

Using a virtual reality interview simulator to explore factors influencing people's behavior

Xinyi Luo*	Yuyang Wang†	Lik-Hang Lee‡	Zihan Xing
University of Electronic Science and Technology of China	Hong Kong University of Science and Technology (Guangzhou)	The Hong Kong Polytechnic University	Beijing Normal University Hong Kong Baptist University United International College
Shan Jin	Boya Dong	Yuanyi Hu	Zeming Chen
Hong Kong University of Science and Technology (Guangzhou)	South China University of Technology	Guangdong Medical University	South China University of Technology
Jing Yan	Pan Hui		
Taiyuan University of Technology	Hong Kong University of Science and Technology (Guangzhou)		

ABSTRACT

Virtual reality interview simulator (VRIS) provides an effective and manageable approach for candidates prone to being very nervous during interviews, yet, the major anxiety-inducing elements remain unknown. During an interview, the anxiety levels, overall experience, and performance of interviewees might be affected by various circumstances. By analyzing electrodermal activity and questionnaire, we investigated the influence of five variables: (I) *Realism*; (II) *Question type*; (III) *Interviewer attitude*; (IV) *Timing*; and (V) *Preparation*. As such, an orthogonal design $L_8(4^1 \times 2^4)$ with eight experiments (OA_8 matrix) was implemented, in which 19 college students took part in the experiments. Considering the anxiety, overall experience, and performance of the interviewees, results indicate that *Question type* plays a major role; secondly, *Realism*, *Preparation*, and *Interviewer attitude* all have some degree of influence; lastly, *Timing* have little to no impact. Specifically, professional interview questions elicited a greater degree of anxiety than personal ones among the categories of interview questions. This work contributes to our understanding of anxiety-stimulating factors during job interviews in virtual reality and provides cues for designing future VRIS.

Index Terms: I.3.7 [Three-Dimensional Graphics and Realism]: Virtual reality; H.5.2 [User Interfaces]: Training, help, and documentation

1 INTRODUCTION

The global pandemic in recent years has aggravated the situation where many companies are hiring fewer employees, which has led to a series of anxieties among college students [48], one of which is interview anxiety disorder (IAD)-anxiety manifests itself in the form of speech disturbances, socially inappropriate behaviors as well as other nervous jitters [44]. Furthermore, past research has shown that as interview anxiety increases, people tend to develop more protective self-presentational tactics, making it difficult for interviewees to perform well and decreasing competitiveness [19]. Virtual reality has become a common tool in the therapeutic field, such as substance [3], high-functioning autism [14,39] and eating disorders [12]. Virtual reality exposure therapy (VRET) is an important treatment

for various anxiety disorders. By mimicking social circumstances for social distress patients [16] and using computer-generated virtual scenes [38], patients can be exposed to an environment with virtual social situations and interactions that target diverse social fears in a controllable way. For example, one study indicated the utility of VR in inducing stressful reactions through a combination of stressors for healthy subjects [40]. However, the factors that actually influence an interviewee's anxiety, overall experience, and interview performance in a VRET interview are still unclear and poorly understood, especially when multiple factors are assembled in one system; thus, our research serves as a pioneer for such an investigation.

So far, researchers have also conducted extensive research on systems for VR interview training and evaluation. For example, a virtual job interviewing practice system has been designed for high-anxiety populations like people with Autism Spectrum Disorder and former convicts [28], a VR-based job interview training platform has been developed for autistic individuals to practice interviewing skills in a less anxiety-inducing virtual context [2], an agent-based VR training and multidimensional evaluation system has been created for introverted college students to cope with interview anxiety [36], a job training simulation environment has been presented for young people who are out of employment, education, or training with social cue recognition techniques [22].

Nevertheless, most studies today validate and evaluate an entire product without examining the multiple factors separately. Not only do we know very little about the significant elements that truly influence interview anxiety and the overall experience (e.g., cognitive load, discomfort [58, 59]), but we also need a solid understanding of how each factor affects the interviewee's external performance (e.g., verbal expressions, eye contact, body movements). For graduates to be competitive in the job market and for employers to pick prospective workers, they must demonstrate strong performance during job interviews and have practical anxiety management abilities. Therefore, a study on factors contributing to people's interview anxiety is required to create successful training programs and develop tailored therapy approaches.

In terms of anxiety, previous studies have found that visual display [43], interview questions [29], interviewer's attitude [42], timing [53] and preparation [20] can all have an impact on interviewee's anxiety. In addition to feelings of anxiety, other studies have investigated the impact of virtual reality on overall experience (e.g., discomfort, eyestrain, psychosocial stress). Some studies have found that apparatus has an impact on quality of experience, with HMDs experiencing higher eyestrain and visual discomfort than PCs [56]; some have found the V-TSST (virtual environment versions of the TSST) is effective at inducing psychosocial stress which can lead to poor physical and psychological health outcomes, though the

*e-mail: 2020090903013@std.uestc.edu.cn

†e-mail: yuyangwang@ust.hk

‡e-mail: lhleeac@connect.ust.hk

magnitude of this response is less than the traditional TSST [31]; Kothgassner et al. have indicated that the perceived social presence did not differ over time in the VR TSST conditions as the main effect of time was not significant [41].

Meanwhile, interviewees' performances have also been proven to be influenced by factors within a virtual environment. For example, prior studies have identified the relationship between the number of completed virtual interviews and improved interviewing skills or performance as the mechanism for getting a job offer [55]; the results of M Barreda-Ángeles et al. have shown that, compared to the neutral audience, the negative audience elicited increases in skin conductance level and heart rate variability, decreases in voice intensity, and a higher ratio of silent parts in the speech, as well as a more negative self-reported valence, higher anxiety, and lower social presence [6].

As such, we developed the VRIS, where the above factors were introduced and investigated within an orthogonal design to examine their significance separately.

In this paper, we question whether the above five factors (I) REALISM; (II) QUESTION TYPE; (III) INTERVIEWER ATTITUDE; (IV) TIMING; and (V) PREPARATION will indeed significantly influence interviewee's anxiety during a job interview. We hypothesized that all of these factors could potentially have a pronounced effect on interview anxiety. Consider the above mentioned, the current study proposed five research questions (RQs) with the following hypothesis (H):

1. RQ1: How do different interview questions affect the interviewee?

H1: Professional interview questions can cause more anxiety, worse overall experience, and poorer performance than personal questions.

2. RQ2: How does preparation for an interview affect the interviewee?

H2: Being unprepared for an interview can be more nervous and uncomfortable than being well prepared for the content of the interview.

3. RQ3: How does the timing of the answers to interview questions affect the interviewee?

H3: Timed answers can be more nerve-wracking than untimed answers, leading to worse performance.

4. RQ4: How do different levels of realism affect the interviewee?

H4: A more realistic scenario would make interviewees more nervous and reduce their eye contact with the interviewer.

5. RQ5: How does the interviewer's attitude affect the interviewee?

H5: Compared to an interviewer with a positive attitude, an interviewer with a negative attitude will elicit increases in skin conductance response, decreases in eye contact, a higher cognitive load during the interview, as well as more unsatisfactory performance.

We then conducted an orthogonal experiment design with eight different interview conditions using a mixed level $L_8(4^1 \times 2^4)$ orthogonal table including all five factors above to examine the significance of each factor on interview anxiety, overall experience, and performance. In addition, we measured the interviewee's electrodermal activity(EDA) during the interview, given that increased EDA has been associated with anxiety. Finally, in all eight conditions, we asked the interviewee to fill out a self-rated anxiety questionnaire and NASA-TLX criteria once the interview was completed. The

interviewer would also rate the interviewee's performance during the interview. Followed by the mixed-effects model and the associated post-hoc analysis for the data collected from questionnaires and electrodermal activity, we found that all five of the above factors had different levels of influence on interviewees' anxiety, overall experience and interview performance, among which TYPE OF INTERVIEW QUESTIONS had the most significant impact, in particular, the professional questions significantly increased the interviewee's anxiety, discomfort, electrodermal activity and cognitive workload.

The proposed work aims to identify anxiety-stimulating factors during job interviews in virtual reality and provides detailed insights into designing future VRIS.

2 RELATED WORK

Relevant prior work includes studies of psychotherapy, online interview systems, and virtual reality interview training. This section summarizes those works separately.

2.1 Psychotherapy in Virtual Reality

Since the inception of virtual reality, several psychotherapy pieces of research have been undertaken in virtual settings, and the idea of employing them to treat psychiatric illnesses has been investigated. As social concern about social phobia gradually increased, North et al. first used virtual reality exposure therapy for social phobia: they developed a virtual auditorium that could be triggered in real-time [50]. The audience and audio clips would respond to the experimenter's voice in the auditorium, prompting the experimenter to speak louder and more loudly. Based on feedback from the questionnaire results, the experiment demonstrated that Virtual Reality Exposure Therapy (VRET) effectively mitigated anxiety symptoms during presentations. Although the feedback from the experiment mainly was auditory information, it highlights the need for further research on the relationship between VRET and human psychology.

Further studies have implied that there are three preconditions for the treatment of anxiety disorders through VRET, including immersion, anxiety, and presence [34]. Parsons et al. reported 21 case studies [51] confirming that VRET can effectively treat arachnophobia (Garcia et al. [21]), flying phobia (Banos and Botella [5]), phobia of public places (Botella et al. [8]), acrophobia(Coelho et al. [13])

Some researchers have emphasized social cognition interaction training for autistic youngsters. Didehbani et al. designed a virtual reality social cognitive training to improve the social skills of children with ASD [14]. The research findings, which tested emotion detection, social attribution, attention, and executive functioning, revealed that the virtual reality platform successfully ameliorated the social deficits typical of ASD.

In addition to virtual reality social cognitive training for children with ASD, Burke et al. designed Virtual Interactive Training Agents (ViTA) to develop social skills and reduce anxiety in young people with ASD and other developmental disorders [11]. According to studies, experimenters who have received ViTA training are much better at recognizing their abilities, promoting themselves, self-promoting, self-advocating, and responding to situational queries. In contrast, our study relates virtual reality to job interviews, the most prevalent situation in business practice. In the era of the metaverse, we anticipate migrating job interviews to the virtual realm. At the same time, our research exposes user behaviors in the settings of psychological factors and virtual job interviews.

2.2 Virtual Reality and Online Interview Systems

Due to the global pandemic and rising economic costs, firms are progressively incorporating online video interviews into their hiring and appointment procedures. In addition, the rapid development of artificial intelligence (AI) has driven the applications of automatic scoring, where AI interview systems can score interviewees' hard

and soft skills based on the content of their answers, facial movements, eye contact, and speaking tones in the video [30], which successfully increase the effectiveness of interviews and assist professionals in assessing their candidates more quickly.

Online interview formats have driven the development of online interview simulation training. For example, Aysina et al. created Job Interview Simulation Training (JIST) to improve psychological preparation for job interviews among the pre-retirement unemployed [4]. The experiment showed that having interviewees practice interviews over and over in a stress-free environment made them much more psychologically ready for the actual interview(s), which could help demonstrate the relationship between JIST and increased re-employment among pre-retirement job seekers in the future.

Other studies have found that highly interactive virtual reality role-play training based on behavioral learning principles is more effective than traditional role-play training in training other types of interpersonal skills. Smith et al. developed a study to test a role-play simulation “virtual reality job interview training” (VR-JIT) for the feasibility and effectiveness of improving job-related interview content and interviewees’ performance-related interview skills in individuals with ASD [54]. Their findings demonstrate that VR-JIT can improve job interview skills in individuals with ASD. Among current online interview systems, our research investigates the viability of conducting job interview preparations in immersive settings. Additionally, we specifically focus on the effect of different factors on interviewees’ anxiety levels.

3 VRIS FRAMEWORK

3.1 Orthogonal experimental design

The orthogonal experimental design is an efficient method to study the effect of multiple factors within one system simultaneously compared to the conventional methods of studying each factor separately by selecting one of the variables to change its parameters and fixing the rest of the variables. It selects some representative combinations from a full-scale test according to the modern algebra of Galois theory [1, 35]. The orthogonal table based on orthogonality ensures that the effects of all factors are obtained with a minimum number of trials.

We identified five independent variables that potentially affect interviewee’s anxiety levels during a job interview in virtual reality: (I) LEVEL OF REALISM (4 levels); (II) TYPE OF INTERVIEW QUESTIONS (2 levels); (III) INTERVIEWER ATTITUDE (2 levels); (IV) TIMED OR UNTIMED ANSWERS (2 levels); and (V) WITH OR WITHOUT PREPARATION (2 levels). In order to determine the relative importance of these five variables and find out what factor most stimulates the interviewee’s anxiety, we constructed an orthogonal fractional factorial design to arrange the tests by using a mixed level $L_8(4^1 \times 2^4)$ orthogonal table with all these five variables in a total of eight sets of conditions, as shown in Table 1. For example, to conduct the seventh experimental group, the participants had to test with the Oculus Quest 2 HMD in a virtual job interview environment and be asked ten professional questions without preparation before the interview began. Each question will be timed 30 seconds for the answer by an interviewer with a negative attitude (e.g., passive body movements and negative verbal feedback).

The first independent variable (I) LEVEL OF REALISM corresponded to the visual display of the interviewer. We were aware that a video conference(e.g., through Skype, Tencent Meeting) would be considered quite “real” as an online meeting using video conferencing is generally a well-accepted form of a remote interview. In our context, “realism” refers to the level of immersion. Previous studies have found that higher visual display levels provoke more anxiety and the sense of presence [43]. Therefore, in order to create four different kinds of realism, we set four conditions representing a continuous spectrum of immersiveness, i.e., the least to the most

immersive, including an interviewer presented by video conference on a laptop computer (*PC*), a low-poly and cartoon-like 3D avatar representing the human interviewer (*VR1*), a realistic interviewer with a high fidelity 3D human avatar (*VR2*) and a face-to-face real human interviewer (*REAL*) (see Fig. 2). In the experiments under *PC* condition, interviewees conducted video interviews with live interviewers via video conference. In the experiments under *VR1* and *VR2* condition, animated sequences of the interviewer’s avatar were played automatically by the software in a VR headset, and a conversation with a live interviewer was conducted via voice conference. Each of the interviewers’ avatars had a full-body presentation, but each of the 19 interviewees only had both hands as a physical presence in the virtual environment. In the experiments under *REAL* condition, the interviewee had a face-to-face interview with a live interviewer (see Fig. 1).



Figure 1: Four experiment settings for job interview simulation under the conditions of (a) video-conferencing(*PC*), (b) cartoon VR (*VR1*), (c) realistic VR (*VR2*); (d) onsite meeting (*REAL*).

The second independent variable (II) TYPE OF INTERVIEW QUESTIONS corresponded to two categories of questions: professional inquiries and personnel interview questions, since the previous study has also shown that different types of interview questions have an impact on interviewee’s performance [29]. The professional inquiries consisted of the professional knowledge the interviewee has learned in college, which involved computer networking, operating systems, programming and algorithms, linear algebra, database, and principles of computer composition. Each question had the corresponding correct answers and would examine the interviewee’s memory, logical thinking ability, reaction speed, and mastery of professional knowledge. These professional questions were based on university final exams, internship interviews, job interviews, and interview questions for the graduate school review. Since there was no established standard for what kind of job the participants were applying for, the personnel questions used in the simulated interviews were general and typical job interview questions. The questions were divided into five categories: basic personal information, personality assessment, emotional control, organizing and planning skills, and creative questions. The pressure and difficulty of the questions in these five categories increased, and they were eventually divided into four sets of ten questions each, with a similar level of difficulty and no repetitions. The answers to these questions were mainly based on the interviewee’s review and summary of their experience and evaluation of themselves and were open-ended. Moreover, for each of the eight groups of interviews, we prepared different interview questions accordingly, so there were eight separate sets of interview questions in total, and each category(i.e., professional and personnel) had four sets of interview questions of the same difficulty with each

set including ten interview questions. All the interview questions were in the additional materials.

The third independent variable (III) INTERVIEWER ATTITUDE corresponded to the interviewer's attitude (mainly body language and tone of voice) and response to the interviewee's performance during the interview process. A previous study discovered that the participants exhibited more anxiety by the attitude of virtual avatars than the avatar's level of realism [42]. We designed two types of interviewers with positive and negative attitudes. Both interviewers would give interviewees real-time responses based on their performances. The positive interviewer would respond with positive feedback on the interviewee's answers (If the interviewee did well, the interviewer would respond "Excellent, exactly right." If the interviewee did not perform well, the interviewer would reply "It's okay, there is no rush, please take your time to think about it.") with positive animations (e.g., greeting, handshaking, listening with full attention, nodding, acknowledging, see Fig. 2). The negative interviewer would start the interview by emphasizing "I will only ask all the questions once and will not repeat them, so listen carefully." During an interview, the interviewer would give negative feedback on the interviewee's answers (If the interviewee is unable to answer or answers incorrectly, the interviewer will respond "You can't answer such a simple question?" or "Totally wrong, it's all learned knowledge." or "Organize your language more clearly, time is up."), and with negative animations (e.g., shaking head, yawning, pouting, rubbing shoulders, looking around impatiently, talking on the phone, or texting, see Fig. 2).

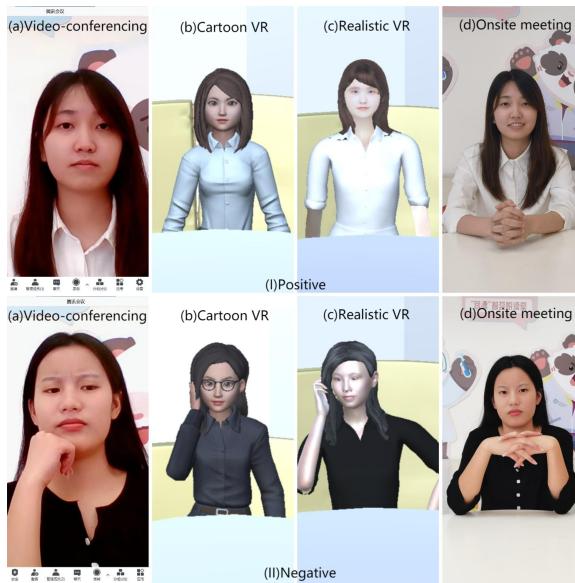


Figure 2: Four levels of realism from the least to most immersive (from left to right): video-conferencing (VR1), cartoon VR (VR1), realistic VR (VR2), onsite meeting (REAL). Two types of attitude: positive and negative (from top to bottom).

The fourth independent variable (IV) TIMED OR UNTIMED ANSWERS corresponded to whether to time each interviewee's answer. There is also past literature on the effects of timed and untimed questions on student performance and anxiety [49, 53]. In the case of a task without time limitation (i.e., no timing), each interview question could be answered for any length of time; in the case of a time-sensitive task (i.e., timing), for each interview question, the interviewer would time the interviewee for 30 seconds and interrupt the interviewer's answer as soon as the time is up.

The fifth independent variable (V) WITH OR WITHOUT PREPARATION corresponded to whether the interviewee was given 5 min-

utes to prepare for that round of the interview before the job interview began. During the 5 minutes, the interviewee could review the ten interview questions for that round and could search for information or memorize the relevant materials distributed by the staff to structure their answers in advance.

3.2 Apparatus

Fig. 1 shows our experiment setup. The whole system had four different settings for the four levels of realism. In all conditions, participants were asked to wear an E4 wristband on their left wrist, and a smartphone on the side would display the real-time physiological data acquisition without being seen by the participant. For the video-conferencing condition, interviewers conducted video conferences with the interviewee via Tencent meeting on a laptop computer with Windows 10 operating system, NVIDIA GeForce RTX 3060 GPU, and a 16.1" monitor with a resolution of 1920 × 1080. The cartoon VR and realistic VR conditions used the same experimental equipment setup. Interviewers conducted a Tencent meeting with the interviewee on a laptop computer while wearing a Meta Oculus Quest 2, a standalone headset with an internal, Android-based operating system, graphics of 1832 × 1920 pixels per eye at 90 Hz, and a 6 GB of LPDDR4X RAM processor. Through a fiber-optic link cable, we connected the socket of the VR headset to the USB socket of the laptop computer, which allowed us to cast the scene rendered in the VR headset directly to the laptop computer through the SideQuest application. Then, the picture on the laptop computer would be screen shared with the interviewer through Tencent meeting so that the interviewer could give verbal feedback in real-time according to the animations of the avatar and the interviewee's performance. The interviewee could also hear the interviewer's voice reply in the Tencent meeting on the laptop computer linked to the VR headset. For the onsite meeting condition, interviewees had a face-to-face interview with real human interviewers in the real site as Fig. 3 demonstrated.

3.3 Application

The Unity applications consisted of scenes and interviewers' avatars, and the Unity3D version we used for developing the application is 2020.3.25. In order to focus only on the realism of the avatar itself, we excluded the interference of different environments by making them the same as the physical environment. We built a virtual interview scene based on a real interview scene by using the abundant 3D models in the Unity Asset Store. Fig. 3 demonstrates real interview site and virtual interview scenario. We designed the corresponding avatars based on two real females (see Fig. 2). In order to present avatars with different realism (cartoon and realistic), we used two different modeling approaches. The cartoon avatar was developed by using *Ready Player Me*¹, a free web platform that supports users to automatically generate an avatar that resembles a real person by uploading a selfie. The realistic avatar was created by using *Avatar SDK*², which is an advanced avatar creation toolkit using AI to create photorealistic and lifelike 3D avatars from selfie photos. In order to rig the skeleton and animate the avatars, we used the Mixamo auto rigging tool to rig and animate our characters by uploading the models to the Mixamo website and selecting the animations needed by the corresponding avatars for downloads (e.g., handshaking for the positive interviewer and pouting for the negative interviewer). Avatars of interviewers were programmed to achieve various autonomous animations using the corresponding animator controller. With the animator components added to our avatars, the avatars would automatically play the pre-customized animation sequence with natural transitions. Using the Unity XR Interaction

¹<https://readyplayer.me>

²<https://avatarsdk.com>

Table 1: Orthogonal design with multi-factors and mixed levels.

Experiment group	Interviewer attitude	Question type	Timing	Preparation	Realism
1	Positive	Personal	Yes	Yes	Video-conferencing
2	Positive	Personal	No	No	Realistic VR
3	Positive	Professional	Yes	No	Cartoon VR
4	Positive	Professional	No	Yes	Onsite meeting
5	Negative	Personal	Yes	No	Onsite meeting
6	Negative	Personal	No	Yes	Cartoon VR
7	Negative	Professional	Yes	Yes	Realistic VR
8	Negative	Professional	No	No	Video-conferencing

Toolkit³, we built our application on the Android platform as “apk” files, which would run on an Oculus Quest 2 headset.



Figure 3: Real interview site (L) & virtual scenario built in Unity (R).

3.4 Participants

The participants were recruited from our university campuses with voluntary consensus. They were undergraduate students facing internship, job, and graduate school review interviews in the next year or two. They were, therefore, likely to be the primary users of the VRIS. Nineteen university students ($M = 11$, $F = 8$) participated in this experiment, aged between 19 and 21 ($M = 19.9$, $SD = 0.64$). Participants received ¥100 each for their participation. Each participant performed the interviews in all eight experiment conditions within four days (twice daily). Each participant was interviewed twice a day (i.e., once in the morning and once in the evening) with a 12-hour interval between the two sessions. Each interview is approximately 5-10 minutes long. The serial numbers of the experiment conducted each time were counterbalanced to reduce sequence effects.

3.5 Procedure

The flow chart of the experimental protocol is represented in Fig. 4. Experiments were conducted in a controlled studio setting where the room temperature was set to 25°C-27°C with indoor air conditioning. The studio was also designed to resemble a pleasant, calm lounge where users may use VRIS to mimic job interviews. We supposed that many real-life interviews would take place in a similar scenario. Before all interviews began, participants were asked to fill out the Measure of Anxiety in Selection Interviews (MASI) and General Self-Efficacy (GSE) scale in order to filter out exceptional cases of being too nervous about the interview (or even suffering from a related illness) or not nervous at all. In our interview experiment, participants were told to imagine that this was a real job interview scenario and that each interview question asked by the interviewer needed to be thought through and answered carefully. During the interview, a staff member sitting beside the interviewee would collect physiological data from the bracelet. At the end of each interview, the interviewees were asked to fill out two questionnaires (a NASA Task Load Index (NASA TLX) and an anxiety self-assessment questionnaire).

³<https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@2.2/manual/index.html>

a self-assessment anxiety questionnaire) based on their experience during the interview they just completed, and the interviewer will rate the interviewees’ performance on a score sheet for the round.

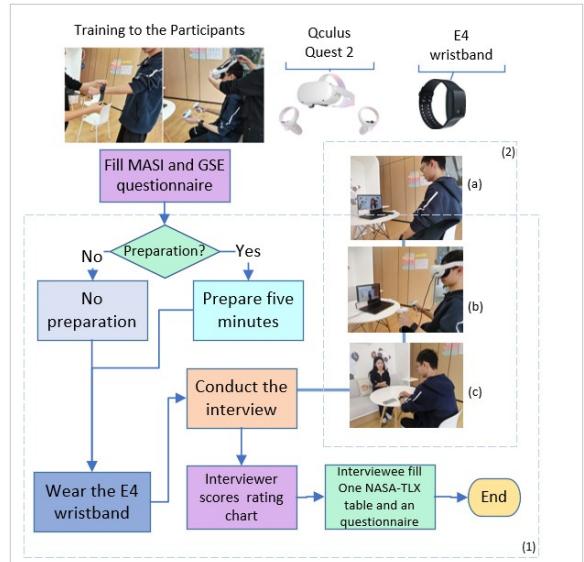


Figure 4: Flow chart of the experiment. (1) Repeat two times in one day (one in the morning and one in the evening with a 12-hour interval). Each participant performed the interviews in all 8 different experiment conditions within 4 days. (2) Choose one of the experiment setup for each interview.

3.6 Data Analysis

In order to comprehensively assess interviewer anxiety, overall experience, and interview performance, we combined subjective questionnaires and objective physiological signals, where the questionnaires were divided into interviewees’ self-perceptions and interviewers’ observations and evaluations; and we used electrodermal activity, in particular skin conductance response, as a quantitative approach to measure interviewees’ anxiety.

Questionnaires: Interviewees were asked to fill in a paper version of the questionnaires to measure self-rated anxiety both before and after the interview. Before starting the experiment, each interviewee was asked to fill in personal information, the Measure of Anxiety in Selection Interviews (MASI), and the General Self-Efficacy (GSE) scale. Then after each set of experiments, interviewees were asked to complete the NASA Task Load Index (TLX) and an anxiety self-assessment questionnaire.

The Measure of Anxiety in Selection Interviews (MASI) was used to measure interviewees’ self-rated interview anxiety. MASI is a concise and practical measurement tool that comprehensively

assesses multiple aspects of job interview anxiety [47]. The MASI includes measures of interview anxiety across five dimensions: communication anxiety, appearance anxiety, social anxiety, performance anxiety, and behavioral anxiety. Each dimension includes six questions, for a total of 30 questions. More than 35% of the answers to the MASI questions scored above three on the five-point response scale (1= strongly disagree, 5= strongly agree), indicating that interviewees displayed considerable anxiety in at least some aspects of the interview process [62]. Based on this criterion, i.e., “11 questions with a score greater than 3”, the eligible samples of MASI results were screened and aggregated, with nine interviewees experiencing substantial interview anxiety, representing 47% of the total number of interviewees.

Self-efficacy is a central concept in Bandura’s social cognitive theory, where “efficacy expectations are presumed to influence the level of performance by enhancing intensity and persistence of effort”. Further examination revealed that self-efficacy significantly influences human behavior (e.g., stress reactions, self-regulation, coping, achievement striving, career pursuits) [60]. A 10-item version of the General Self-Efficacy (GSE) scale [37] of which the total score range is 10-40 using a 5-point Likert metric with options “1 = Not at all true, 2 = Hardly true, 3 = Moderately true, 4 = Exactly true” was used to measure the interviewees’ ability to deal with various stressors in their life and, in particular, to have control over their actions in the interview setting. The Chinese version of GSE [61] has been validated by Zhang and Schwarzer to have good reliability and validity. People with different levels of self-efficacy feel, think and act differently. At the level of feelings, self-efficacy is often associated with depression, anxiety, and helplessness. Generally, a score of 20 or below on the GSE scale indicates low self-efficacy. After analyzing the data, we calculated that the average score of the participants was 25.21, and 18 participants scored between 20 and 30, i.e., they met the criteria of “high self-efficacy,” while only one participant scored 17 indicating “low self-efficacy and sometimes low confidence.”

National Aeronautics and Space Administration Task Load Index (NASA TLX) [27] is a subjective workload rating scale. It uses six dimensions to assess workload: mental demand, physical demand, temporal demand, performance, effort, and frustration. Interviewees are asked to rate each dimension on a twenty-step bipolar scale with a score from 0 to 100 (0= Very Low, 100=Very High). NASA TLX has been proven to have the highest factor validity rated the best in its ability to represent workload in the four subjective workload scales [32]. This scale measures the interview’s mental, performance, and psychological effects on the interviewer across different experimental groups.

We also designed a self-assessment questionnaire with three questions for interviewees to rate their levels of anxiety, subjective discomfort, and eye contact avoidance, as studies have found that less eye contact is strongly associated with the interviewee being more anxious, more uncomfortable and less well behaved, which equates to more nervousness [33]. The scale of both anxiety and discomfort ranged from 0 to 100, and interviewees were given four options to measure their eye contact: no eye contact, glance, gaze, or gaze with a smile. Each choice was eventually mapped to a scale of 0-100 for analysis.

Prior research has suggested that interviewers can detect interview anxiety with reasonable accuracy [18]. Therefore, for interviewers to evaluate each interview session based on the interviewees’ performance, a performance rating scale with three dimensions, each with a score range from 0 to 100, was applied to ascertain the participants’ perceptions of their anxiety and behavior during the interview. First, the interviewee’s nervousness was scored through non-verbal behaviors, such as shaking hands, stiff movements, or demonstrating little eye contact. Second, the interviewee’s performance was scored on their answers’ accuracy, logic, and time control. Third, the inter-

viewee’s communication skill is scored by the ratio of pauses, errors, stammering, slurring, and verbal chanting throughout the interview. The self-assessment questionnaire and performance rating scale can be found in the appendix.

Physiology Measures: Physiological measures related to feelings of anxiety are mainly composed of EDA(Electrodermal Activity). We used the Empatica E4 bracelet with two electrode sensors to measure these data. It captures conductance (inverse resistance) through the skin, passing a minimal amount of current between two electrodes in contact with the skin to obtain EDA signal data. When experiencing emotional activation, increased cognitive load, or physical exertion, the brain sends signals to the skin, the pores begin to fill below the surface, and conductance increases in a measurable way [25]. One component of EDA is the phasic component, which refers to the faster-changing elements of the signal - the Skin Conductance Response (SCR) [10]. Tonic data in EDA signals such as SCL(Skin conductance level) levels vary according to individual differences and changes in the experimental setting and thus need to record baseline for further analysis. Yet, our analysis mainly focused on ER-SCR(Event-related skin conductance response), which is one kind of phasic data when specific events (e.g., visual stimuli or stressful events) induce corresponding SCRs where individual differences and changes in time and environment play little role. So recording the baseline is not mandatory when analyzing ER-SCR in our experiment. The processing for extracting the ER-SCR from the EDA data is as follows:

- (1) Raw data collection: The data from the E4 wristband website was collected right after each interview. And then, raw EDA data with Unix timestamps were converted to the local time of the lab.
- (2) ER-SCR extraction: we used the neurokit2 package⁴, a python toolbox for neurophysiological signal processing. Through functions neurokit2 provided, we could feed raw EDA signal and got returns like the number of occurrences of Skin Conductance Response (SCR), the mean amplitude of the SCR peak occurrences, and other SCR information for future analysis [46].

4 RESULTS

All participants completed the orthogonal experiment, and therefore we opted for the mixed-effects model (also named “multilevel model” or “hierarchical model”) to analyze the data from repeated measures [7]. The mixed-effects model includes fixed effects associated with the response variable and random effects not related to the response variable. In order to get over the effect of individual differences on the interview experience, we set it as a random effect in this model. Other variables presented in Table 1 are fixed effects. Mixed-effects model can study whether one factor has a higher impact on the response variable compared to the other factors. In comparison, the post-hoc analysis can compare the effect of different levels in one factor. Therefore, we carry out the analysis of the mixed-effects model and post-hoc analysis. Data analysis was performed in R, and no significant interaction effects of any combinations of independent variables were detected. The significance level was set to .05. We also used corrections when performing the post-hoc analysis.

4.1 Feedback on Anxiety

Table 2 presents the results of mixed-effects model for analyzing the effect of five independent variables (*Interviewer attitude*, *Question type*, *Time keeping*, *Preparation*, *Realism*) on interviewees’ anxiety. We measured interviewees’ anxiety in three ways, which were **self-perceived anxiety**, **physiological anxiety** and **interviewer-rated anxiety**; **self-perceived anxiety** was from the self-rated anxiety scores in questionnaires for interviewees as a subjective indicator, **physiological anxiety** was the ER-SCR data extracted from electrodermal activities embodying the level of anxiety as an objective indicator, and **interviewer-rated anxiety** was

⁴<https://pypi.org/project/neurokit2/>

from the rating chart for interviewers. Results from mixed-model rANOVA indicated that *Question type* ($F_{1,133} = 89.31, p = .01, \eta_p^2 = .40$), *Preparation* ($F_{1,133} = 5.45, p = .02, \eta_p^2 = .04$) and *Realism* ($F_{1,133} = 3.14, p = .02, \eta_p^2 = .02$) all have significant influence on **self-perceived anxiety**; while *Question type* ($F_{1,133} = 4.30, p = .04, \eta_p^2 = .03$) and *Interviewer attitude* ($F_{1,133} = 4.38, p = .04, \eta_p^2 = .03$) both significantly affect *physiological anxiety*; followed with **interviewer-rated anxiety**, from the interviewer's point, interviewees' anxiety is affected by *Question type* ($F_{1,133} = 22.61, p < .01, \eta_p^2 = .15$), *Interviewer attitude* ($F_{1,133} = 6.78, p = .01, \eta_p^2 = .05$), and *Time keeping* ($F_{1,133} = 10.04, p < .01, \eta_p^2 = .07$), which is quite different to the results reported by interviewees.

Table 2: Analysis of factors that affecting interviewee's anxiety using mixed-effects model(NumDF=1, DenDF=133)

Interviewer attitude	Question type	Time keeping	Preparation	Realism
Self-perceived Anxiety	SS=55.7	SS=23032.3	SS=1.3	SS=1344.1
	F=89.31	F=.005	F=5.45	F=3.14
	p<.01**	p=.94	p=.02*	p=.02*
ER-SCR (Gromba2008)	$\eta_p^2=.40$	$\eta_p^2=.00$	$\eta_p^2=.04$	$\eta_p^2=.02$
	SS=.0074	SS=.0072	SS=.0034	SS=.0005
	F=4.30	F=2.05	F=.32	F=.28
Interviewer-rated Anxiety	p=.04*	p=.15	p=.57	p=.53
	$\eta_p^2=.03$	$\eta_p^2=.02$	$\eta_p^2=.00$	$\eta_p^2=.00$
	SS=1392.1	SS=4642.1	SS=2063.2	SS=65.8
	F=6.78	F=22.61	F=10.04	F=.32
	p<.01**	p<.01**	p=.01**	p=.07
	$\eta_p^2=.05$	$\eta_p^2=.15$	$\eta_p^2=.07$	$\eta_p^2=.02$

Further post-hoc analysis is shown in Table 3. Pairwise comparisons were performed using the t-test. Compared to *Personal question*, *Professional question* significantly leads to higher **self-perceived anxiety** ($MD = 24.07, t_{133} = 9.45, p < .01, \eta_p^2 = .40$). *Without preparation* also leads to higher **self-perceived anxiety** ($MD = -5.95, t_{133} = -2.33, \mathbf{p=.02^*}, \eta_p^2 = .04$), compared with the *Realistic VR* interview, *PC* interview are linked to lower *self-perceived anxiety* ($MD = -7.5, t_{133} = -2.08, p = .03, \eta_p^2 = .03$), similarly, *Real person* interview leads to significantly lower *self-perceived anxiety* ($MD = -10.7, t_{133} = -2.97, p < .01, \eta_p^2 = .06$) than *Realistic VR*.

When it comes to **physiological anxiety**, only *Question type* and *Interviewer attitude* have significant influence with a negative interviewer ($MD = .01, t_{133} = 2.10, p = .04, \eta_p^2 = .03$) inducing more **physiological anxiety** than the positive one; yet unlike **self-assessed anxiety**, personal questions ($MD = -0.01, t_{133} = -2.07, p = .04, \eta_p^2 = .03$) tends to induce more **physiological anxiety** than professional ones.

In terms of **interviewer-rated anxiety**, same as **self-assessed anxiety**, *Professional question* significantly leads to higher **interviewer-rated anxiety** ($MD = 11.05, t_{133} = 4.75, p < .01, \eta_p^2 = .15$), also *Real person* ($MD = -7.63, t_{133} = 2.32, p = .02, \eta_p^2 = .04$) and *PC* ($MD = -7.63, t_{133} = 2.32, p = .02, \eta_p^2 = .04$) interview both induce less **interviewer-rated anxiety** than *Realistic VR* which is in line with the results of **self-assessed anxiety**; but inconsistent with **self-assessed anxiety**, *Interviewer attitude* and *Time keeping* also play a role with a negative interviewer ($MD = -6.05, t_{133} = -2.60, p = .01, \eta_p^2 = .05$) and keeping time for 30s each question ($MD = 7.37, t_{133} = 3.17, p < .01, \eta_p^2 = .07$) inducing more **interviewer-rated anxiety** than the positive one and without

time keeping.

4.2 Overall Experience

To get a full picture of the interviewee's experience, we collected their **Cognitive load**, **Discomfort** and **Avoidance of eye contact** through NASA-TLX and subjective self-assessment questionnaires.

Table 4 demonstrates the effect of interview factors on the cognitive workload measured through the NASA-TLX in six dimension using mixed-effects model. We can observe that *Interviewer attitude* doesn't have any significant influence on the NASA-TLX criteria, while *Question type* has significantly affected **Mental demand** ($F_{1,133} = 37.79, p < .01, \eta_p^2 = .22$), **Physical demand** ($F_{1,133} = 7.72, p < .01, \eta_p^2 = .05$), **Temporal demand** ($F_{1,133} = 20.85, p < .01, \eta_p^2 = .14$), **Performance** ($F_{1,133} = 33.75, p < .01, \eta_p^2 = .14$), **Effort** ($F_{1,133} = 3.78, p = .05, \eta_p^2 = .02$), **Frustration** ($F_{1,133} = 34.87, p < .01, \eta_p^2 = .21$). Among six criteria, *Timekeeping* only has significant effect on **Physical demand** ($F_{1,133} = 4.50, p = .03, \eta_p^2 = .03$) and **Temporal demand** ($F_{1,133} = 15.11, p < .01, \eta_p^2 = .10$). *Preparation* and *Realism* have similar significant effect on **Performance** ($F_{1,133} = 6.57, p = .01, \eta_p^2 = .05$, and $F_{1,133} = 3.34, p = .02, \eta_p^2 = .02$ respectively) and **Frustration** ($F_{1,133} = 7.75, p < .01, \eta_p^2 = .06$, and $F_{1,133} = 3.83, p = .01, \eta_p^2 = .03$ respectively).

Table 4: Analysis of cognitive workload of the interviewee using mixed-effects model (NumDF=1, DenDF=133)

	Mental demand	Physical demand	Temporal demand	Performance	Effort	Frustration
Interviewer attitude	Ss=112.9	SS=206.5	Ss=30.8	SS=265.8	SS=720.8	SS=154.0
	F=3.11	F=11	F=.003	F=.87	F=.85	F=.45
	p=.54	p=.73	p=.05	p=.35	p=.09	p=.50
Question type	T=.00	n=.00	n=.00	n=.00	n=.02	n=.00
	Ss=11620.0	SS=1428.7	SS=5508.1	SS=10263.2	SS=9550.1	SS=2000.7
	F=37.79	F=7.72	F=20.85	F=33.75	F=.38	F=34.87
Timekeeping	p<.01**	p<.01**	p<.01**	p<.01**	p=.05*	p<.01**
	T=.22	n=.22	n=.14	n=.20	n=.02	n=.21
	Ss=720.8	SS=833.79	SS=399.4	SS=857.4	SS=265.8	SS=519.5
Preparation	F=2.34	F=1.50	F=1.82	F=1.05	F=.51	F=.22
	p=.13	p=.03*	p<.01**	p=.09	p=.30	p=.22
	T=.02	n=.02	n=.10	n=.02	n=.00	n=.01
Realism	Ss=628.2	SS=1119.3	SS=87.0	SS=1997.4	SS=351.06	SS=2669.5
	F=2.04	F=6.4	F=.33	F=6.57	F=.39	F=.75
	p=.15	p=.43	p=.57	p=.01*	p=.24	p<.01**
	T=.02	n=.00	n=.00	n=.05	n=.00	n=.06
	Ss=1026.8	SS=261.0	SS=1085.0	SS=3044.0	SS=673.7	SS=3952.5
	F=1.11	F=4.7	F=1.37	F=.34	F=.89	F=.383
	p=.34	p=.70	p=.26	p=.02*	p=.44	p=.01*
	T=.00	n=.00	n=.01	n=.02	n=.00	n=.03

The post-hoc analysis was also performed for the NASA-TLX criteria, given in Table 5. *Professional question* tends to increase more **Mental demand** ($MD = 17.49, t_{133} = 6.15, p < .01, \eta_p^2 = .22$), **Physical demand** ($MD = 6.13, t_{133} = 7.32, p < .01, \eta_p^2 = .29$), **Temporal demand** ($MD = 12.04, t_{133} = 4.57, p < .01, \eta_p^2 = .14$), **Effort** ($MD = 5.01, t_{133} = 1.94, p = .05, \eta_p^2 = .03$), and **Frustration** ($MD = 17.78, t_{133} = 5.91, p < .01, \eta_p^2 = .21$), but reduce **Performance** ($MD = 12.04, t_{133} = 4.57, p < .01, \eta_p^2 = .14$). With *Timekeeping*, interviewee experienced more **Physical Demand** ($MD = 4.68, t_{133} = .18, p = .03, \eta_p^2 = .14$) and more **Temporal demand** ($MD = 10.25, t_{133} = 3.89, p < .01, \eta_p^2 = .10$). Also, *Preparation* can significantly improve the **Performance** ($MD =$

Table 3: Post-hoc analysis for factors affecting interviewee's anxiety using t test (MD=mean difference, df=133).

Interviewer attitude: Negative - Positive	Question type: Professional - Personal	Timekeeping: Yes - No	Preparation: Yes - No	Realism: Cartoon VR - PC	Realism: Cartoon VR - Real Person	Realism: Cartoon VR - Realistic VR	Realism: PC - Real Person	Realism: PC - Realistic VR	Realism: Real Person - Realistic VR
Self-perceived Anxiety	MD=1.21 t=4.8 $p < .01^{**}$	MD=24.07 t=9.45 $p < .01^{**}$	MD=.018 t=.07 $p = .64$	MD=-5.95 t=-2.33 $p = .02^{*}$	MD=2.39 t=.66 $p = .51$	MD=5.61 t=1.56 $p = .12$	MD=-5.11 t=-1.42 $p = .15$	MD=3.21 t=.89 $p = .37$	MD=7.5 t=2.08 $p < .01^{**}$
	$\eta^2_{\text{eff}} = .00$	$\eta^2_{\text{eff}} = .40$	$\eta^2_{\text{eff}} = .00$	$\eta^2_{\text{eff}} = .04$	$\eta^2_{\text{eff}} = .00$	$\eta^2_{\text{eff}} = .02$	$\eta^2_{\text{eff}} = .01$	$\eta^2_{\text{eff}} = .03$	$\eta^2_{\text{eff}} = .06$
	MD=.01 t=2.10 $p = .04^{*}$	MD=.01 t=2.07 $p = .04^{*}$	MD=-.004 t=-1.43 $p = .57$	MD=-.004 t=-.81 $p = .57$	MD=.0007 t=.32 $p = .41$	MD=.0003 t=.03 $p = .75$	MD=.0003 t=.50 $p = .62$	MD=.0004 t=.78 $p = .44$	MD=.0002 t=.28 $p = .78$
ER-SCR (Gombiova 2008)	$\eta^2_{\text{eff}} = .03$	$\eta^2_{\text{eff}} = .03$	$\eta^2_{\text{eff}} = .02$	$\eta^2_{\text{eff}} = .00$	$\eta^2_{\text{eff}} = .00$	$\eta^2_{\text{eff}} = .00$	$\eta^2_{\text{eff}} = .00$	$\eta^2_{\text{eff}} = .00$	$\eta^2_{\text{eff}} = .00$
	MD=6.05 t=2.60 $p = .01^{**}$	MD=11.05 t=4.75 $p < .01^{**}$	MD=-.37 t=.31 $p < .01^{**}$	MD=1.32 t=.88 $p = .57$	MD=.289 t=.88 $p = .38$	MD=-2.89 t=-.88 $p = .38$	MD=-4.74 t=-1.44 $p = .15$	MD=5.79 t=1.76 $p = .08$	MD=.763 t=.23 $p = .02^{*}$
	$\eta^2_{\text{eff}} = .05$	$\eta^2_{\text{eff}} = .15$	$\eta^2_{\text{eff}} = .07$	$\eta^2_{\text{eff}} = .00$	$\eta^2_{\text{eff}} = .00$	$\eta^2_{\text{eff}} = .00$	$\eta^2_{\text{eff}} = .02$	$\eta^2_{\text{eff}} = .04$	$\eta^2_{\text{eff}} = .04$

$7.75, t_{133} = 2.56, p < .01, \eta_p^2 = .05$) and reduce **Frustration** ($MD = -8.38, t_{133} = -2.78, p < .01, \eta_p^2 = .05$). Compared to the interview via *Real person*, *Cartoon VR* interview has an adverse effect on the *Performance* ($MD = -11.95, t_{133} = -2.99, p < .01, \eta_p^2 = .06$), *Realistic VR* will increase the **Frustration** ($MD = 14.13, t_{133} = -3.32, p < .01, \eta_p^2 = .08$).

Table 5: Post-hoc analysis for factors affecting interviewee's cognitive workload using t test (MD=mean difference, df=133).

	Mental demand	Physical demand	Temporal demand	Performance	Effort	Frustration
Interviewer attitude: Negative - Positive	MD=-1.72 t=-.61 $\eta_p^2=.03$	MD=-0.74 t=.55 $\eta_p^2=.02$	MD=-0.14 t=.05 $\eta_p^2=.01$	MD=-2.64 t=-.94 $\eta_p^2=.00$	MD=-4.36 t=-1.69 $\eta_p^2=.02$	MD=-2.01 t=-.67 $\eta_p^2=.00$
Question type: Professional - Personal	MD=-17.49 t=-6.15 $\eta_p^2=.22$	MD=-10.13 t=-7.32 $\eta_p^2=.29$	MD=-12.04 t=-5.57 $\eta_p^2=.14$	MD=-16.43 t=-5.81 $\eta_p^2=.20$	MD=-17.78 t=-1.04 $\eta_p^2=.03$	MD=-17.78 t=-.91 $\eta_p^2=.21$
Timekeeping: Yes- No	MD=-4.35 t=1.53 $\eta_p^2=.02$	MD=-4.68 t=1.18 $\eta_p^2=.14$	MD=-4.05 t=3.89 $\eta_p^2=.02$	MD=-4.75 t=1.68 $\eta_p^2=.00$	MD=-3.64 t=1.03 $\eta_p^2=.01$	MD=-3.70 t=1.23 $\eta_p^2=.00$
Preparation: Yes - No	MD=-4.06 t=.15 $\eta_p^2=.02$	MD=-1.56 t=.42 $\eta_p^2=.02$	MD=-2.25 t=.57 $\eta_p^2=.57$	MD=-3.04 t=.25 $\eta_p^2=.05$	MD=-3.88 t=.24 $\eta_p^2=.05$	MD=-3.78 t=.28 $\eta_p^2=.05$
Realism: Cartoon VR - PC	MD=-1.45 t=.36 $\eta_p^2=.00$	MD=-0.89 t=.72 $\eta_p^2=.00$	MD=-1.61 t=.43 $\eta_p^2=.00$	MD=-3.37 t=1.84 $\eta_p^2=.02$	MD=-4.87 t=1.34 $\eta_p^2=.00$	MD=-2.05 t=.48 $\eta_p^2=.00$
Realism: Cartoon VR - Real Person	MD=-3.42 t=.85 $\eta_p^2=.00$	MD=-2.66 t=.40 $\eta_p^2=.16$	MD=-0.61 t=.87 $\eta_p^2=.02$	MD=-11.95 t=-2.91 $\eta_p^2=.00$	MD=-3.84 t=1.05 $\eta_p^2=.00$	MD=-6.66 t=.56 $\eta_p^2=.02$
Realism: Cartoon VR - Realistic VR	MD=-3.68 t=.91 $\eta_p^2=.00$	MD=-5.29 t=.74 $\eta_p^2=.00$	MD=-3.29 t=1.42 $\eta_p^2=.00$	MD=-3.37 t=.82 $\eta_p^2=.00$	MD=-5.37 t=1.47 $\eta_p^2=.02$	MD=-7.47 t=1.76 $\eta_p^2=.02$
Realism: PC - Real Person	MD=-1.97 t=.49 $\eta_p^2=.00$	MD=-3.55 t=.62 $\eta_p^2=.02$	MD=-1.06 t=.27 $\eta_p^2=.00$	MD=-4.58 t=1.14 $\eta_p^2=.00$	MD=-1.03 t=.28 $\eta_p^2=.00$	MD=-.61 t=.08 $\eta_p^2=.00$
Realism: PC - Realistic VR	MD=-5.13 t=1.27 $\eta_p^2=.01$	MD=-1.71 t=.52 $\eta_p^2=.00$	MD=-6.89 t=.52 $\eta_p^2=.00$	MD=-4.08 t=1.02 $\eta_p^2=.00$	MD=-5.50 t=.14 $\eta_p^2=.02$	MD=-9.53 t=2.24 $\eta_p^2=.04$
Realism: Real Person - Realistic VR	MD=-7.10 t=1.76 $\eta_p^2=.02$	MD=-1.84 t=.41 $\eta_p^2=.11$	MD=-5.88 t=1.58 $\eta_p^2=.21$	MD=-1.53 t=.42 $\eta_p^2=.00$	MD=-14.13 t=-3.32 $\eta_p^2=.08$	

Table 6 presents the self-perceived discomfort and the level of eye contact avoidance.

Discomfort is significantly influenced by *Question type* ($F_{1,133} = 53.62, p < .01, \eta_p^2 = .29$), *Preparation* ($F_{1,133} = 3.81, p = .05, \eta_p^2 = .03$), and *Realism* ($F_{1,133} = 6.51, p < .01, \eta_p^2 = .05$), while **Avoidance of eye contact** is greatly affected by *Interviewer attitude* ($F_{1,133} = 4.03, p = .04, \eta_p^2 = .03$), *Question type* ($F_{1,133} = 8.34, p < .01, \eta_p^2 = .06$), *Preparation* ($F_{1,133} = 6.91, p < .01, \eta_p^2 = .05$) and *Realism* ($F_{1,133} = 3.69, p = .01, \eta_p^2 = .03$).

Table 6: Analysis of interviewee's discomfort and avoidance of eye contact using mixed-effects model(NumDF=1, DenDF=133)

	Interviewer attitude	Question type	Time keeping	Preparation	Realism
Discomfort	SS=920.2 F=3.17 $\eta_p^2=.02$	SS=15562.1 F=53.62 $p<.01**$ $\eta_p^2=.29$	SS=9.5 F=.03 $p=.85$ $\eta_p^2=.00$	SS=1105.9 F=3.81 $p=.05*$ $\eta_p^2=.03$	SS=5673.5 F=6.51 $p<.01**$ $\eta_p^2=.05$
	SS=1422.5 F=4.03 $p=.04*$ $\eta_p^2=.03$	SS=2944.5 F=8.34 $p<.01**$ $\eta_p^2=.06$	SS=469 F=1.32 $p=.25$ $\eta_p^2=0.00$	SS=2440.0 F=6.91 $p<.01**$ $\eta_p^2=.05$	SS=3908.7 F=3.69 $p=.01*$ $\eta_p^2=.03$

Further post-hoc analysis of **Discomfort** and **Avoidance of eye contact** in Table 7 demonstrates that more **Discomfort** discomfort tends to arise under conditions with *Professional questions* than *Personal questions* ($MD = 20.24, t_{133} = 7.32, p < .01, \eta_p^2 = .29$),

Table 7: Post-hoc analysis for factors affecting interviewee's discomfort and eye contact using t test (MD=mean difference, df=133).

	Interviewer attitude: Negative - Positive	Question type: Professional - Personal	Timekeeping: Yes- No	Preparation: Yes - No	Realism: Cartoon VR - PC	Realism: Cartoon VR - Real Person	Realism: Cartoon VR - Realistic VR	Realism: PC - Real Person	Realism: PC - Realistic VR	Realism: PC - Real Person - Realistic VR
Discomfort	MD=4.92 t=1.78 $\eta_p^2=.02$	MD=20.24 t=7.32 $\eta_p^2=.29$	MD=-0.50 t=.18 $\eta_p^2=0.00$	MD=-5.40 t=.85 $\eta_p^2=.03$	MD=6.94 t=1.95 $\eta_p^2=.03$	MD=12.79 t=3.27 $\eta_p^2=.07$	MD=-2.89 t=-0.74 $\eta_p^2=.06$	MD=5.84 t=1.49 $\eta_p^2=.02$	MD=-9.84 t=-2.52 $\eta_p^2=.05$	MD=-15.68 t=-4.01 $\eta_p^2=.11$
	MD=6.12 t=2.01 $\eta_p^2=.03$	MD=8.80 t=2.89 $\eta_p^2=.06$	MD=-3.51 t=1.15 $\eta_p^2=.00$	MD=-8.01 t=2.63 $\eta_p^2=.05$	MD=2.55 t=.59 $\eta_p^2=.03$	MD=4.37 t=1.01 $\eta_p^2=.00$	MD=-8.84 t=-0.74 $\eta_p^2=.02$	MD=1.82 t=.46 $\eta_p^2=.02$	MD=-11.40 t=-2.64 $\eta_p^2=.05$	MD=-13.21 t=-3.06 $\eta_p^2=.07$
Avoidance of eye contact	MD=-0.48 t=-0.08 $\eta_p^2=.03$	MD=-1.80 t=-0.25 $\eta_p^2=.00$	MD=-0.50 t=-0.25 $\eta_p^2=.05$	MD=-0.50 t=-0.07 $\eta_p^2=.00$	MD=-0.50 t=-0.07 $\eta_p^2=.00$	MD=-0.50 t=-0.07 $\eta_p^2=.00$	MD=-0.50 t=-0.06 $\eta_p^2=.00$	MD=-0.50 t=-0.05 $\eta_p^2=.00$	MD=-0.50 t=-0.05 $\eta_p^2=.00$	MD=-0.50 t=-0.11 $\eta_p^2=.00$

Without preparation than *With preparation* ($MD = -5.40, t_{133} = -1.95, p = .05*, \eta_p^2 = .03$), *Cartoon VR* than *Real person* ($MD = 12.79, t_{133} = 3.27, p < .01, \eta_p^2 = .07$), *Realistic VR* than *PC* ($MD = -9.84, t_{133} = -2.52, p = .01, \eta_p^2 = .05$), and *Realistic VR* than *Real person* ($MD = -15.68, t_{133} = 4.01, p < .01, \eta_p^2 = .11$). Also, *Avoidance of eye contact* which means less eye contact with the interviewer is associated with a *Negative attitude* ($MD = 6.12, t_{133} = 2.01, p = .04, \eta_p^2 = .03$) interviewer, *Personal Question* ($MD = 8.80, t_{133} = 2.89, p < .01, \eta_p^2 = .06$) and *No preparation* ($MD = -8.01, t_{133} = -2.63, p < .01, \eta_p^2 = .05$); surprisingly, *Realistic VR* can greatly reduce eye contact than any other conditions including *Cartoon VR* ($MD = -8.84, t_{133} = -2.05, p = .04, \eta_p^2 = .03$), *PC* ($MD = -11.40, t_{133} = -2.64, p < .01, \eta_p^2 = .05$) and *<Real Person* ($MD = -13.21, t_{133} = -3.06, p < .01, \eta_p^2 = .07$).

4.3 Interview Performance

The interviewers' feedback on the interviewees' performances and ability includes *Communication skill* (e.g., the ratio of pauses, errors, stammering, slurring and verbal chanting) and *Overall performance* (e.g., accuracy, logic and time control); a better communication skill means more fluent, accurate and constructive oral presentation to interview questions while a better performance indicates more correct, logical and adequate answers within time limits. The mixed-effect rANOVA results are given in Table 8.

From the interviewers' point, interviewers' *Overall performance* can be significantly influenced by *Interviewer attitude* ($F_{1,133} = 3.98, p = .04, \eta_p^2 = .03$), *Time keeping* ($F_{1,133} = 5.83, p = .02, \eta_p^2 = .04$) and *Preparation* ($F_{1,133} = 4.91, p = .02, \eta_p^2 = .04$); and their **Communication skill** tends to be affected by the *Question type* ($F_{1,133} = 6.41, p = .01, \eta_p^2 = .05$) and *Realism* ($F_{1,133} = 3.05, p = .03, \eta_p^2 = .02$).

Table 8: Analysis of interviewers' evaluation on interviewees' performance using mixed-effects model(NumDF=1, DenDF=133)

	Interviewer attitude	Question type	Time keeping	Preparation	Realism
Overall performance	SS=862.13 F=3.98 $p=.04*$ $\eta_p^2=.03$	SS=.03 F=.001 $p=.99$ $\eta_p^2=.00$	SS=1262.13 F=5.83 $p=.02*$ $\eta_p^2=.14$	SS=1063.18 F=4.91 $p=.02*$ $\eta_p^2=.04$	SS=1203.24 F=1.85 $p=.78$ $\eta_p^2=.01$
	SS=716.5 F=3.44 $p=.07$ $\eta_p^2=.03$	SS=132.2 F=6.41 $p=.53$ $\eta_p^2=.06$	SS=111.18 F=3.45 $p=.46$ $\eta_p^2=.04$	SS=716.45 F=3.05 $p=.06$ $\eta_p^2=.03$	SS=1902.0 F=3.05 $p=.03*$ $\eta_p^2=.02$
Communication skill	SS=716.5 F=3.44 $p=.07$ $\eta_p^2=.03$	SS=132.2 F=6.41 $p=.53$ $\eta_p^2=.06$	SS=111.18 F=3.45 $p=.46$ $\eta_p^2=.04$	SS=716.45 F=3.05 $p=.06$ $\eta_p^2=.03$	SS=1902.0 F=3.05 $p=.03*$ $\eta_p^2=.02$
	SS=716.5 F=3.44 $p=.07$ $\eta_p^2=.03$	SS=132.2 F=6.41 $p=.53$ $\eta_p^2=.06$	SS=111.18 F=3.45 $p=.46$ $\eta_p^2=.04$	SS=716.45 F=3.05 $p=.06$ $\eta_p^2=.03$	SS=1902.0 F=3.05 $p=.03*$ $\eta_p^2=.02$

The post-hoc analysis was carried out to study the effect of different levels of significant factors from the interviewer's opinion, results are given in Table 9. According to the results, interviewers' **Communication skill** tends to be better with *Personal question* than *Professional question* ($MD = -5.92, t_{133} = -2.53, p < .01, \eta_p^2 = .05$), but it can be worse under condition *Real Person* than *Cartoon VR* ($MD = 6.58, t_{133} = 1.99, p < .05, \eta_p^2 = .03$) and *PC* ($MD = 9.74, t_{133} = 2.94, p < .01, \eta_p^2 = .06$). Meanwhile, interviewees' **Overall performance** is worsen by *Negative interviewer*, *Timekeeping*, *Without preparation*, *Cartoon VR* than *Positive interviewer* ($MD = -4.76, t_{133} = -2.00, p = .05, \eta_p^2 = .03$), *No preparation* ($MD = -14.13, t_{133} = -3.32, p < .01, \eta_p^2 = .08$), *Real Person* than *Realistic VR* ($MD = -13.21, t_{133} = -3.06, p < .01, \eta_p^2 = .07$), *Real Person* than *Realistic VR* ($MD = -13.21, t_{133} = -3.06, p < .01, \eta_p^2 = .07$), *Real Person* than *Real Person - Realistic VR* ($MD = -13.21, t_{133} = -3.06, p < .01, \eta_p^2 = .07$).

Table 9: Post-hoc analysis for factors affecting interviewees' performances using t test (MD=mean difference, df=133).

	Interviewer attitude: Negative - Positive	Question type: Professional - Personal	Timekeeping: Yes- No	Preparation: Yes - No	Realism: Cartoon VR - PC	Realism: Cartoon VR - Real Person	Realism: Cartoon VR- Realistic VR	Realism: PC - Real Person	Realism: PC - Realistic VR	Realism: Real Person - Realistic VR
Communication skill	MD=4.34 t=1.86 $\eta_p^2=.03$	MD=5.92 p<.01** $\eta_p^2=.05$	MD=1.71 t=.73 $\eta_p^2=.00$	MD=4.34 p=.47 $\eta_p^2=.00$	MD=3.16 t=.95 $\eta_p^2=.03$	MD=6.58 t=.99 $\eta_p^2=.00$	MD=3.11 t=.64 $\eta_p^2=.03$	MD=9.74 p<.01** $\eta_p^2=.06$	MD=5.26 t=.59 $\eta_p^2=.02$	MD=4.47 t=1.35 $\eta_p^2=.01$
	MD=4.76 t=2.00 $\eta_p^2=.03$	MD=0.03 p=.05* $\eta_p^2=.00$	MD=5.76 t=.01 $\eta_p^2=.04$	MD=5.29 p=.02* $\eta_p^2=.04$	MD=1.11 t=-.33 $\eta_p^2=.00$	MD=7.37 t=-.22 $\eta_p^2=.03$	MD=3.16 t=.94 $\eta_p^2=.00$	MD=6.26 t=.61 $\eta_p^2=.03$	MD=2.05 t=1.25 $\eta_p^2=.00$	MD=4.21 t=1.25 $\eta_p^2=.01$

timekeeping($MD = -5.76, t_{133} = -2.42, p = .02, \eta_p^2 = .04$), With *preparation*($MD = 5.29, t_{133} = 2.42, p = .02, \eta_p^2 = .04$), *Real Person*($MD = -7.37, t_{133} = -2.18, p = .03, \eta_p^2 = .03$).

5 DISCUSSION

To answer the five research questions and the corresponding hypotheses about the effect of different variables, we found that each of the factors played a specific role, with *Question Type* having the greatest impact, followed by *Interview Attitude*, *Preparation* and *Realism* all having approximately the same effect, and finally *Timekeeping* having the smallest effect. Our findings both negate and support some of the hypotheses. Our results showed that professional questions, being unprepared, timed answers, and negative interviewers can indeed cause more anxiety than their respective opposites. However, in terms of *Realism*, it was predicted that a greater level of realism would result in greater anxiety, but this turned out not to be the case, where *Realistic VR* was found to have the greatest anxiety-inducing effect.

Regarding the independent variables, we found that *Question type* has the most significant effect among other independent variables. In particular, professional questions can lead to higher anxiety on each dependent variable and dimension. For example, considering anxiety, professional questions can lead to more self-perceived, SCR-embodied, and interviewer-rated anxiety; for overall experience, professional questions can cause more discomfort, more cognitive load, and less eye contact; and for interview performance, professional questions can cause decreased communication skills. The only exception is the overall performance on which *Question Type* has no significant impact. There has been little research on the impact of question types on interview anxiety; also, the impact of many question variables remains poorly understood, for example, question variables including whether the question is open or closed [23], experience-based or situational questions [15], ‘lower-order’ or ‘higher-order’ thinking [9]. Our research focused on professional and personal questions related to job interviews. Our data suggest that professional questions can cause more cognitive load than personal questions. Susan Gee et al. [23] mentioned that a recall question requires more cognitive processing than an answer to a recognition question, which offered valuable insight into our study. The professional questions in our survey required a memory search and thus can be defined as recall questions. In contrast, personal questions with specific cues provided were very familiar with recognition questions. However, Susan Gee used a sample of 157 children aged nine to thirteen, and our experiments mainly targeted college students. Next, *Interview Attitude*, *Preparation*, and *Realism* all have considerable effects on each dependent variable. Apparently, a negative interviewer can cause more SCR-embodied anxiety, interviewer-rated anxiety, less eye contact, and worse performances. Joung Huem Kwon’s finding considered that anxiety level was affected more by the attitude of the virtual interviewer than its level of realism [42], whereas our findings do not support that the attitude’s impact necessarily outweighed the level of realism. Joung Huem Kwon’s experiment only focused on virtual humans and did not include a real human interviewer. Also, the indicator of the interviewee’s anxiety used in Joung Huem Kwon’s study was

simply physiological measurements(i.e., the percent rate of gaze fixation and eye blink). In a similar study, Patrick Gebhard [22] designed two types of virtual recruiters: a sympathetic one with friendly facial expressions and a warm tone; and a demanding one with unfriendly facial expressions and a cold tone. They found that the participants felt that the demanding character induced a higher stress level than the understanding character and felt less comfortable, which is in line with our discovery in terms of the effect on the overall experience. Similarly, no preparation before an interview can lead to more self-perceived anxiety, more discomfort, a higher cognitive load regarding frustration, and worse performance. This is consistent with a previous study which suggested that job-seekers perform better in job interviews when they are better prepared and have rehearsed answers to common interview questions, and the experiential practice of mock interviews may enhance students’ preparation for real-world job interviewing [26]. The influence of *Realism* is much more complicated since this variable has four levels, mixed-effect model indicates that *Realism* has a significant impact on self-perceived anxiety, discomfort, cognitive load, eye contact, and communication skills, with further post-hoc analysis, we discovered that *Realistic VR* induces more self-perceived anxiety, more discomfort, higher cognitive load regarding frustration, and less eye contact than *PC* and even *Real Person*, which is against our prior hypotheses that *Real Person* should have caused more anxiety than *Realistic VR*, yet, it is reasonable and in line with many previous findings that VR is effective in inducing stress [17, 57, 63]. Also, there is no significant difference shown in any dependent variables between *Realistic VR* and *Cartoon VR* except for *Realistic VR* can reduce eye contact compared to *Cartoon VR*, which aligns with Jean-Luc Lugrin’s finding that graphical details or level of realism for avatar visual display reveal no significant differences [45]. Lastly and unexpectedly, *Timekeeping* has the least impact, only shown in interviewer-rated anxiety and performance; specifically, keeping time increases interviewer-rated anxiety and cognitive load considering physical and temporal demand, also leads to worse performance. Nevertheless, the ability to finish tasks under time urgency is crucial; thus, a previous study has validated a Virtual Training System for improving time-limited decision skills and learning performances [52], while our research mainly focused on interview performance instead of learning performance as the previous study, both studies indicated the potential of virtual reality as a training tool.

Regarding the dependent variables, results indicate that **Anxiety** is greatly influenced by *Question type*, secondly *Interviewer attitude*, and lastly *Timekeeping*, *Preparation* and *Realism*; while **Overall experience** is greatly influenced by *Question type*, *Preparation* and *Realism*, secondly *Interviewer attitude* and *Timekeeping*; yet **Performance** is effected by all five variables and with almost the same level of influence with *Preparation* slightly having more impact. We further investigated the association between dependent variables and found consistent associations between self-perceived anxiety, SCR-embodied anxiety, and interviewer-rated anxiety, especially in *Question Type* and *Realism* where *Realistic VR* tends to induce more anxiety than *PC* and *Real Person*. However, we found an inconsistency between self-perceived performance collected in NASA-TLX and interviewer-rated performance; interviewees tend to believe their performances are influenced by *Question type*, *Preparation* and *Re-*

alism while interviewers think that their performances are mainly affected by *Interviewer attitude*, *Timekeeping* and *Preparation* even though both sides found *Preparation* influenced performances. The inconsistency might be because the interviewer had a full-body avatar. However, the interviewee only had both hands as a physical presence in virtual reality, where the interviewers could only judge the interviewee's voice without facial expressions or eye contact to rate their performances. Therefore, the interviewer's evaluation might need to be completed in virtual reality. For the interviewer to evaluate the interviewee's performance more comprehensively, the interviewees could also interact with more expressive avatars, such as a customized avatar with facial and motion capture that can deliver their feelings, facial expressions, and body movements in real-time. The previous study also showed that facial animation could increase the enfacement illusion and avatar self-identification [24].

Our findings have a few implications for the optimization and development of VRIS: (1) Professional questions and an interviewer with a negative attitude can remarkably induce anxiety during an interview; (2) VR interviews can indeed be effectively used to produce a similar interview experience, inducing the same or even more anxiety and discomfort than real person interviews to the interviewees; (3) low-fidelity avatars can provide the same user experience, anxiety level, cognitive load as the high ones while having lower requirements for computational performance, time latency, network load, and hardware; (4) preparation is still the critical element to have good performance; (5) during an interview, self-perceived anxiety and the interviewer's evaluated anxiety are approximately the same which means that the interviewer can detect the interviewee's tension level well.

6 LIMITATION

The quantification of anxiety is tricky, and electrodermal activity responses may not accurately capture the transient nature of anxiety and are influenced by irrelevant factors, including food and drink intake. Thus for future work, we consider using eye movements, facial expressions, voice intonation, physical motions, or nerve center activity to quantify anxiety comprehensively. In addition, the influence of long-term studies is not reflected in our short-term, sequential experiments; a long-term consecutive interview study may reveal additional implications for designing VRIS. Moreover, we can further investigate the relationships between dependent variables to identify whether the higher the anxiety level, the worse the experience or interview performance must be.

7 CONCLUSION

We developed and evaluated a virtual interview simulator to investigate the possible causes of anxiety in job interviews within VRIS. By conducting an orthogonal experimental design with eight job interview conditions and evaluating it with 19 college students to assess the significance of five possible anxiety-inducing factors, our study sheds light on understanding the fundamental factors creating people's anxiety and influencing experience and performance during an interview. Results confirm the significance of specific variables and emphasize the need to consider question types in VRIS. We also identified the effectiveness of the VR interviews regarding anxiety-inducing and overall experience compared to real-person interviews; therefore, VRIS could be a promising tool in training and practicing for interviews.

8 APPENDIX

8.1 Interview questions and questionnaires

Our supplementary materials include eight sets of interview questions (i.e., 4 sets of personal questions and 4 sets of professional

questions), a "Self-assessment Questionnaire" and a "Performance Rating Scale" see ⁵.

REFERENCES

- [1] S. Addelman. Orthogonal main-effect plans for asymmetrical factorial experiments. *Technometrics*, 4(1):21–46, 1962.
- [2] D. Adiani, A. Itzkovitz, D. Bian, H. Katz, M. Breen, S. Hunt, A. Swanson, T. J. Vogus, J. Wade, and N. Sarkar. Career interview readiness in virtual reality (cirvr): a platform for simulated interview training for autistic individuals and their employers. *ACM Transactions on Accessible Computing (TACCESS)*, 15(1):1–28, 2022.
- [3] G. Albright, N. Khalid, K. Shockley, K. Robinson, K. Hughes, B. Paced-Danley, et al. Innovative virtual role play simulations for managing substance use conversations: Pilot study results and relevance during and after covid-19. *JMIR Formative Research*, 5(4):e27164, 2021.
- [4] 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 Journal of Psychology*, 13(2):251, 2017.
- [5] R. M. Baños, C. Botella, C. Perpiñá, M. Alcañiz, J. A. Lozano, J. Osma, and M. Gallardo. Virtual reality treatment of flying phobia. *IEEE Transactions on Information Technology in Biomedicine*, 6(3):206–212, 2002.
- [6] M. Barreda-Ángeles, S. Aleix-Guillaume, and A. Pereda-Baños. Users' psychophysiological, vocal, and self-reported responses to the apparent attitude of a virtual audience in stereoscopic 360-video. *Virtual Reality*, 24(2):289–302, 2020.
- [7] D. Bates, M. Mächler, B. Bolker, and S. Walker. Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*, 2014.
- [8] C. Botella, A. García-Palacios, H. Villa, R. Baños, S. Quero, M. Alcañiz, and G. Riva. Virtual reality exposure in the treatment of panic disorder and agoraphobia: A controlled study. *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice*, 14(3):164–175, 2007.
- [9] M. E. Bradley, L. R. Thom, J. Hayes, and C. Hay. Ask and you will receive: How question type influences quantity and quality of online discussions. *British Journal of Educational Technology*, 39(5):888–900, 2008.
- [10] J. J. Braithwaite, D. G. Watson, R. Jones, and M. Rowe. A guide for analysing electrodermal activity (eda) & skin conductance responses (scrs) for psychological experiments. *Psychophysiology*, 49(1):1017–1034, 2013.
- [11] S. L. Burke, T. Bresnahan, T. Li, K. Epner, A. Rizzo, M. Partin, R. M. Ahlness, and M. Trimmer. Using virtual interactive training agents (vita) with adults with autism and other developmental disabilities. *Journal of autism and developmental disorders*, 48(3):905–912, 2018.
- [12] D. Clus, M. E. Larsen, C. Lemey, S. Berrouiguet, et al. The use of virtual reality in patients with eating disorders: systematic review. *Journal of medical Internet research*, 20(4):e7898, 2018.
- [13] C. M. Coelho, J. A. Santos, J. Silvério, and C. F. Silva. Virtual reality and acrophobia: one-year follow-up and case study. *CyberPsychology & Behavior*, 9(3):336–341, 2006.
- [14] N. Didehbani, T. Allen, M. Kandalaf, D. Krawczyk, and S. Chapman. Virtual reality social cognition training for children with high functioning autism. *Computers in human behavior*, 62:703–711, 2016.
- [15] A. P. Ellis, B. J. West, A. M. Ryan, and R. P. DeShon. The use of impression management tactics in structured interviews: A function of question type? *Journal of applied psychology*, 87(6):1200, 2002.
- [16] P. M. Emmelkamp, K. Meyerbröker, and N. Morina. Virtual reality therapy in social anxiety disorder. *Current psychiatry reports*, 22(7):1–9, 2020.
- [17] M. Fallon, M. M. Riem, L. E. Kunst, W. J. Kop, and N. Kupper. Multimodal responses to the virtual reality trier social stress test: A comparison with standard interpersonal and control conditions. *International Journal of Psychophysiology*, 161:27–34, 2021.
- [18] A. R. Feiler. *Interviewee and interviewer ratings of job interview anxiety: an investigation of the discrepancy using the Brunswik Lens Model*. PhD thesis, University of Guelph, 2010.

⁵Click to view the online supplementary materials

- [19] A. R. Feiler and D. M. Powell. Behavioral expression of job interview anxiety. *Journal of Business and Psychology*, 31(1):155–171, 2016.
- [20] L. T. Gantt. The effect of preparation on anxiety and performance in summative simulations. *Clinical Simulation in Nursing*, 9(1):e25–e33, 2013.
- [21] A. Garcia-Palacios, H. Hoffman, A. Carlin, T. A. Furness III, and C. Botella. Virtual reality in the treatment of spider phobia: a controlled study. *Behaviour research and therapy*, 40(9):983–993, 2002.
- [22] P. Gebhard, T. Baur, I. Damian, G. Mehlmann, J. Wagner, and E. André. Exploring interaction strategies for virtual characters to induce stress in simulated job interviews. 2014.
- [23] S. Gee, M. Gregory, and M.-E. Pipe. ‘what colour is your pet dinosaur?’ the impact of pre-interview training and question type on children’s answers. *Legal and Criminological Psychology*, 4(1):111–128, 1999.
- [24] M. Gonzalez-Franco, A. Steed, S. Hoogendoyk, and E. Ofek. Using facial animation to increase the enface illusion and avatar self-identification. *IEEE transactions on visualization and computer graphics*, 26(5):2023–2029, 2020.
- [25] A. Greco, G. Valenza, and E. P. Scilingo. *Advances in Electrodermal activity processing with applications for mental health*. Springer, 2016.
- [26] K. Hansen and R. S. Hansen. Employment interview preparation: Assessing the writing-to-learn approach. In *Developments in Business Simulation and Experiential Learning: Proceedings of the Annual ABSL conference*, vol. 33, 2006.
- [27] S. G. Hart. Nasa task load index (tlx). 1986.
- [28] A. Hartholt, S. Mozgai, and A. S. Rizzo. Virtual job interviewing practice for high-anxiety populations. In *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*, pp. 238–240, 2019.
- [29] C. J. Hartwell, C. D. Johnson, and R. A. Posthuma. Are we asking the right questions? predictive validity comparison of four structured interview question types. *Journal of Business Research*, 100:122–129, 2019.
- [30] D. Harwell. A face-scanning algorithm increasingly decides whether you deserve the job. In *Ethics of Data and Analytics*, pp. 206–211. Auerbach Publications, 2019.
- [31] E. C. Helminen, M. L. Morton, Q. Wang, and J. C. Felver. A meta-analysis of cortisol reactivity to the trier social stress test in virtual environments. *Psychoneuroendocrinology*, 110:104437, 2019.
- [32] S. G. Hill, H. P. Iavecchia, J. C. Byers, A. C. Bittner Jr, A. L. Zaslade, and R. E. Christ. Comparison of four subjective workload rating scales. *Human factors*, 34(4):429–439, 1992.
- [33] A. N. Howell, D. A. Zibulsky, A. Srivastav, and J. W. Weeks. Relations among social anxiety, eye contact avoidance, state anxiety, and perception of interaction performance during a live conversation. *Cognitive behaviour therapy*, 45(2):111–122, 2016.
- [34] W. Huang, J. Wang, and L. Liu. Application of virtual reality technology in the treatment of anxiety disorder. In *Proceedings of the 2021 5th International Conference on Electronic Information Technology and Computer Engineering*, pp. 845–849, 2021.
- [35] A. I. Huffcutt, C. H. Van Iddekinge, and P. L. Roth. Understanding applicant behavior in employment interviews: A theoretical model of interviewee performance. *Human Resource Management Review*, 21(4):353–367, 2011.
- [36] X. Jin, Y. Bian, W. Geng, Y. Chen, K. Chu, H. Hu, J. Liu, Y. Shi, and C. Yang. Developing an agent-based virtual interview training system for college students with high shyness level. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 998–999. IEEE, 2019.
- [37] M. Johnston, S. Wright, and J. Weinman. Measures in health psychology: a user’s portfolio. 1995.
- [38] I. L. Kampmann, P. M. Emmelkamp, and N. Morina. Meta-analysis of technology-assisted interventions for social anxiety disorder. *Journal of Anxiety Disorders*, 42:71–84, 2016.
- [39] M. R. Kandalaf, N. Didehbani, D. C. Krawczyk, T. T. Allen, and S. B. Chapman. Virtual reality social cognition training for young adults with high-functioning autism. *Journal of autism and developmental disorders*, 43(1):34–44, 2013.
- [40] B. Kerous, R. Bartoček, R. Roman, P. Sojka, O. Bečev, and F. Liarokapis. Examination of electrodermal and cardio-vascular reactivity in virtual reality through a combined stress induction protocol. *Journal of Ambient Intelligence and Humanized Computing*, 11:6033–6042, 2020.
- [41] O. D. Kothgassner, A. Goreis, L. M. Glenk, J. X. Kafka, B. Pfeffer, L. Beutl, I. Kryspin-Exner, H. Hlavacs, R. Palme, and A. Felnhofer. Habituation of salivary cortisol and cardiovascular reactivity to a repeated real-life and virtual reality trier social stress test. *Physiology & Behavior*, 242:113618, 2021.
- [42] J. H. Kwon, C. Alan, S. Czanner, G. Czanner, and J. Powell. A study of visual perception: Social anxiety and virtual realism. In *Proceedings of the 25th Spring Conference on Computer Graphics*, pp. 167–172, 2009.
- [43] J. H. Kwon, J. Powell, and A. Chalmers. How level of realism influences anxiety in virtual reality environments for a job interview. *International journal of human-computer studies*, 71(10):978–987, 2013.
- [44] S. P. Levine and R. S. Feldman. Women and men’s nonverbal behavior and self-monitoring in a job interview setting. *Applied HRM Research*, 7(1):1–14, 2002.
- [45] J.-L. Lugrin, M. Wiedemann, D. Bieberstein, and M. E. Latoschik. Influence of avatar realism on stressful situation in vr. In *2015 IEEE Virtual Reality (VR)*, pp. 227–228. IEEE, 2015.
- [46] D. Makowski, T. Pham, Z. J. Lau, J. C. Brammer, F. Lespinasse, H. Pham, C. Schözel, and S. A. Chen. Neurokit2: A python toolbox for neurophysiological signal processing. *Behavior research methods*, pp. 1–8, 2021.
- [47] J. McCarthy and R. Goffin. Measuring job interview anxiety: Beyond weak knees and sweaty palms. *Personnel Psychology*, 57(3):607–637, 2004.
- [48] K. H. Mok, W. Xiong, and H. Ye. Covid-19 crisis and challenges for graduate employment in taiwan, mainland china and east asia: a critical review of skills preparing students for uncertain futures. *Journal of Education and Work*, 34(3):247–261, 2021.
- [49] L. W. Morris and R. M. Liebert. Effects of anxiety on timed and untimed intelligence tests: another look. *Journal of Consulting and Clinical Psychology*, 33(2):240, 1969.
- [50] M. M. North, S. M. North, and J. R. Coble. Virtual reality therapy: an effective treatment for the fear of public speaking. *International Journal of Virtual Reality*, 3(3):1–6, 1998.
- [51] T. D. Parsons and A. A. Rizzo. Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. *Journal of behavior therapy and experimental psychiatry*, 39(3):250–261, 2008.
- [52] D. M. Romano and P. Brna. Presence and reflection in training: Support for learning to improve quality decision-making skills under time limitations. *CyberPsychology & Behavior*, 4(2):265–277, 2001. doi: 10.1089/109493101300117947
- [53] S. M. Schwartz, C. Evans, and A. M. Agur. Comparison of physical therapy anatomy performance and anxiety scores in timed and untimed practical tests. *Anatomical Sciences Education*, 8(6):518–524, 2015.
- [54] M. J. Smith, E. J. Ginger, K. Wright, M. A. Wright, J. L. Taylor, L. B. Humm, D. E. Olsen, M. D. Bell, and M. F. Fleming. Virtual reality job interview training in adults with autism spectrum disorder. *Journal of autism and developmental disorders*, 44(10):2450–2463, 2014.
- [55] M. J. Smith, J. D. Smith, M. F. Fleming, N. Jordan, C. H. Brown, L. Humm, D. Olsen, and M. D. Bell. Mechanism of action for obtaining job offers with virtual reality job interview training. *Psychiatric Services*, 68(7):747–750, 2017.
- [56] A. D. Souchet, S. Philippe, A. Lévéque, F. Ober, and L. Leroy. Short- and long-term learning of job interview with a serious game in virtual reality: Influence of eyestrain, stereoscopy, and apparatus. *Virtual Reality*, 26(2):583–600, 2022.
- [57] M. Wallergård, P. Jónsson, G. Johansson, and B. Karlson. A virtual reality version of the trier social stress test: A pilot study. *Presence*, 20(4):325–336, 2011.
- [58] Y. Wang, J.-R. Chardonnet, and F. Merienne. Development of a speed protector to optimize user experience in 3d virtual environments. *International Journal of Human-Computer Studies*, 147:102578, 3 2021. doi: 10.1016/j.ijhcs.2020.102578

- [59] Y. Wang, J.-R. Chardonnet, and F. Merienne. Enhanced cognitive workload evaluation in 3d immersive environments with topsis model. *International Journal of Human-Computer Studies*, 147:102572, 3 2021. doi: 10.1016/j.ijhcs.2020.102572
- [60] G. Zeng, S.-f. Fung, J. Li, N. Hussain, and P. Yu. Evaluating the psychometric properties and factor structure of the general self-efficacy scale in china. *Current Psychology*, pp. 1–11, 2020.
- [61] J. X. Zhang and R. Schwarzer. Measuring optimistic self-beliefs: a chinese adaptation of the general self-efficacy scale. *Psychologia: An International Journal of Psychology in the Orient*, 1995.
- [62] A. P. Zielinska, J. A. Mawhinney, N. Grundmann, S. Bratsos, J. S. Y. Ho, and A. Khajuria. Virtual interview, real anxiety: prospective evaluation of a focused teaching programme on confidence levels among medical students applying for academic clinical posts. *Advances in Medical Education and Practice*, 12:675, 2021.
- [63] P. Zimmer, C. C. Wu, and G. Domes. Same same but different? replicating the real surroundings in a virtual trier social stress test (tsst-vr) does not enhance presence or the psychophysiological stress response. *Physiology & behavior*, 212:112690, 2019.