



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

Xinyi Luo<sup>1,2</sup> · Yuyang Wang<sup>2</sup> · Lik-Hang Lee<sup>3</sup> · Zihan Xing<sup>4</sup> · Shan Jin<sup>2</sup> · Boya Dong<sup>2</sup> · Yuanyi Hu<sup>5</sup> · Zeming Chen<sup>2</sup> · Jing Yan<sup>2</sup> · Pan Hui<sup>2</sup>

Received: 16 May 2023 / Accepted: 12 December 2023 / Published online: 28 February 2024  
© The Author(s) 2024

## Abstract

Virtual reality interview simulator (VRIS) is an effective and valid tool that uses virtual reality technology to train people's interview skills. Typically, it offers candidates prone to being very nervous during interviews the opportunity to practice interviews in a safe and manageable virtual environment and realistic settings, providing real-time feedback from a virtual interviewer on their performance. It helps interviewees improve their skills, reduce their fears, gain confidence, and minimize the cost and time associated with traditional interview preparation. 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, we found that *Question type* plays a major role; secondly, *Realism*, *Preparation*, and *Interviewer attitude* all have middle influence; lastly, *Timing* has 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.

**Keywords** Virtual reality · User interfaces · Interview · Anxiety

---

✉ Yuyang Wang  
yuyangwang@hkust-gz.edu.cn

Xinyi Luo  
2020090903013@std.uestc.edu.cn

Lik-Hang Lee  
lik-hang.lee@polyu.edu.hk

Zihan Xing  
p930026137@mail.uic.edu.cn

Shan Jin  
sjin752@connect.hkust-gz.edu.cn

Boya Dong  
201930360251@mail.scut.edu.cn

Yuanyi Hu  
19211080272@gdmu.edu.cn

Zeming Chen  
201930130250@mail.scut.edu.cn

Jing Yan  
yanjingcn01@gmail.com

Pan Hui  
panhui@ust.hk

<sup>1</sup> University of Electronic Science and Technology of China, Jianshe Rd, Chengdu 610000, Sichuan, China

<sup>2</sup> The Hong Kong University of Science and Technology (Guangzhou), Duxue Rd, Guangzhou 511453, Guangdong, China

<sup>3</sup> The Hong Kong polytechnic university, 11 Yuk Choi Rd, Hung Hom 999077, Hong Kong, China

<sup>4</sup> BNU-HKBU United International College, 2000 Jintong Rd, Zhuhai 519087, Guangdong, China

<sup>5</sup> Guangdong Medical University, No.1 City Avenue, Dongguan 523109, Guangdong, China

## 1 Introduction

Researchers have conducted extensive investigations on systems for VR interview training and evaluation. For instance, Hartholt et al. designed a virtual job interview practice system aimed at high-anxiety populations, such as individuals with Autism Spectrum Disorder and former convicts (Hartholt et al. 2019). Similarly, Adiani et al. created a VR-based job interview training platform for autistic individuals, offering a less anxiety-inducing virtual environment for practicing interviewing skills (Adiani et al. 2022). Jin et al. developed an agent-based VR training and multidimensional evaluation system to assist introverted college students in managing interview anxiety (Jin et al. 2019). Additionally, Gebhard et al. presented a job training simulation environment with social cue recognition techniques, targeting young people who are currently not engaged in employment, education, or training (Gebhard et al. 2014).

Studies have found that visual display (Kwon et al. 2013), interview questions (Hartwell et al. 2019), interviewer's attitude (Kwon et al. 2009), timing (Schwartz et al. 2015), and preparation (Gantt 2013) could all potentially influence interviewees' behavior. 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 and discomfort (Wang et al. 2021a, b)), but we also need a solid understanding of how each factor affects the interviewee's external performance (e.g., verbal expressions, eye contact, and body movements), especially when multiple factors are assembled in one system; thus, our research serves as a pioneer for such an investigation. Therefore, a study on factors contributing to people's interview anxiety is required to create successful training programs and develop tailored therapy approaches.

As such, we developed the virtual reality interview simulator (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. Considering the above mentioned, the current study proposed five research questions (RQs) with the following hypothesis (H):

- 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.*

- 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.*
- 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.*
- 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.*
- 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, and 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 interview anxiety, psychotherapy in virtual reality, and virtual reality interview systems. This section summarizes those works.

## 2.1 Interview anxiety

In recent years, the global pandemic has exacerbated the employment landscape, resulting in reduced hiring by many companies. This situation has given rise to a range of anxieties among college students (Mok et al. 2021), including the prevalence of interview anxiety disorder (IAD). IAD is characterized by symptoms like speech disturbances, socially inappropriate behaviors, and general nervousness (Levine and Feldman 2002). Additionally, prior research has indicated that heightened interview anxiety prompts individuals to adopt more defensive self-presentation strategies. This, in turn, can hinder interviewees' performance and reduce their competitiveness (Feiler and Powell 2016). Current studies have indicated that feelings of nervousness, anxiety, frustration, and distress during interviews were evoked by a highly evaluative situation, talking to strangers, and a sense of lack of control (McCarthy and Goffin 2004) since interviews involve both cognitively demanding tasks (responding to interview questions), and socially demanding task (interacting with an interviewer), which is ultimately judged by another person (the ratings made by the interviewer) (Powell et al. 2018).

Measurement of anxiety, nervousness, stress, and frustration typically contains self-perceived questionnaires and physiological indicators. Self-perceived questionnaires, such as the Measure of Anxiety in Selection Interviews (MASI) and NASA-TLX, are commonly used for this purpose. The MASI, for instance, assesses interview anxiety across five dimensions: communication anxiety, appearance anxiety, social anxiety, performance anxiety, and behavioral anxiety (Zielinska et al. 2021; McCarthy and Goffin 2004). National Aeronautics and Space Administration Task Load Index (NASA-TLX) (Hart 1986) is a subjective workload rating scale. It assesses workload through six dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration. It is regarded as the most effective in representing workload when compared to other subjective workload scales (Hill et al. 1992). Physiological indicators of anxiety and nervousness encompass gaze fixation and eye-blink rate captured by an eye tracker (Kwon et al. 2009, 2013), Electrodermal activity, which becomes measurable during emotional activation, heightened cognitive load, or physical exertion (Greco et al. 2016), is also employed. Additional physiological markers include heart rate variability (HRV), blood pressure (BP), electroencephalogram (EEG), and cortisol sampling (Roos et al. 2021).

## 2.2 Psychotherapy in virtual reality

Since the inception of virtual reality, numerous research studies in psychotherapy have explored virtual settings, particularly investigating their potential for treating psychiatric

conditions. As awareness of social phobia grew, North et al. conducted groundbreaking research by introducing virtual reality exposure therapy (VRET) as a treatment for social phobia. They developed a virtual auditorium capable of real-time activation (North et al. 1998). In this virtual environment, the audience and audio responses adjusted in response to the experimenter's voice, prompting the experimenter to speak louder and more forcefully. The experiment, guided by feedback from questionnaires, revealed that VRET effectively alleviated anxiety symptoms associated with public speaking and presentations. While the experiment primarily focused on auditory stimuli, it underscored the importance of further investigating the relationship between VRET and human psychology.

Virtual reality exposure therapy (VRET) has become an important treatment for various anxiety disorders. By using computer-synthesized virtual scenes (Kampmann et al. 2016) to emulate social circumstances (Emmelkamp et al. 2020), we can expose patients to virtual social situations and interactions that could induce different levels of social fears. For example, one study indicated the utility of VR in inducing stressful reactions through a combination of stressors for healthy subjects (Kerous et al. 2020).

Further studies have implied that there are three preconditions for the treatment of anxiety disorders through VRET, including immersion, anxiety, and presence (Huang et al. 2021). Parsons et al. reported 21 case studies (Parsons and Rizzo 2008) confirming that VRET can effectively treat arachnophobia (Garcia-Palacios et al. 2002), flying phobia (Baños et al. 2002), phobia of public places (Botella et al. 2007), and acrophobia (Coelho et al. 2006). Also, research has shown that VRET has been used in treating substance abuse (Albright et al. 2021), high-functioning autism (Didehbani et al. 2016), and eating disorders (Clus et al. 2018).

Several researchers have highlighted the significance of social cognition interaction training for children with Autism Spectrum Disorder (ASD). For instance, Didehbani et al. developed a virtual reality-based social cognitive training program to enhance the social skills of children with ASD (Didehbani et al. 2016). Their research, which assessed aspects such as emotion detection, social attribution, attention, and executive functioning, demonstrated that the virtual reality platform effectively improved the social deficits commonly associated with ASD. In addition to virtual reality social cognitive training for children with ASD, Burke et al. introduced Virtual Interactive Training Agents (ViTA) as a tool to enhance social skills and reduce anxiety among young individuals with ASD and other developmental disorders (Burke et al. 2018). The results indicate that individuals who experienced the ViTA training have demonstrated improved self-promotion, self-advocacy, and responding to situational queries.

### 2.3 Virtual reality interview systems

Virtual reality can promote job interview skills with some unique features that other platforms such as desktops or tablets can hardly provide, including (1) realistic interview environments allowing candidates to practice, (2) immersive interaction with virtual interviewers to receive real-time feedback, and (3) high-pressure interview atmospheres helping candidates accommodate interview-related stress and anxiety.

Virtual reality training systems have demonstrated significant effects on individuals' anxiety levels, overall experiences, and performance in various scenarios. Regarding anxiety, previous studies have found that visual display (Kwon et al. 2013), interview questions (Hartwell et al. 2019), interviewer's attitude (Kwon et al. 2009), timing (Schwartz et al. 2015), and preparation (Gantt 2013) could potentially affect interviewee's anxiety. In addition to feelings of anxiety, the impact of virtual reality on overall experience (e.g., discomfort, eyestrain, and psychosocial stress) has also been reported. For example, apparatus affects the quality of experience, with HMDs experiencing higher eyestrain and visual discomfort than PCs (Souchet et al. 2022). Helminen et al. have found the V-TSST (virtual trier social stress test) is effective at inducing psychosocial stress, which can lead to poor physical and psychological health outcomes, though the effect size is less than that from the traditional TSST (Helminen et al. 2019); Kothgassner et al. have indicated that the perceived social presence did not differ over time in the VR TSST conditions as time was not a significant factor (Kothgassner et al. 2021). Furthermore, the virtual environment can influence interviewee performance. For example, prior studies have identified a positive relationship between the number of completed virtual interviews and improved interview performance (Smith et al. 2017); 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 (Barreda-Ángeles et al. 2020).

Virtual reality interview systems have showcased their effectiveness in assessing and training interview-related skills within simulated environments. For instance, Aysina et al. introduced the job interview simulation training (JIST) to enhance the psychological preparedness of pre-retirement unemployed individuals for job interviews (Aysina et al. 2017). The research revealed that repetitive practice of interviews in a stress-free virtual environment significantly increased psychological readiness for actual interviews. This suggests the potential for JIST to contribute to increased re-employment among pre-retirement job seekers in the future.

Furthermore, studies have shown that highly interactive virtual reality role-play training is more effective than traditional role-play training for improving various interpersonal skills. Smith et al. conducted a study to evaluate the feasibility and effectiveness of a role-play simulation titled "Virtual Reality Job Interview Training" (VR-JIT) in enhancing job-related interview content and performance-related interview skills among individuals with Autism Spectrum Disorder (ASD) (Smith et al. 2014). Their findings underscored the capacity of VR-JIT to enhance job interview skills in individuals with ASD.

## 3 VRIS framework

### 3.1 Orthogonal experimental design

The orthogonal experimental design is an efficient method of studying 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 (Addelman 1962; Huffcutt et al. 2011). 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'S ATTITUDE (2 levels); (IV) TIMED OR UNTIMED ANSWERS (2 levels); and (V) WITH OR WITHOUT PREPARATION (2 levels). 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 and 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

**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

found that higher visual display levels provoke more anxiety and the sense of presence (Kwon et al. 2013). Therefore, to create four different kinds of realism, we set four conditions representing a continuous spectrum of immersion, 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 Figure 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'

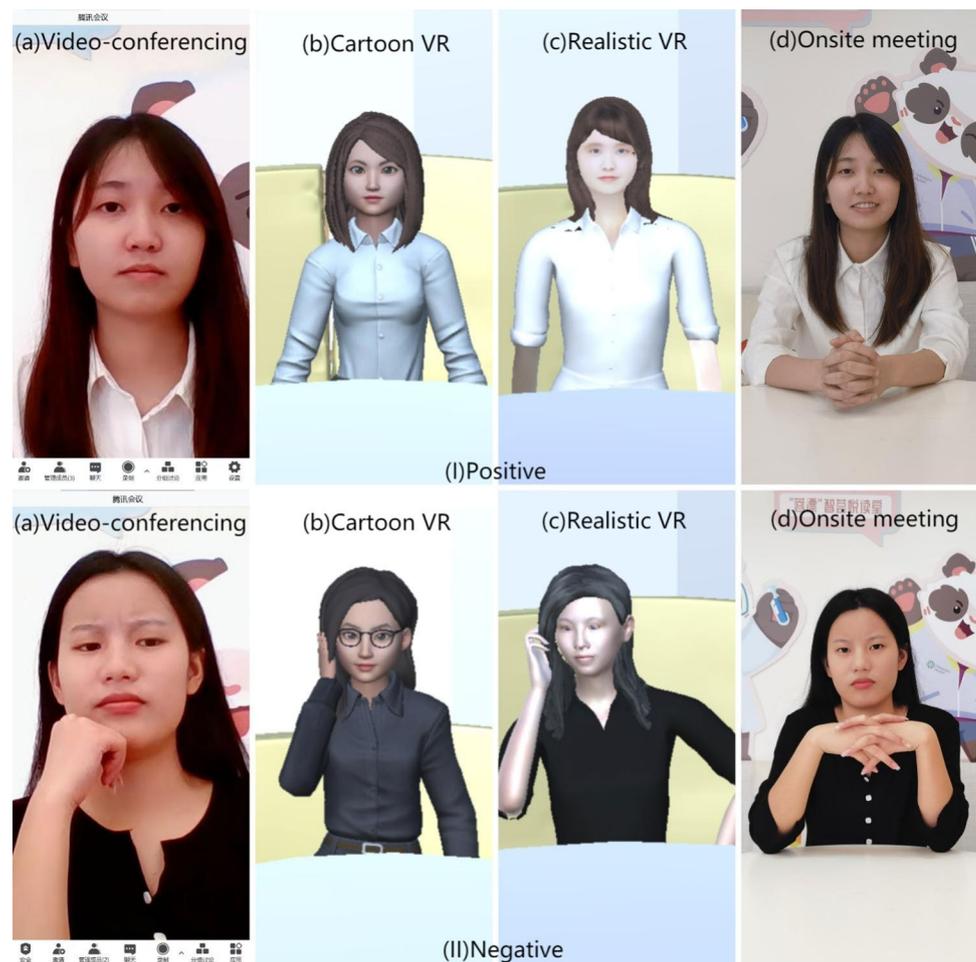
avatars had a full-body presentation, but each of the 19 interviewees only had both hands as a physical presence in the virtual environment. In the experiments under *REAL* condition, the interviewee had a face-to-face interview with a live interviewer (see Figure 1).

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 (Hartwell et al. 2019). 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

**Fig. 1** Four experiment settings for job interview simulation under the conditions: video-conferencing (*PC*), cartoon VR (*VR1*), realistic VR (*VR2*), onsite meeting (*REAL*)



**Fig. 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)



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

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, databases, 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 collected from final exams, internship interviews, job interviews, and interview questions for the graduate school review. Each of the four professional question sets had an equal number of questions of the same difficulty level. However, as personnel questions did not come with standard answers, we had to divide them into five categories: basic personal information, personality assessment, emotional control, organizing and planning skills, and creative questions; then, we formulated the four personal question sets by choosing the same number of questions from the five categories.

The third independent variable (III) INTERVIEWER'S ATTITUDE corresponded to the interviewer's attitude (body language and tone of voice) and response to the interviewee's performance during the interview process. In our experiment, instead of using appearance, we employed body language and verbal feedback to differentiate interviewer attitudes. A previous study discovered that the participants exhibited more anxiety by the attitude of virtual avatars than the avatar's level of realism (Kwon et al. 2009). 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, and acknowledging, see Figure 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."

**Fig. 3** Real interview site (L) & virtual scenario built in Unity (R)



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 Figure 2). As presented in Table 1, there were two real human interviewers, each appearing four times to represent either the negative or positive attitude. The positive interviewer appeared in experimental groups labeled 1-4, and the negative interviewer appeared in experimental groups labeled 5-8.

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 (Schwartz et al. 2015; Morris and Liebert 1969). 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 interviewee'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 minutes 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

Figure 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  $1920 \times 1080$ . 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  $1832 \times 1920$  pixels per eye at 90 Hz, and 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, allowing 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 Figure 3 demonstrates.

### 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. Figure 3 demonstrates a real interview site and virtual interview scenario.

We designed the corresponding avatars based on two real females (see Figure 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*<sup>1</sup>, 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 *AvatarSDK*<sup>2</sup>, which is an advanced avatar creation toolkit using AI to create photorealistic and lifelike 3D avatars from selfie photos.

We initially used the XR Interaction Toolkit and Animation Rigging packages to incorporate VR components and inverse kinematics into the interviewers' avatars. This allowed for the tracking of upper body movements in real time. However, this approach was eventually abandoned due to the significant training required for interviewers to operate the system and make the avatars' movements look natural. For instance, animating hand gestures or finger movements by pressing the grip and trigger buttons resulted in less fluidity compared to pre-animated sequences. In our case, animation was used to distinguish between the interviewers' attitudes (positive and negative). However, real-time tracking implies more random behaviors, making it challenging to control the magnitude and frequency of movements for each interview. Eventually, we opted for the *Mixamo* auto-rigging tool to rig and animate our characters with animations required by different attitude types (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 Toolkit<sup>3</sup>, we built our application on the Android platform as "apk" files, which would run on an Oculus Quest 2 headset.

### 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. They were sophomores and juniors from the School of Information and Software Engineering and had taken the same required computer science courses. All computer-related questions in the interview were selected from compulsory undergraduate

courses; thus, participants were supposed to have sufficient knowledge and expertise to answer these questions. 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 as a reward. 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. Participants were instructed to interact via VR controllers and to familiarize themselves with the entire VR experiment, as in Figure 4.

### 3.5 Procedure

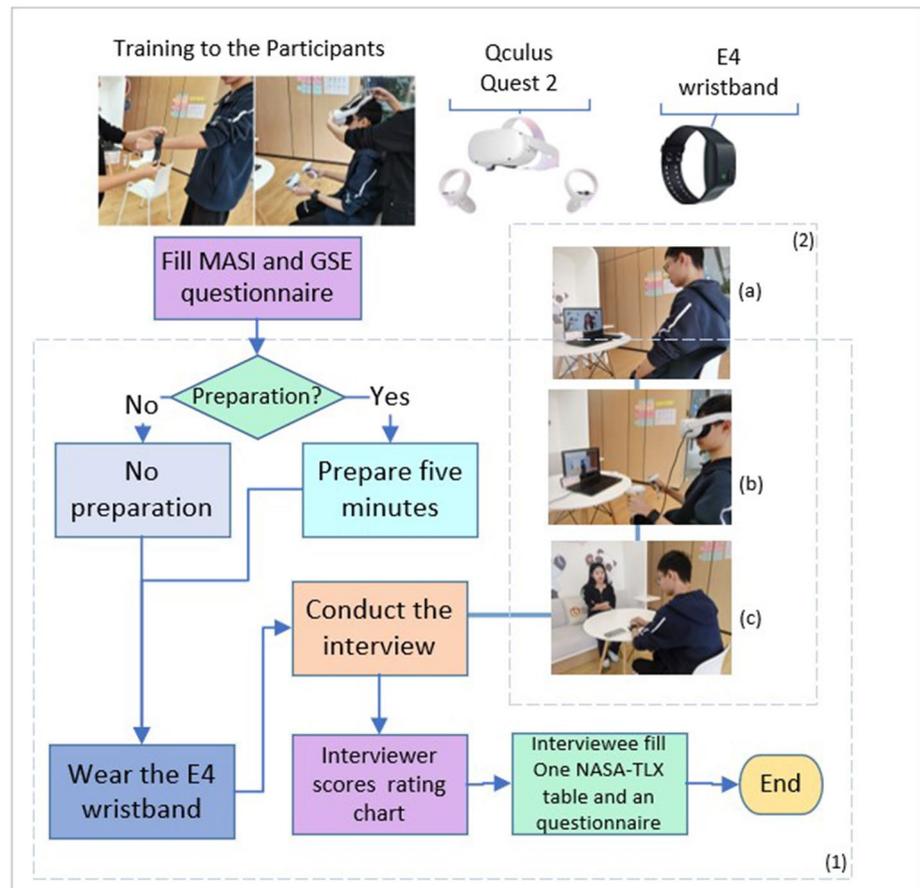
The flow chart of the experimental protocol is represented in Figure 4. The experiments were conducted in a controlled studio environment, maintaining a room temperature between 25°C and 27°C using indoor air conditioning. The studio was purposefully designed to resemble a serene and inviting lounge, creating an atmosphere in which users could utilize the Virtual Reality Interview System (VRIS) to simulate job interviews. We supposed that many real-life interviews occur in comparable settings. Before commencing the interviews, participants were required to complete the Measure of Anxiety in Selection Interviews (MASI) and the General Self-Efficacy (GSE) scale. This step was taken to identify exceptional cases of individuals who might be excessively nervous about the interview or potentially suffering from a related condition, as well as those who exhibited little to no nervousness. During our interview experiment, participants were instructed to envision the scenario as a genuine job interview. They were advised that each interview question posed by the interviewer should be considered thoughtfully and answered carefully. Throughout the interview, a staff member seated beside the interviewee was responsible for gathering physiological data using a bracelet. At the end of each interview, the interviewees were requested to complete two questionnaires, namely the NASA Task Load Index (NASA-TLX) and a self-assessment anxiety questionnaire. Additionally, the interviewer provided performance ratings for the interviewees on a score sheet for that particular round. In the two VR scenarios, the interviewer could not see the interviewees who were wearing VR headsets, so the ratings were given based on the interviewees' vocal responses and the VR casting from a 2D monitor. When interviewees were not looking into the interviewer's avatar and turned their heads randomly elsewhere (e.g., floor, ceiling, and desk) to avoid eye contact, the interviewer could observe such behavior from the VR casting.

<sup>1</sup> <https://readyplayer.me>

<sup>2</sup> <https://avatarsdk.com>

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

**Fig. 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 eight different experiment conditions within four days. (2) Choose one of the experiment setup for each interview



### 3.6 Data analysis

In order to comprehensively assess interview 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** Each of the five dimensions (communication anxiety, appearance anxiety, social anxiety, performance anxiety, and behavioral anxiety) of the MASI 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) could indicate that the participant once displayed considerable anxiety in at least some aspects of the interview. Self-efficacy significantly influences human behavior (e.g., stress reactions, self-regulation, coping, achievement striving, and career pursuits) (Zeng et al. 2020). A 10-item version of the General Self-Efficacy (GSE) scale (Johnston et al. 1995) 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 (Zhang and Schwarzer 1995) has been validated by Zhang and Schwarzer to have good reliability and validity. Generally, a score of 20 or below on the GSE scale indicates low self-efficacy. As for NASA-TLX, interviewees were asked to rate each of the six dimensions on a twenty-step bipolar scale with a score from 0 to 100 (0= Very Low, 100=Very High). 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 (ranging from 0 to 100), subjective discomfort (ranging from 0 to 100), 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 (Howell et al. 2016). Prior research has suggested that interviewers can detect interview anxiety with reasonable accuracy (Feiler 2010). Therefore, for interviewers to assess interviewees, a performance rating scale with three dimensions

was applied, each rated from 0 to 100. These dimensions include: non-verbal behavior, assessing signs of nervousness, such as limited eye contact; answers' quality, regarding correctness, logic, and time management; communication skills, considering pause duration, and pause rate. The self-assessment questionnaire and performance rating scale can be found in the appendix.

The MASI and GSE scales were applied only once before participant recruitment to screen out extreme cases. Following each interview, participants were requested to complete the NASA Task Load Index (TLX) and the anxiety self-assessment questionnaire, as depicted in Figure 4.

**Physiology measures:** The physiological measures related to feelings of anxiety in our study 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. One component of EDA is the phasic component, which refers to the faster-changing elements of the signal—the skin conductance response (SCR) (Braithwaite et al. 2013). 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. (Kritikos et al. 2019) indicated that electrodermal activity was an effective tool for anxiety detection within a virtual reality interactive scenario. 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 package<sup>4</sup>, 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 (Makowski et al. 2021).

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

## 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<sup>5</sup> from repeated measures (Bates et al. 2014). 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. A 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.

Based on the eligible samples of MASI, nine interviewees have experienced substantial interview anxiety, representing 47% of the total number of interviewees. According to the GSE questionnaire, we found that participants achieved an average score of 25.21, and eighteen participants acquired a score between 20 and 30, i.e., they met the criteria of “high self-efficacy,” while only one participant got a score of 17, indicating “low self-efficacy and sometimes low confidence.”

### 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 from the rating chart for interviewers. Results from the mixed-effects model 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*

<sup>5</sup> Available on reasonable request

**Table 2** Analysis of factors 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=22032.2	SS=1.3	SS=1344.1	SS=2322.7
	F=.23	F=89.31	F=.005	F=5.45	F=3.14
	p=.63	p<.01**	p=.94	p=.02*	p=.02*
	$\eta_p^2=.40$	$\eta_p^2=.00$	$\eta_p^2=.00$	$\eta_p^2=.04$	$\eta_p^2=.02$
ER-SCR (Gomboaa2008)	SS=.0074	SS=.0072	SS=.0034	SS=.0005	SS=.0014
	F=4.38	F=4.30	F=2.05	F=.32	F=.28
	p=.04*	p=.04*	p=.15	p=.57	p=.84
	$\eta_p^2=.03$	$\eta_p^2=.03$	$\eta_p^2=.02$	$\eta_p^2=.00$	$\eta_p^2=.00$
Interviewer-rated Anxiety	SS=1392.1	SS=4642.1	SS=2063.2	SS=65.8	SS=1276.3
	F=6.78	F=22.61	F=10.04	F=.32	F=2.07
	p=.01*	p<.01**	p<.01**	p=.57	p=.11
	$\eta_p^2=.05$	$\eta_p^2=.15$	$\eta_p^2=.07$	$\eta_p^2=.00$	$\eta_p^2=.02$

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.

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** ( $MD = -5.95, t_{133} = -2.33, p = .02, \eta_p^2 = .04$ ), 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 interviewees’ 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 dimensions using a mixed-effects model. We can observe that *Interviewer attitude* does not 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,)

**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=.48	MD=24.07 t=9.45 $p < .01^{**}$	MD=0.18 t=.07	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 $p = .03^*$	MD=-10.7 t=-2.97 $p < .01^{**}$
ER-SCR (Gomboa2008)	MD=.01 t=2.10 $p = .04^*$	MD=.01 t=2.07 $p = .04^*$	MD=-.01 t=-1.43 $p = .15$	MD=-.004 t=-.57 $p = .57$	MD=-.007 t=-.81 $p = .41$	MD=.003 t=-.32 $p = .75$	MD=-.0003 t=-.03 $p = .97$	MD=.004 t=.50 $p = .62$	MD=.007 t=.78 $p = .44$	MD=.002 t=.28 $p = .78$
Interviewer-rated Anxiety	MD=6.05 t=2.60 $p = .01^{**}$	MD=11.05 t=4.75 $p < .01^{**}$	MD=7.37 t=3.17 $p < .01^{**}$	MD=1.32 t=.57 $p = .57$	MD=2.89 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=-7.63 t=-2.32 $p = .02^*$	MD=-7.63 t=-2.32 $p = .02^*$

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).

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** ( $MD = 7.75, t_{133} = 2.56, p < .01, \eta_p^2 = .05$ ) 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 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* ( $F_{1,133} = 3.69, p = .01, \eta_p^2 = .03$ ).

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* ( $MD = 20.24, t_{133} = 7.32, p < .01, \eta_p^2 = .29$ ), *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

**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 F=.37 p=.54 $\eta^2=.00$	SS=20.63 F=.11 p=.73 $\eta^2=.00$	SS=0.8 F=.003 p=.95 $\eta^2=.00$	SS=265.8 F=.87 p=.35 $\eta^2=.00$	SS=720.8 F=2.85 p=.09 $\eta^2=.02$	SS=154.0 F=.45 p=.50 $\eta^2=.00$
Question type	SS=11620.0 F=37.79 <b>p&lt;.01**</b> $\eta^2=.22$	SS=1428.7 F=7.72 <b>p&lt;.01**</b> $\eta^2=.05$	SS=5508.1 F=20.85 <b>p&lt;.01**</b> $\eta^2=.14$	SS=10263.2 F=33.75 <b>p&lt;.01**</b> $\eta^2=.20$	SS=955.01 F=3.78 <b>p=.05*</b> $\eta^2=.02$	SS=12007.9 F=34.87 <b>p&lt;.01**</b> $\eta^2=.21$
Timekeeping	SS=720.8 F=2.34 p=.13 $\eta^2=.02$	SS=833.79 F=4.50 <b>p=.03*</b> $\eta^2=.03$	SS=3992.4 F=15.11 <b>p&lt;.01**</b> $\eta^2=.10$	SS=857.4 F=2.82 p=.09 $\eta^2=.02$	SS=265.8 F=1.05 p=.30 $\eta^2=.00$	SS=519.5 F=1.51 p=.22 $\eta^2=.01$
Preparation	SS=628.2 F=2.04 p=.15 $\eta^2=.02$	SS=119.13 F=.64 p=.43 $\eta^2=.00$	SS=87.0 F=6.57 p=.57 $\eta^2=.00$	SS=1997.4 F=6.57 <b>p=.01*</b> $\eta^2=.05$	SS=351.06 F=1.39 p=.24 $\eta^2=.00$	SS=2669.5 F=7.75 <b>p&lt;.01**</b> $\eta^2=.06$
Realism	SS=1026.8 F=1.11 p=.34 $\eta^2=.00$	SS=261.0 F=.47 p=.70 $\eta^2=.00$	SS=1085.0 F=1.37 p=.26 $\eta^2=.01$	SS=3044.0 F=3.34 <b>p=.02*</b> $\eta^2=.02$	SS=673.7 F=.89 p=.44 $\eta^2=.00$	SS=3952.5 F=3.83 <b>p=.01*</b> $\eta^2=.03$

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-effects model 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 ( $F_{1,133} = 3.05, p = .03, \eta_p^2 = .02$ ).

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

**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 p=.55 $\eta_p^2=.00$	MD=-0.74 t=1.78 p=.74 $\eta_p^2=.02$	MD=-0.14 t=-.05 p=.96 $\eta_p^2=0$	MD=-2.64 t=-.94 p=.35 $\eta_p^2=.00$	MD=-4.36 t=-1.69 p=.09 $\eta_p^2=.02$	MD=2.01 t=-.67 p=.50 $\eta_p^2=.00$
Question type: Professional - Personal	MD=17.49 t=6.15 <b>p&lt;.01**</b> $\eta_p^2=.22$	MD=6.13 t=7.32 <b>p&lt;.01**</b> $\eta_p^2=.29$	MD=12.04 t=4.57 <b>p&lt;.01**</b> $\eta_p^2=.14$	MD=-16.43 t=-5.81 <b>p&lt;.01**</b> $\eta_p^2=.20$	MD=5.01 t=1.94 <b>p=.05*</b> $\eta_p^2=.03$	MD=17.78 t=5.91 <b>p&lt;.01**</b> $\eta_p^2=.21$
Timekeeping: Yes- No	MD=4.35 t=1.53 p=.13 $\eta_p^2=.02$	MD=4.68 t=.18 <b>p=.03*</b> $\eta_p^2=.14$	MD=10.25 t=3.89 <b>p&lt;.01**</b> $\eta_p^2=.10$	MD=-4.75 t=-1.68 p=.10 $\eta_p^2=.02$	MD=-2.64 t=-1.03 p=.31 $\eta_p^2=.00$	MD=3.70 t=1.23 p=.22 $\eta_p^2=.01$
Preparation: Yes - No	MD=-4.06 t=-1.43 p=.15 $\eta_p^2=.02$	M=1.76 t=-1.95 p=.42 $\eta_p^2=.03$	MD=-1.51 t=-.57 p<.57 $\eta_p^2=0$	MD=7.25 t=2.56 <b>p=.01**</b> $\eta_p^2=.05$	MD=3.04 t=1.18 p=.24 $\eta_p^2=.01$	MD=-8.38 t=-2.78 <b>p&lt;.01**</b> $\eta_p^2=.05$
Realism: Cartoon VR - PC	MD=1.45 t=.36 p=.72 $\eta_p^2=.00$	MD=-0.89 t=1.78 p=.77 $\eta_p^2=.00$	MD=1.61 t=.43 p=.67 $\eta_p^2=0$	MD=-7.37 t=-1.84 p=.07 $\eta_p^2=.02$	MD=-4.87 t=-1.34 p=.18 $\eta_p^2=.01$	MD=2.05 t=.48 p=.63 $\eta_p^2=.00$
Realism: Cartoon VR - Real Person	MD=3.42 t=.85 p=.40 $\eta_p^2=.00$	MD=2.66 t=5.06 p=.40 $\eta_p^2=.16$	MD=0.61 t=.16 p=.87 $\eta_p^2=0$	MD=-11.95 t=-2.99 <b>p&lt;.01**</b> $\eta_p^2=.06$	MD=-3.84 t=-1.05 p=.29 $\eta_p^2=.00$	MD=6.66 t=1.56 p=.12 $\eta_p^2=.02$
Realism: Cartoon VR - Realistic VR	MD=-3.68 t=-.91 p=.36 $\eta_p^2=.00$	MD=0.82 t=-0.74 p=.79 $\eta_p^2=.00$	MD=-5.29 t=-1.42 p=.16 $\eta_p^2=.01$	MD=-3.29 t=-.82 p=.41 $\eta_p^2=.00$	MD=-5.37 t=-1.47 p=.14 $\eta_p^2=.02$	MD=-7.47 t=-1.76 p=.08 $\eta_p^2=.02$
Realism: PC - Real Person	MD=1.97 t=.49 p=.62 $\eta_p^2=.00$	MD=3.55 t=1.49 p=.25 $\eta_p^2=.02$	MD=-1.00 t=-0.27 p=.79 $\eta_p^2=.00$	MD=-4.58 t=-1.14 p=.25 $\eta_p^2=.00$	MD=1.03 t=.28 p=.78 $\eta_p^2=.02$	MD=4.61 t=1.08 p=.28 $\eta_p^2=.00$
Realism: PC - Realistic VR	MD=-5.13 t=-1.27 p=.20 $\eta_p^2=.01$	MD=1.71 t=-2.52 p=.58 $\eta_p^2=.05$	MD=-6.89 t=-1.85 p=.07 $\eta_p^2=.03$	MD=4.08 t=1.02 p=.31 $\eta_p^2=.00$	MD=-0.50 t=-.14 p=.89 $\eta_p^2=.02$	MD=-9.53 t=-2.24 <b>p=.03*</b> $\eta_p^2=.04$
Realism: Real Person - Realistic VR	MD=-7.10 t=-1.76 p=.08 $\eta_p^2=.02$	MD=-1.84 t=-4.01 p=.56 $\eta_p^2=.11$	MD=-5.89 t=-1.58 p=.12 $\eta_p^2=.21$	MD=8.66 t=2.16 <b>p=.03*</b> $\eta_p^2=.03$	MD=-1.53 t=-.42 p=.68 $\eta_p^2=.00$	MD=-14.13 t=-3.32 <b>p&lt;.01**</b> $\eta_p^2=.08$

### 5 Discussion

Regarding user experience, participants were asked post-experiment if the interviewer’s movements were live or pre-recorded sequences. Four participants believed the movements were live, while twelve were uncertain due to their focus on answering the interview questions. Interestingly,

only three participants suspected the movements were pre-recorded, noting some appeared to repeat.

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

**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	SS=15562.1	SS=9.5 F=.03 p=.85 $\eta^2=.00$	SS=1105.9 F=3.81 p=.05* $\eta^2=.03$	SS=5673.5 F=6.51 p<.01** $\eta^2=.05$
	F=3.17 p=.07 $\eta^2=.02$	F=53.62 p<.01** $\eta^2=.29$			
Avoidance of eye contact	SS=1422.5	SS=2944.5	SS=469	SS=2440.0	SS=3908.7 F=3.69 p=.01* $\eta^2=.03$
	F=4.03 p=.04* $\eta^2=.03$	F=8.34 p<.01** $\eta^2=.06$	F=1.32 p=.25 $\eta^2=0.00$	F=6.91 p<.01** $\eta^2=.05$	

approximately the same effect, and finally *Timekeeping* having the smallest effect. Our findings both negate and support some of the hypotheses. Our results implied 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 Type* 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 (Gee et al. 1999), experience-based or situational questions (Ellis et al. 2002), “lower-order” or “higher-order” thinking (Bradley et al. 2008). 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. (Gee et al. 1999) 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 (Kwon et al. 2009), 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 (Gebhard et al. 2014) 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 aligns with our discovery regarding 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'

**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 Realistic VR	Realism: PC - Realistic VR	Realism: Real Person - Realistic VR
Discomfort	MD=4.92 t=1.78	MD=20.24 t=7.32	MD=-0.50 t=.18	MD=-5.40 t=-1.95	MD=6.94 t=1.78	MD=12.79 t=3.27	MD=-2.89 t=-0.74	MD=5.84 t=-2.52	MD=-15.68 t=-4.01
	p=.07	p<.01**	p=.85	p=.05*	p=.078	p<.01**	p=.46	p=.01*	p<.01**
	$\eta_p^2=.02$	$\eta_p^2=.29$	$\eta_p^2=.00$	$\eta_p^2=.03$	$\eta_p^2=.02$	$\eta_p^2=.07$	$\eta_p^2=.06$	$\eta_p^2=.05$	$\eta_p^2=.11$
Avoidance of eye contact	MD=6.12 t=2.01	MD=8.80 t=2.89	MD=3.51 t=1.15	MD=-8.01 t=-2.63	MD=2.55 t=.59	MD=4.37 t=1.01	MD=-8.84 t=-2.05	MD=-11.40 t=-2.64	MD=-13.21 t=-3.06
	p=.04*	p<.01**	p=.25	p<.01**	p=.55	p=.31	p=.04*	p<.01**	p<.01**
	$\eta_p^2=.03$	$\eta_p^2=.06$	$\eta_p^2=.00$	$\eta_p^2=.05$	$\eta_p^2=.00$	$\eta_p^2=.00$	$\eta_p^2=.03$	$\eta_p^2=.05$	$\eta_p^2=.07$

**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* $\eta^2=.03$	SS=.03 F=.0001 p=.99 $\eta^2=.00$	SS=1262.13 F=5.83 p=.02* $\eta^2=.04$	SS=1063.18 F=4.91 p=.02* $\eta^2=.04$	SS=1203.24 F=1.85 p=.14 $\eta^2=.01$
Communication skill	SS=716.5 F=3.44 p=.07 $\eta^2=.03$	SS=1332.2 F=6.41 p=.01* $\eta^2=.00$	SS=111.18 F=.53 p=.46 $\eta^2=.00$	SS=716.45 F=3.45 p=.06 $\eta^2=.03$	SS=1902.0 F=3.05 p=.03* $\eta^2=.02$

**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 - Realistic VR	Realism: Real Person - Realistic VR
Communication skill	MD=4.34 t=1.86 p=.07 $\eta_p^2=.03$	MD=-5.92 t=-2.53 p<.01** $\eta_p^2=.05$	MD=1.71 t=.73 p=.47 $\eta_p^2=.00$	MD=-4.34 t=-1.86 p=.07 $\eta_p^2=.03$	MD=-3.16 t=.95 p=.34 $\eta_p^2=.00$	MD=6.58 t=1.99 p<.05* $\eta_p^2=.03$	MD=-2.11 t=-.64 p=.53 $\eta_p^2=.00$	MD=9.74 t=2.94 p<.01** $\eta_p^2=.06$	MD=-4.47 t=-1.35 p=.18 $\eta_p^2=.01$
Overall performance	MD=-4.76 t=-2.00 p=.05* $\eta_p^2=.03$	MD=-0.03 t=-.01 p=.99 $\eta_p^2=.00$	MD=-5.76 t=-2.42 p=.02* $\eta_p^2=.04$	MD=5.29 t=2.22 p=.03* $\eta_p^2=.04$	MD=-1.11 t=-.33 p=.74 $\eta_p^2=.00$	MD=-7.37 t=-2.18 p=.03* $\eta_p^2=.03$	MD=-3.16 t=-.94 p=.35 $\eta_p^2=.00$	MD=-6.26 t=-1.86 p=.07 $\eta_p^2=.03$	MD=4.21 t=1.25 p=.21 $\eta_p^2=.01$

preparation for real-world job interviewing (Hansen and Hansen 2006). 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 (Wallergård et al. 2011; Zimmer et al. 2019; Fallon et al. 2021). 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 Lugin's finding that graphical details or level of realism for avatar visual display reveal no significant differences (Lugin et al. 2015). 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 (Romano and Brna 2001), 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 , *Realism* 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 (Gonzalez-Franco et al. 2020).

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 a 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

Quantifying anxiety presents challenges. Electrodermal activity responses may not provide an accurate representation of anxiety and can be influenced by extraneous factors like food and drink intake. Therefore, for future research, we recommend considering alternative measures for a comprehensive quantification of anxiety. These alternatives may include eye movements, facial expressions, voice intonation, physical gestures, or neural activity. Furthermore, our short-term, sequential experiments do not account for the potential influence of long-term studies. Conducting extended, consecutive interview studies may reveal additional insights relevant to the design of Virtual Reality Interview Systems (VRIS). Additionally, investigating the relationships between dependent variables can help elucidate whether higher anxiety levels correlate with worse interview experiences or performance.

## 7 Conclusion

We developed and evaluated a virtual interview simulator to investigate the possible causes of anxiety in job interviews within VRIS. Employing an orthogonal experimental design comprising eight job interview conditions and evaluating it with 19 college students, our study aimed to discern the significance of five potential anxiety-inducing factors. The research provides valuable insights into the core factors that contribute to interview-related anxiety and influence the overall interview experience and performance. Results affirm the significance of specific variables and emphasize the necessity of considering the "Question Type" within VRIS. Additionally, we identified the effectiveness of VR interviews in comparison to traditional in-person interviews in terms of anxiety induction. This suggests that VRIS holds promise as a valuable tool for interview training and practice.

## Appendix A 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<sup>6</sup>.

**Acknowledgements** The authors appreciate the financial support for research from the *Human Resources and Social Security Department of Guangzhou City and Nansha District, Guangdong, China*.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest. The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

<sup>6</sup> Click to view the online supplementary materials

## References

- Hartholt A, Mozgai S, Rizzo AS (2019) Virtual job interviewing practice for high-anxiety populations. In: Proceedings of the 19th ACM international conference on intelligent virtual agents, pp. 238–240
- Adiani D, Itzkovitz A, Bian D, Katz H, Breen M, Hunt S, Swanson A, Vogus TJ, Wade J, Sarkar N (2022) Career interview readiness in virtual reality (cirvr): a platform for simulated interview training for autistic individuals and their employers. *ACM Trans Accessible Comput (TACCESS)* 15(1):1–28
- Jin X, Bian Y, Geng W, Chen Y, Chu K, Hu H, Liu J, Shi Y, Yang C (2019) 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
- Gebhard P, Baur T, Damian I, Mehlmann G, Wagner J, André E (2014) Exploring interaction strategies for virtual characters to induce stress in simulated job interviews
- Kwon JH, Powell J, Chalmers A (2013) How level of realism influences anxiety in virtual reality environments for a job interview. *Int J Hum Comput Stud* 71(10):978–987
- Hartwell CJ, Johnson CD, Posthuma RA (2019) Are we asking the right questions? predictive validity comparison of four structured interview question types. *J Bus Res* 100:122–129
- Kwon JH, Alan C, Czanner S, Czanner G, Powell J (2009) A study of visual perception: Social anxiety and virtual realism. In: Proceedings of the 25th spring conference on computer graphics, pp. 167–172
- Schwartz SM, Evans C, Agur AM (2015) Comparison of physical therapy anatomy performance and anxiety scores in timed and untimed practical tests. *Anat Sci Educ* 8(6):518–524
- Gantt LT (2013) The effect of preparation on anxiety and performance in summative simulations. *Clin Simul Nurs* 9(1):25–33
- Wang Y, Chardonnet J-R, Merienne F (2021) Enhanced cognitive workload evaluation in 3D immersive environments with topsis model. *Int J Human-Comput Stud* 147:102572
- Wang Y, Chardonnet J-R, Merienne F (2021) Development of a speed protector to optimize user experience in 3D virtual environments. *Int J Human-Comput Stud* 147:102578
- Mok KH, Xiong W, Ye H (2021) 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. *J Educ Work* 34(3):247–261
- Levine SP, Feldman RS (2002) Women and men's nonverbal behavior and self-monitoring in a job interview setting. *Appl HRM Res* 7(1):1–14
- Feiler AR, Powell DM (2016) Behavioral expression of job interview anxiety. *J Bus Psychol* 31(1):155–171
- McCarthy J, Goffin R (2004) Measuring job interview anxiety: beyond weak knees and sweaty palms. *Pers Psychol* 57(3):607–637
- Powell DM, Stanley DJ, Brown KN (2018) Meta-analysis of the relation between interview anxiety and interview performance. *Can J Behav Sci* 50(4):195
- Zielinska AP, Mawhinney JA, Grundmann N, Bratsos S, Ho JSY, Khajuria A (2021) Virtual interview, real anxiety: prospective evaluation of a focused teaching programme on confidence levels among medical students applying for academic clinical posts. *Adv Med Educ Pract* 12:675
- Hart SG (1986) Nasa task load index (tlx)
- Hill SG, Iavecchia HP, Byers JC, Bittner AC Jr, Zaklade AL, Christ RE (1992) Comparison of four subjective workload rating scales. *Hum Factors* 34(4):429–439
- Greco A, Valenza G, Scilingo EP (2016) Advances in electrodermal activity processing with applications for mental health. Springer, Cham
- Roos A-L, Goetz T, Voracek M, Krannich M, Bieg M, Jarrell A, Pekrun R (2021) Test anxiety and physiological arousal: a systematic review and meta-analysis. *Educ Psychol Rev* 33:579–618
- North MM, North SM, Coble JR (1998) Virtual reality therapy: an effective treatment for the fear of public speaking. *Int J Virtual Reality* 3(3):1–6
- Kampmann IL, Emmelkamp PM, Morina N (2016) Meta-analysis of technology-assisted interventions for social anxiety disorder. *J Anxiety Disord* 42:71–84
- Emmelkamp PM, Meyerbröker K, Morina N (2020) Virtual reality therapy in social anxiety disorder. *Curr Psychiatry Rep* 22(7):1–9
- Kerous B, Barteček R, Roman R, Sojka P, Bečev O, Liarokapis F (2020) Examination of electrodermal and cardio-vascular reactivity in virtual reality through a combined stress induction protocol. *J Ambient Intell Humaniz Comput* 11:6033–6042
- Huang W, Wang J, Liu L (2021) 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
- Parsons TD, Rizzo AA (2008) Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: a meta-analysis. *J Behav Ther Exp Psychiatry* 39(3):250–261
- García-Palacios A, Hoffman H, Carlin A, Furness TA III, Botella C (2002) Virtual reality in the treatment of spider phobia: a controlled study. *Behav Res Ther* 40(9):983–993
- Baños RM, Botella C, Perpiñá C, Alcañiz M, Lozano JA, Osma J, Gallardo M (2002) Virtual reality treatment of flying phobia. *IEEE Trans Inf Technol Biomed* 6(3):206–212
- Botella C, García-Palacios A, Villa H, Baños R, Quero S, Alcañiz M, Riva G (2007) Virtual reality exposure in the treatment of panic disorder and agoraphobia: a controlled study. *Clinic Psychol Psychother Int J Theor Pract* 14(3):164–175
- Coelho CM, Santos JA, Silvério J, Silva CF (2006) Virtual reality and acrophobia: one-year follow-up and case study. *CyberPsychol Behav* 9(3):336–341
- Albright G, Khalid N, Shockley K, Robinson K, Hughes K, Pace-Danley B et al (2021) Innovative virtual role play simulations for managing substance use conversations: Pilot study results and relevance during and after covid-19. *JMIR Formative Res* 5(4):27164
- Didehbani N, Allen T, Kandalaf M, Krawczyk D, Chapman S (2016) Virtual reality social cognition training for children with high functioning autism. *Comput Hum Behav* 62:703–711
- Clus D, Larsen ME, Lemey C, Berrouguet S et al (2018) The use of virtual reality in patients with eating disorders: systematic review. *J Med Internet Res* 20(4):7898
- Burke SL, Bresnahan T, Li T, Epnere K, Rizzo A, Partin M, Ahlness RM, Trimmer M (2018) Using virtual interactive training agents (vita) with adults with autism and other developmental disabilities. *J Autism Dev Disord* 48(3):905–912
- Souchet AD, Philippe S, Lévêque A, Ober F, Leroy L (2022) Short-and long-term learning of job interview with a serious game in virtual reality: influence of eyestrain, stereoscopy, and apparatus. *Virtual Real* 26(2):583–600
- Helminen EC, Morton ML, Wang Q, Felver JC (2019) A meta-analysis of cortisol reactivity to the trier social stress test in virtual environments. *Psychoneuroendocrinology* 110:104437
- Kothgassner OD, Goreis A, Glenk LM, Kafka JX, Pfeffer B, Beutl L, Kryspin-Exner I, Hlavacs H, Palme R, Felnhofner A (2021) Habituation of salivary cortisol and cardiovascular reactivity to a repeated real-life and virtual reality trier social stress test. *Physiol Behav* 242:113618

- Smith MJ, Smith JD, Fleming MF, Jordan N, Brown CH, Humm L, Olsen D, Bell MD (2017) Mechanism of action for obtaining job offers with virtual reality job interview training. *Psychiatr Serv* 68(7):747–750
- Barreda-Ángeles M, Aleix-Guillaume S, Pereda-Baños A (2020) 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
- Aysina RM, Efremova GI, Maksimenko ZA, Nikiforov MV (2017) Using a computer simulation to improve psychological readiness for job interviewing in unemployed individuals of pre-retirement age. *Eur J Psychol* 13(2):251
- Smith MJ, Ginger EJ, Wright K, Wright MA, Taylor JL, Humm LB, Olsen DE, Bell MD, Fleming MF (2014) Virtual reality job interview training in adults with autism spectrum disorder. *J Autism Dev Disord* 44(10):2450–2463
- Adelman S (1962) Orthogonal main-effect plans for asymmetrical factorial experiments. *Technometrics* 4(1):21–46
- Huffcutt AI, Van Iddekinge CH, Roth PL (2011) Understanding applicant behavior in employment interviews: A theoretical model of interviewee performance. *Hum Resour Manag Rev* 21(4):353–367
- Morris LW, Liebert RM (1969) Effects of anxiety on timed and untimed intelligence tests: another look. *J Consult Clin Psychol* 33(2):240
- Zeng G, Fung S-f, Li J, Hussain N, Yu P (2020) Evaluating the psychometric properties and factor structure of the general self-efficacy scale in china. *Curr Psychol* 1–11
- Johnston M, Wright S, Weinman J (1995) Measures in health psychology: a user's portfolio
- Zhang JX, Schwarzer R (1995) Measuring optimistic self-beliefs: a Chinese adaptation of the general self-efficacy scale. *Psychol Int J Psychol Orient*
- Howell AN, Zibulsky DA, Srivastav A, Weeks JW (2016) Relations among social anxiety, eye contact avoidance, state anxiety, and perception of interaction performance during a live conversation. *Cogn Behav Ther* 45(2):111–122
- Feiler AR (2010) Interviewee and interviewer ratings of job interview anxiety: an investigation of the discrepancy using the Brunswik lens model. PhD thesis, University of Guelph
- Braithwaite JJ, Watson DG, Jones R, Rowe M (2013) A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRS) for psychological experiments. *Psychophysiology* 49(1):1017–1034
- Kritikos J, Tzannetos G, Zoitaki C, Pouloupoulou S, Koutsouris D (2019) Anxiety detection from electrodermal activity sensor with movement & interaction during virtual reality simulation. In: 2019 9th International IEEE/EMBS conference on neural engineering (NER), pp. 571–576 . IEEE
- Makowski D, Pham T, Lau ZJ, Brammer JC, Lespinasse F, Pham H, Schölzel C, Chen SA (2021) Neurokit2: a python toolbox for neurophysiological signal processing. *Behav Res Methods* 1–8
- Bates D, Mächler M, Bolker B, Walker S (2014) Fitting linear mixed-effects models using lme4. arXiv preprint [arXiv:1406.5823](https://arxiv.org/abs/1406.5823)
- Gee S, Gregory M, Pipe M-E (1999) what colour is your pet dinosaur? the impact of pre-interview training and question type on children's answers. *Legal Criminol Psychol* 4(1):111–128
- Ellis AP, West BJ, Ryan AM, DeShon RP (2002) The use of impression management tactics in structured interviews: a function of question type? *J Appl Psychol* 87(6):1200
- Bradley ME, Thom LR, Hayes J, Hay C (2008) Ask and you will receive: how question type influences quantity and quality of online discussions. *Br J Edu Technol* 39(5):888–900
- Hansen K, Hansen RS (2006) Employment interview preparation: assessing the writing-to-learn approach. In: *Developments in business simulation and experiential learning: proceedings of the annual ABSEL conference*, vol. 33
- Wallergård M, Jönsson P, Johansson G, Karlson B (2011) A virtual reality version of the trier social stress test: a pilot study. *Presence* 20(4):325–336
- Zimmer P, Wu CC, Domes G (2019) 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. *Physiol Behav* 212:112690
- Fallon M, Riem MM, Kunst LE, Kop WJ, Kupper N (2021) Multimodal responses to the virtual reality trier social stress test: A comparison with standard interpersonal and control conditions. *Int J Psychophysiol* 161:27–34
- Lugrin J-L, Wiedemann M, Bieberstein D, Latoschik ME (2015) Influence of avatar realism on stressful situation in vr. In: 2015 IEEE virtual reality (VR), pp. 227–228 . IEEE
- Romano DM, Brna P (2001) Presence and reflection in training: support for learning to improve quality decision-making skills under time limitations. *CyberPsychol Behav* 4(2):265–277
- Gonzalez-Franco M, Steed A, Hoogendyk S, Ofek E (2020) Using facial animation to increase the enfacement illusion and avatar self-identification. *IEEE Trans Vis Comput Graph* 26(5):2023–2029

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.