job interview: Job interview self-statement schedule.," *J. Counseling Psychol.*, vol. 33, no. 2, 1986, Art. no. 190.

[8] A. R. Feiler and D. M. Powell, "Behavioral expression of job interview anxiety," *J. Bus. Psychol.*, vol. 31, no. 1, pp. 155–171, 2016.

[9] C. J. Budnick, E. M. Anderson, A. M. Santuzzi, A. J. Grippo, and L. Matuszewich, "Social anxiety and employment interviews: Does nonverbal feedback differentially predict cortisol and performance?," *Anxiety Stress Coping*, vol. 32, no. 1, pp. 67–81, 2019.

[10] K. Anderson et al., "The TARDIS framework: Intelligent virtual agents for social coaching in job interviews," in *Proc. Int. Conf. Adv. Comput. Entertainment Technol.*, 2013, pp. 476–491.

[11] R. M. Aysina, G. I. Efremova, Z. A. Maksimenko, and M. v Nikiforov, "Using a computer simulation to improve psychological readiness for job interviewing in unemployed individuals of pre-retirement age," *Europe's J. Psychol.*, vol. 13, no. 2, 2017, Art. no. 251.

[12] M. J. Smith et al., "Virtual reality job interview training and 6-month employment outcomes for individuals with schizophrenia seeking employment," *Schizophrenia Res.*, vol. 166, no. 1–3, pp. 86–91, 2015.

[13] M. J. Smith et al., "Virtual reality job interview training for veterans with posttraumatic stress disorder," *J. Vocational Rehabil.*, vol. 42, no. 3, pp. 271–279, 2015.

[14] S. Provoost et al., "Embodied conversational agents in clinical psychology: A scoping review," *J. Med. Internet Res.*, vol. 19, no. 5, 2017, Art. no. e6553.

[15] N. Takeuchi and T. Koda, "Initial assessment of job interview training system using multimodal behavior analysis," in *Proc. 9th Int. Conf. Hum.-Agent Interaction*, 2021, pp. 407–411.

[16] Ö. Ayduk and E. Kross, "From a distance: Implications of spontaneous self-distancing for adaptive self-reflection.," *J. Pers. Social Psychol.*, vol. 98, no. 5, 2010, Art. no. 809.

[17] E. Kross, D. Gard, P. Deldin, J. Clifton, and O. Ayduk, "'Asking why' from a distance: Its cognitive and emotional consequences for people with major depressive disorder," *J. Abnorm. Psychol.*, vol. 121, no. 3, 2012, Art. no. 559.

[18] R. R. Vallacher and D. M. Wegner, "What do people think they're doing? Action identification and human behavior," *Psychol. Rev.*, vol. 94, no. 1, 1987, Art. no. 3.

[19] L. K. Libby, E. M. Shaeffer, and R. P. Eibach, "Seeing meaning in action: A bidirectional link between visual perspective and action identification level," *J. Exp. Psychol.: Gen.*, vol. 138, no. 4, 2009, Art. no. 503.

[20] E. R. Watkins, N. J. Moberly, and M. L. Moulds, "When the ends outweigh the means: Mood and level of identification in depression," *Cogn. Emotion*, vol. 25, no. 7, pp. 1214–1227, 2011.

[21] N. Yee and J. Bailenson, "The Proteus effect: The effect of transformed self-representation on behavior," *Hum. Commun. Res.*, vol. 33, no. 3, pp. 271–290, 2007.

[22] N. Yee and J. N. Bailenson, "The difference between being and seeing: The relative contribution of self-perception and priming to behavioral changes via digital self-representation," *Media Psychol.*, vol. 12, no. 2, pp. 195–209, 2009.

[23] N. Yee and J. N. Bailenson, "Walk a mile in digital shoes: The impact of embodied perspective-taking on the reduction of negative stereotyping in immersive virtual environments," in *Proc. PRESENCE*, vol. 24, 2006, Art. no. 26.

[24] T. C. Peck, S. Seinfeld, S. M. Aglioti, and M. Slater, "Putting yourself in the skin of a black avatar reduces implicit racial bias," *Conscious Cogn.*, vol. 22, no. 3, pp. 779–787, 2013.

[25] G. Gorisse, O. Christmann, S. Houzangbe, and S. Richir, "From robot to virtual doppelganger: Impact of visual fidelity of avatars controlled in third-person perspective on embodiment and behavior in immersive virtual environments," *Front. Robot. AI*, vol. 6, 2019, Art. no. 8.

[26] S. J. Ahn and J. N. Bailenson, "Self-endorsing versus other-endorsing in virtual environments," *J. Advertising*, vol. 40, no. 2, pp. 93–106, 2011.

[27] B. van der Hoort, A. Guterstam, and H. H. Ehrsson, "Being Barbie: The size of one's own body determines the perceived size of the world," *PLoS One*, vol. 6, no. 5, 2011, Art. no. e20195.

[28] D. Banakou, R. Groten, and M. Slater, "Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes," in *Proc. Nat. Acad. Sci. USA*, vol. 110, no. 31, pp. 12846–12851, 2013.

[29] M. Slater, B. Spanlang, M. v Sanchez-Vives, and O. Blanke, "First person experience of body transfer in virtual reality," *PLoS One*, vol. 5, no. 5, 2010, Art. no. e10564.

[30] D. Banakou, P. D. Hanumanthu, and M. Slater, "Virtual embodiment of white people in a black virtual body leads to a sustained reduction in their implicit racial bias," *Front. Hum. Neurosci.*, vol. 10, 2016, Art. no. 601.

[31] S. A. Osimo, R. Pizarro, B. Spanlang, and M. Slater, "Conversations between self and self as Sigmund Freud—A virtual body ownership paradigm for self counselling," *Sci. Rep.*, vol. 5, 2015, Art. no. 13899.

[32] I. Hudson and J. Hurter, "Avatar types matter: Review of avatar literature for performance purposes," in *Proc. Int. Conf. Virtual Augmented Mixed Reality*, 2016, pp. 14–21.

[33] J. N. Bailenson and A. C. Beall, "Transformed social interaction: Exploring the digital plasticity of avatars," in *Avatars At Work and Play*, Berlin, Germany: Springer, 2006, pp. 1–16.

[34] R. R. Vallacher and D. M. Wegner, "Levels of personal agency: Individual variation in action identification.," *J. Pers. Social Psychol.*, vol. 57, no. 4, 1989, Art. no. 660.

[35] R. R. Vallacher, D. M. Wegner, and J. Frederick, "The presentation of self through action identification," *Social Cogn.*, vol. 5, no. 3, pp. 301–322, 1987.

[36] R. R. Vallacher, D. M. Wegner, S. C. McMahan, J. Cotter, and K. A. Larsen, "On winning friends and influencing people: Action identification and self-presentation success," *Social Cogn.*, vol. 10, no. 3, pp. 335–355, 1992.

[37] M.-A. Crocq, "A history of anxiety: From Hippocrates to DSM," *Dialogues Clin. Neurosci.*, vol. 17, pp. 319–325, 2022.

[38] S. L. Rynes, R. D. Bretz Jr, and B. Gerhart, "The importance of recruitment in job choice: A different way of looking," *Personnel Psychol.*, vol. 44, no. 3, pp. 487–521, 1991.

[39] D. B. Jones and J. W. Pinkney, "An Exploratory Assessment of the Sources of Job-Interviewing Anxiety in College Students.," *J. College Student Develop.*, vol. 30, no. 6, pp. 553–560, 1989.

[40] J. Ayres, T. Keereetaweep, P.-E. Chen, and P. A. Edwards, "Communication apprehension and employment interviews," *Commun. Educ.*, vol. 47, pp. 1–17, 1998.

[41] T. L. Rodebaugh, R. M. Holaway, and R. G. Heimberg, "The treatment of social anxiety disorder," *Clin. Psychol. Rev.*, vol. 24, no. 7, pp. 883–908, 2004.

[42] R. M. Rapee and K. Hayman, "The effects of video feedback on the self-evaluation of performance in socially anxious subjects," *Behav. Res. Ther.*, vol. 34, no. 4, pp. 315–322, 1996.

[43] A. G. Harvey, D. M. Clark, A. Ehlers, and R. M. Rapee, "Social anxiety and self-impression: Cognitive preparation enhances the beneficial effects of video feedback following a stressful social task," *Behav. Res. Ther.*, vol. 38, no. 12, pp. 1183–1192, 2000.

[44] H.-Y. Kim, L.-G. Lundh, and A. Harvey, "The enhancement of video feedback by cognitive preparation in the treatment of social anxiety. A single-session experiment," *J. Behav. Ther. Exp. Psychiatry*, vol. 33, no. 1, pp. 19–37, 2002.

[45] I. M. Aderka, "Factors affecting treatment efficacy in social phobia: The use of video feedback and individual vs. group formats," *J. Anxiety Disord.*, vol. 23, no. 1, pp. 12–17, 2009.

[46] ZENKIGEN Co., Ltd., Accessed: Apr. 18, 2023. [Online]. Available: https://harutaka.jp/

[47] Ltd MIDAS IT Japan Co., Accessed: Apr. 18, 2023. [Online]. Available: https://www.inair.co.jp/

[48] T. Baur, I. Damian, P. Gebhard, K. Porayska-Pomsta, and E. André, "A job interview simulation: Social cue-based interaction with a virtual character," in *Proc. Int. Conf. Social Comput.*, 2013, pp. 220–227.

[49] N. Goda, K. Ishihara, and T. Kojiri, "Job Interview Support System Based on Analysis of Nonverbal Behavior," *IEICE Tech. Rep.*, vol. 116, no. 517, pp. 25–30, 2017.

[50] M. Langer, C. J. König, P. Gebhard, and E. André, "Dear computer, teach me manners: Testing virtual employment interview training," *Int. J. Selection Assessment*, vol. 24, no. 4, pp. 312–323, 2016.

[51] H. Tanaka et al., "Automated social skills trainer," in *Proc. 20th Int. Conf. Intell. User Interfaces*, 2015, pp. 17–27.

[52] A. Garcia-Palacios, C. Botella, H. Hoffman, and S. Fabregat, "Comparing acceptance and refusal rates of virtual reality exposure vs. in vivo exposure by patients with specific phobias," *Cyberpsychol. Behav.*, vol. 10, no. 5, pp. 722–724, 2007.

[53] A. Garcia-Palacios, H. G. Hoffman, S. Kwong See, A. M. Y. Tsai, and C. Botella, "Redefining therapeutic success with virtual reality exposure therapy," *CyberPsychol. Behav.*, vol. 4, no. 3, pp. 341–348, 2001.

[54] P. L. Anderson, E. Zimand, L. F. Hodges, and B. O. Rothbaum, "Cognitive behavioral therapy for public-speaking anxiety using virtual reality for exposure," *Depress Anxiety*, vol. 22, no. 3, pp. 156–158, 2005.

[55] S. R. Harris, R. L. Kemmerling, and M. M. North, "Brief virtual reality therapy for public speaking anxiety," *Cyberpsychol. Behav.*, vol. 5, no. 6, pp. 543–550, 2002.

[56] J. Bissonnette, F. Dubé, M. D. Provencher, and M. T. Moreno Sala, "Virtual reality exposure training for musicians: Its effect on performance anxiety and quality," *Med. Problems Performing Artists*, vol. 30, no. 3, pp. 169–177, 2015.

[57] H. S. Wallach, M. P. Safir, and M. Bar-Zvi, "Virtual reality cognitive behavior therapy for public speaking anxiety: A randomized clinical trial," *Behav. Modification*, vol. 33, no. 3, pp. 314–338, 2009.
[58] L. Aymerich-Franch and J. Bailenson, "The use of doppelgangers in virtual reality to treat public speaking anxiety: A gender comparison," in *Proc. Int. Soc. Presence Res. Annu. Conf.*, 2014, pp. 173–186.
[59] L. Aymerich-Franch, R. F. Kizilcec, and J. N. Bailenson, "The relationship between virtual self similarity and social anxiety," *Front. Hum. Neurosci.*, vol. 8, 2014, Art. no. 944.
[60] D. Banakou, S. Kishore, and M. Slater, "Virtually being Einstein results in an improvement in cognitive task performance and a decrease in age bias," *Front. Psychol.*, vol. 9, 2018, Art. no. 917.
[61] H. Tanaka, H. Negoro, H. Iwasaka, and S. Nakamura, "Embodied conversational agents for multimodal automated social skills training in people with autism spectrum disorders," *PLoS One*, vol. 12, no. 8, 2017, Art. no. e0182151.
[62] J. Wang, H. Yang, R. Shao, S. Abdullah, and S. S. Sundar, "Alexa as coach: Leveraging smart speakers to build social agents that reduce public speaking anxiety," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2020, pp. 1–13.
[63] Z. Fang, L. Cai, and G. Wang, "MetaHuman creator the starting point of the metaverse," in *Proc. Int. Symp. Comput. Technol. Inf. Sci.*, 2021, pp. 154–157.
[64] EPIC GAMES DEV COMMUNITY, "Animating with LiveLink." Accessed: Apr. 18, 2023. [Online]. Available: https://dev.epicgames.com/documentation/en-us/metahuman/animating-metahumans-with-livelink-in-unreal-engine .
[65] M. Panaghiston and K. Davis, "Voice activity detection javascript." Accessed: Apr. 18, 2023. [Online]. Available: https://github.com/kdavis-mozilla/vad.js/
[66] J. McCarthy and R. Goffin, "Measuring job interview anxiety: Beyond weak knees and sweaty palms," *Personnel Psychol.*, vol. 57, no. 3, pp. 607–637, 2004.
[67] P. J. Lang et al., "International affective picture system (IAPS): Technical manual and affective ratings," *NIMH Center Study Emotion Attention*, vol. 1, no. 39/58, 1997, Art. no. 3.
[68] A. Burns et al., "SHIMMER^{TM}: An extensible platform for physiological signal capture," in *Proc. IEEE Annu. Int. Conf. Eng. Med. Biol.*, 2010, pp. 3759–3762.
[69] J. Bakker, M. Pechenizkiy, and N. Sidorova, "What's your current stress level? Detection of stress patterns from GSR sensor data," in *Proc. IEEE 11th Int. Conf. Data Mining Workshops*, 2011, pp. 573–580.
[70] D. Makowski et al., "NeuroKit2: A Python toolbox for neurophysiological signal processing," *Behav. Res. Methods*, vol. 53, no. 4, pp. 1689–1696, 2021.
[71] P. V. Washburn and M. D. Hakel, "Visual cues and verbal content as influences on impressions formed after simulated employment interviews," *J. Appl. Psychol.*, vol. 58, no. 1, 1973, Art. no. 137.
[72] M. LaFrance, *Why Smile?: The Science Behind Facial Expressions*. New York, NY, USA: WW Norton, 2011.
[73] T. DeGroot and J. Gooty, "Can nonverbal cues be used to make meaningful personality attributions in employment interviews?," *J. Bus. Psychol.*, vol. 24, no. 2, pp. 179–192, 2009.
[74] J. G. Hollandsworth Jr, R. Kazelskis, J. Stevens, and M. E. Dressel, "Relative contributions of verbal, articulative, and nonverbal communication to employment decisions in the job interview setting," *Personnel Psychol.*, vol. 32, no. 2, pp. 359–367, 1979.
[75] T. Baltrušaitis, P. Robinson, and L.-P. Morency, "Openface: An open source facial behavior analysis toolkit," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, 2016, pp. 1–10.
[76] P. Ekman, W. v Friesen, and J. C. Hager, "The facial action coding system Second edition," London, U.K: Weidenfeld & Nicolson, 2002.
[77] Y. Jadoul, B. Thompson, and B. de Boer, "Introducing parselmouth: A python interface to praat," *J-Phone*, vol. 71, pp. 1–15, 2018.
[78] M. Hazel, C. McMahon, and N. Schmidt, "Immediate feedback: A means of reducing distracting filler words during public speeches," *Basic Commun. Course Annu.*, vol. 23, no. 1, 2011, Art. no. 6.
[79] E. Loper and S. Bird, "NLTK: The natural language toolkit," 2002, *arXiv:cs/0205028*.
[80] A. I. Huffcutt, C. H. van Iddekinge, and P. L. Roth, "Understanding applicant behavior in employment interviews: A theoretical model of interviewee performance," *Hum. Resource Manage. Rev.*, vol. 21, no. 4, pp. 353–367, 2011.
[81] J. Horn, "Why Americans fear the job interview," *San Diego Union-Tribune*, 2013. Accessed: Apr. 18, 2023. [Online]. Available: https://www.sandiegouniontribune.com/business/economy/sdut-job-interview-economy-employment-harris-everest-2013aug20-story.html
[82] T. A. Brown, D. H. Barlow, and P. A. DiNardo, *Anxiety Disorders Interview Schedule Adult Version: Client Interview Schedule*. New York, NY, USA: Graywind Publications Incorporated, 1994.
[83] I. Y. Zhang, D. M. Powell, and S. Bonaccio, "The role of fear of negative evaluation in interview anxiety and social-evaluative workplace anxiety," *Int. J. Selection Assessment*, vol. 30, no. 2, pp. 302–310, 2022.
[84] E. Warnock-Parkes, J. Wild, R. Stott, N. Grey, A. Ehlers, and D. M. Clark, "Seeing is believing: Using video feedback in cognitive therapy for social anxiety disorder," *Cogn. Behav. Pract.*, vol. 24, no. 2, pp. 245–255, May 2017, doi: 10.1016/j.cbpra.2016.03.007.
[85] A. Wells, D. M. Clark, P. Salkovskis, J. Ludgate, A. Hackmann, and M. Gelder, "Social phobia: The role of in-situation safety behaviors in maintaining anxiety and negative beliefs," *Behav. Ther.*, vol. 26, no. 1, pp. 153–161, 1995.
[86] N. Benschop, A. L. P. Nuijten, M. Keil, K. I. M. Rohde, J. S. Lee, and H. R. Commandeur, "Construal level theory and escalation of commitment," *Theory Decis.*, vol. 91, no. 1, pp. 135–151, 2021.
[87] J. Leong et al., "Exploring the use of real-time camera filters on embodiment and creativity," in *Proc. Extended Abstr. CHI Conf. Hum. Factors Comput. Syst.*, 2021, pp. 1–7.
[88] J. Guegan, S. Buisine, F. Mantelet, N. Maranzana, and F. Segonds, "Avatar-mediated creativity: When embodying inventors makes engineers more creative," *Comput Hum. Behav*, vol. 61, pp. 165–175, 2016.
[89] T. C. Peck, J. J. Good, and K. A. Bourne, "Inducing and mitigating stereotype threat through gendered virtual body-swap illusions," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2020, pp. 1–13.
[90] T. L. Rodebaugh and D. L. Chambless, "The effects of video feedback on self-perception of performance: A replication and extension," *Cogn. Ther. Res.*, vol. 26, pp. 629–644, 2002.
[91] I. L. Janis and B. T. King, "The influence of role playing on opinion change," *J. Abnorm. Social Psychol.*, vol. 49, no. 2, 1954, Art. no. 211.
[92] D. W. Johnson, "Role reversal: A summary and review of the research," *Int. J. Group Tensions*, vol. 1, pp. 318–334, 1971.
[93] H. Abeditehrani, C. Dijk, M. D. Neyshabouri, and A. Arntz, "Beneficial effects of role reversal in comparison to role-playing on negative cognitions about other's judgments for social anxiety disorder," *J. Behav. Ther. Exp. Psychiatry*, vol. 70, 2021, Art. no. 101599.
[94] K. Liu, I. Perov, D. Gao, N. Chervoniy, W. Zhou, and W. Zhang, "Deepfacelab: Integrated, flexible and extensible face-swapping framework," *Pattern Recognit.*, vol. 141, 2023, Art. no. 109628.

Sarinasadat Hosseini received the BSc degree in electrical engineering from the Shahid Beheshti University, Iran, in 2016, the MSc degree in computer science from the Tokyo Institute of Technology, Japan, in 2019, and the PhD degree in computer science from the Tokyo Institute of Technology, Japan, in 2023. She is currently working with Panasonic Corporation, Japan. Her research interests include human-computer interaction, multimedia and signal processing, and social robots.

Jingyu Quan received the bachelor's degree in intelligence science and technology from Central South University, China, in 2019, and the master's degree in artificial intelligence from the Tokyo Institute of Technology, Japan, in 2022. He is currently working toward the PhD degree in artificial intelligence with the Tokyo Institute of Technology. He is interested in using and understanding machine learning technology to improve human-human interaction. Multimodal affective computing, explainable artificial intelligence, and interpersonal synchronization are central points to his research activities.

Xiaoqi Deng received the BS degree in biology from Nanjing University, China, in 2009, and the MS degree in finance from EDHEC Business School, France, in 2014. She is currently working toward the PhD degree in computer science with the Tokyo Institute of Technology, Japan. Her research interest includes human-computer interaction, nonverbal communication behaviors, and cognitive science.

Yoshihiro Miyake (Member, IEEE) received the BS, MS, and PhD degrees in pharmaceutical science from the University of Tokyo, Tokyo, Japan, in 1983, 1985, 1989, respectively. He is a professor with the School of Computing, Tokyo Institute of Technology. From 1989 to 1996, he was an associate professor with the Kanazawa Institute of Technology. In 1996, he joined the Tokyo Institute of Technology. He is engaged in the fundamental research of Human Communication Science and Human-Computer Interaction. He is now studying the design principle of co-creation system based on the embodied interaction between humans and artifacts. He is also developing gait assist robot WALK-MATE by using interpersonal synchrony. He is a member of ACM, IEICE, and SICE.

Takayuki Nozawa (Member, IEEE) received the BSc, MSc, and PhD degrees in science from the Tokyo Institute of Technology, Japan, in 1997, 1999, and 2002, respectively. He was affiliated with Kyoto Sangyo University, the National Institution for Academic Degrees and University Evaluation, Japan Women's University, Tokyo University of Agriculture and Technology, Tohoku University, and Tokyo Institute of Technology. He is currently a professor with the Department of Intellectual Information Systems Engineering, Faculty of Engineering, University of Toyama, Japan. His research interests include cognitive neuroscience, communication science, complex systems, human-computer interaction, multimodal affective computing, and educational technologies.