Virtual reality body swapping to improve self-assessment in job interview training

Virtual Reality Body Swapping to Improve Self-Assessment in Job Interview Training

Sofia Seinfeld, F. Gabriele Pratticò, Chiara De Giorgi, Fabrizio Lamberti

***Abstract*—Swapping visual perspective in Virtual Reality provides a unique means for embodying different virtual bodies and for self-distancing. Moreover, this technology is a powerful tool for experiential learning and for simulating realistic scenarios, with broad potential in the training of soft skills. However, there is scarce knowledge on how perspective swapping in Virtual Reality might benefit the training of soft skills such as those required in a job interview. The present study investigates the impact of virtual body swapping on the self-assessment of verbal and non-verbal communication skills, emotional states, and embodiment in a simulated job interview context. Three main conditions were compared: a baseline condition in which the participants practiced a job interview from the first-person perspective of a virtual interviewee (No Swap condition); an external point of view condition where, first, the participants answered questions from the interviewee perspective, but then swap visual perspective to re-experience their responses from a non-embodied point of view (Out of Body condition); a condition in which, after answering questions from the interviewee perspective, the participants re-experienced their responses from the embodied perspective of the virtual recruiter (Recruiter condition). The experimental results indicated that the effectiveness of the Out of Body and Recruiter Conditions was superior to the No Swap Condition to self-assess the communication styles used during a job interview. Moreover, all the conditions led to a high level of embodiment towards the interviewee avatar when seen from the first-person perspective; in the case of the Recruiter Condition, the participants also felt embodied in the recruiter avatar. No differences in emotional states were found among conditions, with all sharing a positive valence.**

***Index Terms*—soft skills training, job interview simulation, self-assessment, virtual body swapping, self-distancing**

I. Introduction

Since the early days of Virtual Reality (VR), education and training were identified as areas that could benefit more from such an immersive medium. Over the years, the continuous advancements in technology and the increased availability and accessibility of cost-effective devices has contributed to catalyzing the interest of the community and boosting the diffusion of VR-based learning tools. Today, the most explored application scenarios are those in which the learner is expected to develop practical skills that may be impractical or dangerous to teach in the real world [1], [2], or in which the use of immersive realistic simulations and interactive content might benefit the transfer of theoretical and practical concepts [3].

S. Seinfelf is with Universitat Oberta de Catalunya, Barcelona, Spain (e-mail: sseinfeld@uoc.edu). Work was developed while she was with Universitat de Catalunya, Barcelona, Spain.

F. G. Pratticò, C. De Giorgi and F. Lamberti are with Politecnico di Torino, Turin, Italy (e-mail: filippogabriele.prattico@polito.it, s279109@studenti.polito.it, fabrizio.lamberti@polito.it).

The possibility offered by VR to simulate scenarios with high ecological validity, as well as to let individuals practice in safe, repeatable, cost-effective, and flexible virtual environments (VEs), is making the technology gain potential in the context of soft skills development [4]. The body of literature on this latter field, however, is way less extensive [5], [6] compared to the former one, and the specific need for further research is also highlighted by a recent comprehensive meta-analysis by Howard and Gutworth [6]. Only a few studies have been carried out regarding, for instance, public speaking [7], [8], interpersonal communication abilities [9], and interviewing [10], [11]. The speculation behind these works is that the intrinsic qualities of VR in terms of fostering immersion and presence should enable individuals to experience and exhibit behaviors comparable to those they would have in real-world situations [12], [13]. However, preliminary evidence indicates that even a naive application of VR technology, i.e. providing a virtual and interactive practice space with low fidelity, can lead to more effective learning than traditional methods based on self-video recording or mirror rehearsing [6].

The use of VR for practicing job interview skills is no exception. Villani et al. carried out a study in which it was found that trainees may experience comparable or even higher levels of anxiety in a VR simulated job interview compared to a similar real-world situation [10]. Other studies have devoted efforts to introduce conversational AI technology in VR in order to control an interviewer avatar [14] or to train the interviewer role [15]. Despite this early research, the latent potential of VR for job interview training remains largely unexplored. For instance, focusing on the particular use of VR to create a practice space for interviewee training, so far trainees have been provided with sandbox VEs where they can possibly receive feedback to improve their skills and review their performance from a first-person view [6]. There is a lack, however, of tools capable of mimicking in VR-based job interview simulations the characteristics of revision approaches leveraging video recording techniques. For instance Tailab and Marsh [16] have observed an increased awareness in the development of oral skills in those trainees who were allowed to review their performance using a video recording taken from a third-person perspective (3PP) [16].

In fact, it has been speculated also by Andrade [17] that the possibility for trainees to self-review, from a viewpoint different from the first-person one (not necessarily a 3PP [16]), would allow them to improve their self-assessment abilities. Indeed, perspective swapping is not a novel approach in VR experiences. As a matter of example, the possible advantages of having either a first-person perspective (1PP) or a 3PP, as well as of re-experiencing a virtual experience from a different embodied perspective have been studied in the context of training for an oral presentation [8], cinematic VR [18], gaming [19], sports and dance training [20], [21], and even psychotherapy [22]. Nevertheless, to the best of the authors' knowledge, there is no study that has investigated the possible benefits of perspective and virtual body swapping applied to job interview training in immersive VR.

To cope with this gap, the present work aims to explore the potential of perspective swapping and virtual embodiment for improving job interview training in VR, analyzing how these factors might impact the self-assessment of verbal and non-verbal communication skills required in professional contexts. More specifically, in a VR job interview simulation created with the intent of allowing participants to realistically practice the responses they would provide to different job interview questions, three conditions were compared: a baseline condition, in which the trainees practiced the interview only from the 1PP of an interviewee avatar (*No Swap* condition); a condition, in which the trainees first answered job interview questions from the interviewee perspective, but subsequently had the possibility to swap to a non-embodied external point of view (POV) to re-experience their responses (*Out of Body* condition); finally, a condition in which the trainees answered the questions from the interviewee perspective and then were able to re-experience their responses from the 1PP of the recruiter avatar who asked them these questions (*Recruiter* condition).

The rest of the paper describes the design, methods, and results of the study devised to reach the said goal. More specifically, Section II reviews relevant literature about perspective swapping and job interview training in VR. Section III discusses the VR simulation created to support the study, as well as the design of the experiment in terms of procedure and measures. Section IV presents the results, which are later discussed in Section V together with limitations and possible directions for future research.

II. BACKGROUND

This section provides greater details regarding the concepts related to the topics of perspective swapping and VR job interview training by considering the current body of literature.

A. Embodiment and body swapping in VR

In VR, individuals typically experience presence and plausibility. Presence refers to feeling as if one is really inside the VE, while plausibility regards the perception that the events happening inside the VR scene are actually occurring [23], [24]. A clear example of these phenomena is when individuals report high levels of anxiety when exposed to feared stimuli in VR, analogous to the reactions they would have in a similar real-life situation [25]. There are, indeed, many different VR scenarios and contexts in which presence and plausibility play a key role, with people reporting a strong perception of being, e.g., inside a classroom [26], a bar [27], an emergency situation [28], a date [29], in front of an audience [30], etc.

In an immersive VE, individuals can also experience a strong sense of owning a virtual body (body ownership), of being located somewhere within it (self-location), and of controlling it (agency), a phenomenon also known as virtual embodiment. These bodily illusions are achieved through the inclusion of congruent sensory feedback between a real and a virtual body, seen from a 1PP [31]. Such types of virtual bodily illusions are based on the same sensory principles used to evoke, e.g., the rubber hand illusion [32]. Interestingly, a series of studies have shown that the visual appearance of an embodied virtual body might impact individuals' cognition, attitudes, and behaviors [33], a phenomenon also known as the Proteus effect [34]. For instance, according to the studies conducted by Banakou et al. [35] and Peck et al. [36], when Caucasian individuals are embodied in a dark-skinned avatar, there is a reduction of their implicit racial bias, an effect that might last at least three-weeks after exposure to VR and that does not occur when embodying a light-skin avatar [37]. Similarly, Kilteni et al. [38] demonstrated that when individuals are embodied in an informally dressed dark-skinned avatar that looks like a "stereotypical musician", they perform more complex drumming movements, compared to being embodied in a formally-dressed light-skinned avatar or when being represented by floating hands. A further example is represented by another work by Banakou et al. [39], in which it was found that when individuals solve a cognitive task while seeing and controlling a life-size avatar of Albert Einstein from a 1PP, there is an improvement in their executive functions and problem solving skills, as well as a reduction in their implicit biases towards older adults, in comparison to being embodied in a "neutral" virtual body. A detailed review of the impact of embodiment attitudes and behaviors can be found in the work by Seinfeld et al. [40].

In addition to the influence that the visual appearance of a virtual body might have on attitudes and behaviors, a series of studies have also focused on investigating the combination of this phenomenon with the possibility of inducing out-of-body and real-time body swapping illusions in VR. In this regard, it has been shown that it is possible to experience a strong out-of-body illusion both in the real world [41] and in VR [42]. Moreover, Galvan et al. [43] found that synchronous visuomotor or visuotactile feedback between a virtual and real body allows individuals to toggle between 1PP and 3PP without breaking their sense of being embodied in a virtual body. In a more applied context, Osimo et al. [44] and Slater et al. [45] designed a virtual body swapping paradigm where it is possible to establish a self-dialogue in VR from different embodied perspectives. In these studies, individuals first embodied a lookalike avatar and, from this perspective, explained a personal problem to a Sigmund Freud's avatar. Once they finished explaining their personal problem, they were able to swap virtual bodies in order to embody Freud's avatar and look at themselves from the outside. From Freud's avatar perspective, individuals were able to provide themselves with advice on how to solve their problem. This research showed that virtual body swapping to have a self-dialogue compared with only answering pre-recorded questions asked by a Freud's avatar helped individuals to find novel solutions to their personal problems, as well as to improve their mood. A similar virtual body swapping approach has been used by Falconer et al. [22] to increase self-compassion and reduce self-criticism in patients with depression.

These studies provide preliminary evidence that the use of virtual body swapping in VR may be a potentially effective tool for the self-assessment and training of personal skills related to problem-solving. They also provide evidence that under the presence of the indicated multisensory cues, it is possible to swap virtual bodies and feel embodied in them to a certain extent, although it is still not clear how the strength of embodiment is modulated by this real-time perspective changes. Moreover, further studies are needed to understand whether the benefits of seeing oneself from the outside in VR are due to being embodied in a different avatar (e.g., Freud) or more related to having the possibility of looking oneself from a self-distant perspective [46]. In this regard, mental imagery studies suggests that visualizing a difficult situation from a 3PP compared to 1PP leads to deeper insights and a more objective analysis on how to better cope with it [47], [48].

B. *Soft skills and job interview training in VR*

Since VR enables individuals to practice in a safe environment with high ecological validity, it has shown to be an effective tool for the training of specific soft skills such as the ones implicated in public speaking [49], [50]. Moreover, according to Tan et al. [51], VR also seems to be a more effective tool to enhance perspective-taking and empathy when compared to "pen-and-paper" methods. In this regard, Akdere et al. [52] found higher engagement and knowledge transfer when intercultural competence skills were trained in VR compared with traditional video training. This body of evidence suggests that VR might be a unique and powerful tool for the training of interpersonal skills in higher education and industry, since it allows to actively engage in simulated social interactions, where trainees are able to react to the behaviors of different virtual agents while having a conversation with them [53].

In the real world, a high number of individuals experience anxiety and discomfort when faced with job interviews, since these are evaluative social interactions where they are typically interviewed by a stranger and lack control over the situation [54], [55]. Based on the above-mentioned advantages of VR simulations in helping individuals to cope with social situations that evoke anxiety [56], immersive virtual scenarios are a very promising tool in the context of job interview training [11]. However, most studies performed so far have focused on understating the impact of VR social simulations on negative emotional states based on repeated exposure therapy and habituation mechanisms to the feared situation [12], [56], while the impact of visual perspective changes on anxiety and other emotional states remained largely unexplored.

From mental imagery studies, there is evidence that thinking about a personal difficult situation from 1PP, in contrast to 3PP, increases emotional arousal and exacerbates negative ruminations [57]. This seems to be explained by the fact that events that are mentally pictured from a 1PP, i.e. self-immersed perspective, involve a bottom-up thinking style where individuals seem to focus more strongly on the specific sensory characteristics and constituent aspects of the imagined situation. In contrast, picturing an event from a 3PP seems to involve a top-down thinking style, in which individuals are more prone to interpret the event through its conceptual meaning and considering its broader context [58]. Therefore, it might be the case that having a 3PP in VR might help individuals to self-distance from an anxiety-evoking situation, leading to more positive emotions compared to 1PP.

The summarized studies highlight that there is still a lack of research on how VR might be useful in the training of other types of more complex interpersonal skills like those involved in a job interview. Moreover, it is still unclear what is the actual impact of specific VR design elements on the learning of soft skills and their emotional outcomes [6]. In this regard, to the best of the authors' knowledge, the potential benefits of virtual body swapping and self-distancing in VR for the training of soft skills related to a job interview situation has still to be investigated. Based on the reviewed literature [17], such elements might represent unique and powerful tools to self-assess verbal and non-verbal performance, as well as to regulate negative emotions, in simulated social situations like those this work aims to explore.

III. Material and Methods

This section describes the experimental design used, the different measures included in the study, and the software tool developed to support it.

A. *VR-based job interview training environment*

A VR-based job interview training environment (later abbreviated VR-JITE) was devised to carry out the study. In the simulated VE, the trainees are able to control, in real-time, their avatar through their body in order to interact with the scene in a seated pose. Furthermore, they are allowed to interact with and provide answers to the questions asked by a recruiter avatar who sits in front of them separated by a desk. At the beginning of the virtual experience, the trainees could choose a male or a female avatar for the interviewee role, whereas a single avatar was used for the recruiter role. The layout of the VE, depicted in Fig. 1, mimics an office room with some pieces of furniture and a mirror that was meant to let the trainees observe their own avatar. Free-of-charge assets and custom-made 3D models created with Blender [59] modeling and animation suite were used to populate the environment. The avatar models were generated with Reallusion's Character Creator [60] and further customized with Blender [59].

The VR-JITE was implemented by leveraging the Unity (v2022.3 LTS) [61] game engine together with the SteamVR framework (v2.7.2) [62], and considering the HTC Vive Pro [63] as the target immersive VR head-mounted display (HMD). Its tracking technology leverages infrared laser emitters in the room and supports six degrees-of-freedom for each tracked object. Since non-verbal cues were considered relevant

Fig. 1. Layout of the virtual office room.

for the scenario, particular care was devoted to allow the trainees to control their avatar's virtual body with their real-world movements. Given the fact that only seating avatars were considered, tracking was limited to upper body. Specifically, the trainee's hands were tracked by means of Valve Index controllers [64], and two additional Vive trackers (v2.0) [65] were strapped on the trainee's elbows to increase the arm tracking accuracy. The tracking information of these tracked devices (HMD included) were combined using the FinalIK (v2.2) Unity plugin [66]. In particular, its module named *FullBodyBipedIK* was exploited, constraining its lower body IK controls (e.g., feet, knee and hip positions) with information gathered in a calibration phase (explained below). To control the recruiter avatar, a Wizard-of-Oz approach was adopted. This approach was preferred over alternative methods such as using chatbots, as done in the work by Stanica et al. [14], where the goal was to find the most effective way to control the interviewer avatar (outside the scope of the present work). Moreover, the use of Wizard-of-Oz also permitted a higher degree of control during the experimental evaluation, avoiding potential confounding factors. To minimize bias among trials and study participants, the "wizard", i.e. the experiment administrator, was allowed to control the avatar through a simplified interface designed to trigger events (Fig. 2). Events included asking the interviewee a question among a set of six predefined ones by playing a pre-recorded audio clip (see questions in Section III-B4), and triggering pre-recorded deictic animations, such as nodding or typing at a notebook on the desk. The movements of the avatars' lips, either of the interviewee or of the recruiter, were implemented using the SALSA Unity asset, which generates lip-sync animations from audio sources (audio files or microphone input).

The VR-JITE can be run in three modes. The *calibration mode* is used for configuration purposes, to match the participant's body to that of the user-controlled avatar. This mode asks the participant to assume a T-pose, a T-pose with fully bent arms (thumbs touching the chest), and a seated upright pose with the palms on the knees. From these poses, information about the participant's height, wrist-elbow and elbow-shoulder distance, as well as hip-head and hip-ground distance is extracted. This information is used to make the FinalIK asset produce visually appropriate results.

In the *Job interview practice mode*, the wizard controls the recruiter avatar, and the participant answers the questions of the job interview by controlling the interviewee avatar in real time. During the experience, the movements of the avatars and the answers of the participant are recorded. The recording feature was implemented by customizing the Record and Play Unity asset [67]; in particular, the built-in features of the asset were extended to let it record the movements of all the objects in the VE (avatars included), as well as to store an audio clip created from the participant's utterances.

The recordings are later used in the *Replay mode*, where the participants can replay and observe their behavior and job interview responses from different POVs.

B. Experiment design and procedure

1) Experimental Design: The experiment consisted of a fully counterbalanced within-subjects design with three main conditions: the *No Swap* condition, consisting in a baseline configuration in which the participants embodied an interviewee avatar and had to answer job interview questions asked by a virtual recruiter (Fig. 3a); the *Out of Body* condition, in which, first, the participants answered the questions of the recruiter from the interviewee perspective, then self-assessed their responses (including both utterances and movement recordings) by swapping to an external non-embodied configuration where they saw the scene from a virtual camera

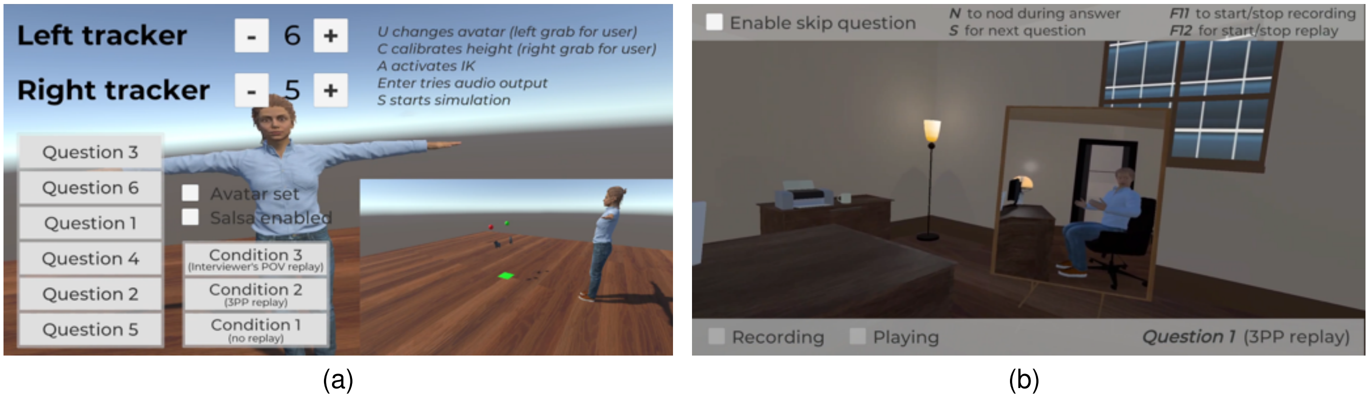


Fig. 2. Administrator interface: (a) configuration options, and (b) Wizard-of-Oz interface with the controls provided (upper and lower panel) and the current first-person view of the user (center panel).

located near the wall, halfway from both the avatars (Fig. 3b); lastly, the *Recruiter* condition, in which, after answering the questions from the 1PP of the interviewee, the participants were able to swap to the visual perspective of the recruiter to self-assess their performance (Fig. 3c). The aim of the study was to explore the effectiveness of self-assessing verbal and non-verbal performance in a job interview simulation in the considered experimental conditions. Moreover, the experiences were also evaluated in terms of perceived degree of embodiment, presence, and plausibility, as well as with respect to their emotional impact in terms of anxiety, valence and arousal.

The reasoning for opting for a within-subjects design was based on the control of potential confounding variables related to inter-individual differences [68], since there is evidence that personal factors such as self-efficacy perceptions, degree of motivation, emotion regulation skills, cognitive abilities, and prior job interview experiences play an important role in how interviewees feel and cope with job interviews, as well that as in their baseline self-assessment abilities [54]. Furthermore, a fully counterbalanced design was also used to further control for order effects.

2) Experimental conditions: The experimental conditions and the reasons behind their selection are illustrated below.

- *No Swap:* This condition consists of a single step in which the participants listen to the recruiter and answer the questions from the perspective of the interviewee. Differently than the other two conditions, this condition does not include a replay or re-experience of the answers provided from an external POV, since it represents the current state of the art of job interview simulations in VR, where there is no perspective swap. Moreover, it also represents the equivalent of a job interview in the real world, therefore acting as a baseline condition. In the other two conditions, each interview question and answer is listened to twice: a first time during the practice, and a second time during the replay where the participants visual perspective is changed. Therefore, since in this condition there is no perspective swap, the step described above is repeated, so that the questions are practiced twice. The choice of a second exposure ensures a fully counterbalanced design and accounts for potential learning effects related to being exposed to the practiced content more than once.
- *Out of Body:* The first step of this condition is the same as in the *No Swap* condition: the recruiter asks a question, and the participant provides an answer. Then, the perspective is changed to an non-embodied external POV. The camera is positioned at eye level next to the virtual desk between the two avatars, as shown in Fig. 3b; the fact that the point of view is external implies that the participants do not have any virtual body. From this non-embodied perspective, the participants watch the recording of the recruiter asking the current question and of them answering it. As said, according to the state of the art, it is speculated that while experiencing the scene from an external perspective, the participants should feel "safer" [69], getting able to self-distance from themselves in order to manage anxiety and decrease it [70], [71]. It was expected that this condition could let the participants learn how to assess their performance from a detached perspective.
- *Recruiter:* Similarly to the previous one, in this condition first the participants practice the answer to a job interview question from the perspective of the interviewee and, then re-watch themselves being asked and answering the same question. However, in this case, they see their own avatar providing answers from the perspective of the recruiter, as shown in Fig. 3c. Importantly, they can control the movement of the recruiter avatar in real time through their own body movements, in order to better trigger the embodiment towards it [31]. The aim was to test if this condition could lead to self-distancing and to different outcomes when the participants self-assess their performance in the job interview. More specifically, the goal was to explore whether the perception that the participants have of themselves and the effectiveness of their self-assessment is altered or not by the fact that they see themselves from the perspective of the person with the role of judging the outcome of the interview and their suitability for a job position, comparing it to the other conditions.

3) Procedure: After the calibration and prior to entering the first experimental condition, the participants were instructed

(a)

(b)

(c)

Fig. 3. Conditions and POVs: participant (a) answering a question from the 1PP of the interviewee (No Swap Condition, and first step of the other experimental conditions), (b) reviewing the provided answer from an external POV (Out of Body Condition), and (c) reviewing the provided answer from the recruiter perspective (Recruiter Condition).

to get familiar with their virtual body (avatar) and the virtual environment, also taking advantage of the mirror in the room (Fig. 4). First, they were invited to gaze around from their seated position, observe the virtual environment, and verbalize what they saw. Afterwards, the experimenter who was executing the study in the real world encouraged the participants to look at themselves in the virtual mirror, focusing on what happens in the virtual environment when they move, and check the coherency with the avatar movements. This operation was expected to enhance embodiment through visuomotor correlations, as discussed by Slater et al. [45].

Each experimental condition started with a phase where the recruiter asked a question that the participant should answer through voice and body gestures, as it can be seen in Fig. 3a and is described in Section III-A for the *Job interview practice mode*. As said, all the information required to later replay the answer was recorded. Then, in the conditions that involve a perspective swap, namely the *Out of Body* and *Recruiter* ones, the participants went through a second phase in which they were able to watch the recording of the recruiter asking them the question and of themselves answering it, as described in the above section for the *Replay mode*.

For each condition, two job interview questions were asked; thus, at the end of the experiment, each participant had answered a total of six questions.

Fig. 5 presents an overview of the experimental design and the procedure followed. As shown, in all the conditions, for each job interview question, first the participants provide an answer from the POV of the interviewee. Afterwards, depending on the experimental condition they either practice again (*No Swap* condition), or see themselves answering the question from an out-of-body POV or from the POV of the recruiter. Moreover, to limit possible biases in the evaluation, the occurrence of each question and each condition had to be balanced. Thus, Latin-square order was followed, as reported in Table I. The pattern was repeated every six participants. This ensured that the results did not depend on the precise sequence of questions and conditions, or on the question-condition pair.

4) Questions of the interview: The questions selected for the job interview simulation were drafted as a result of interviews with experts in the field of Human Resources, who suggested the main topics that come up in these situations. This group of experts was composed by four women with a mean age of 36.35 (SD=3.69). Three of them had postgraduate education in the field of Human Resources and one in the field of Psychology. They had an average of nine years (SD=4.55) of professional experience working as recruiters. The six questions are reported below:

1) Could you tell me about yourself and briefly describe your studies and professional past experience?
2) Do you prefer working independently or on a team? Why?
3) What did you do in the last year to improve your knowledge?
4) When you are balancing multiple projects, how do you keep yourself organized?
5) How do you deal with pressure or stressful situations?
6) What type of work environment do you prefer?

Considering that the participants were all native Italian speakers, during the experiment the questions were asked using their native language, thus avoiding the anxiety connected to language barriers which may have led to the introduction of a confounding factor.

A video showing the whole procedure, including the familiarization phase and the three conditions is provided as supplemental material (http://tiny.cc/xk6evz).

C. Sample

Since the scenario chosen for the experiment is a job interview and the study has been devised in a university context, the main target of the evaluation is represented by students in their

(a) (b)

Fig. 4. Familiarization phase: participant using the mirror in the (a) virtual and (b) real world.

TABLE I
ORDER IN WHICH NO SWAP, OUT OF BODY, AND RECRUITER CONDITIONS AND QUESTIONS WERE ADMINISTERED DURING THE EXPERIMENT. EACH ROW REFERS TO A PARTICIPANT: ENTRIES OF EACH ROW INDICATE THE EXPERIMENTED CONDITION AND THE ORDER IN WHICH QUESTIONS WERE ASKED. THE QUESTIONS ARE IDENTIFIED BY THE NUMBERS (Q1, Q2, ETC.) USED IN SECTION III-B4.

Id	1st cond./quest.	2nd cond./quest.	3rd cond./quest.
1	No Swap (Q1-Q6)	Out Of Body (Q2-Q5)	Recruiter (Q3-Q4)
2	Recruiter (Q2-Q5)	Out of Body (Q3-Q4)	No Swap (Q1-Q6)
3	No Swap (Q3-Q4)	Recruiter (Q1-Q6)	Out of Body (Q2-Q5)
4	Recruiter (Q4-Q3)	No Swap (Q6-Q1)	Our of Body (Q5-Q2)
5	Out of Body (Q5-Q2)	Recruiter (Q4-Q3)	No Swap (Q6-Q1)
6	Out of Body (Q6-Q1)	No Swap (Q5-Q2)	Recruiter (Q4-Q3)

last year, possibly searching for a job. Nevertheless, practicing for a job interview could be useful also to other students as well as to individuals who have completed their studies or even work already. For this reason, while recruiting the sample no restrictions were set on the participation requirements.

An a-priori power analysis was performed using the *G*Power* tool [72] to determine the required sample size. Setting $\alpha = 0.05$ and aiming to detect at least an effect size of medium entity ($\eta^2 \geq 0.06$), it was found that a sample of 23 participants was adequate to reach a power of (1-β) $= 0.80$ for the arranged study design. Moreover, a sample size of 24 participants ($\eta^2 = 0.067$) ensured a fully counterbalanced order of questions and conditions, since it is a multiple of six (Table I). Demographics details of the sample are reported in Section IV.

D. Measures

A total of five questionnaires were included in the study in order to evaluate the impact of the different experimental conditions. Before starting the experiment, the participants filled in a demographic form, and provided the baseline measures for the Self-Assessment Manikin (SAM) tool [73] and for a short version of the State-Trait Anxiety Questionnaire (STAI) [74]. After each experimental condition, the participants completed again the SAM and STAI tools, as well as the post-condition questionnaire, containing several questions related to the VR experience and the perceived usefulness of that specific condition to improve the self-assessment of job interview skills. At the end of the experiment, when they had experienced already all the conditions, the participants also answered a final questionnaire including questions related to job interview experiences in the real world, to social anxiety, and requesting them to compare the different experimental conditions in terms of effectiveness for self-assessment. In the following, the details of each questionnaire are provided. The complete post-condition and final questionnaires are also provided in the Appendix, included as supplemental material.

1) Demographic form: Through this form, the participants were requested to provide demographic information such as age, gender, professional status (e.g., BSc, MSc, or PhD student, post-doc researcher, employee, other), field of expertise, level of Italian and English, as well as to indicate their familiarity with programming, video-games, and VR experiences. The information gathered through this form is summarized in the Appendix. As part of the demographic information gathering, some questions from the Social Phobia Inventory (SPIN) were also included to assess the degree to which the participants experienced social anxiety in their daily life.

2) Post-condition questionnaire (PCC): After each condition, the participants were administered a questionnaire to evaluate the following aspects: degree of experienced embodiment, presence, and plausibility, as well as perceived usefulness of self-assessing verbal and non-verbal skills elicited in a job interview situation. Moreover, even though the emotional impact of the tested conditions was measured through the SAM and STAI tools, in the post-condition questionnaire several questions were included to better understand the type of emotions that were experienced in each condition.

The questions used to assess embodiment were adapted from the ad-hoc questionnaire developed for the study on counseling in VR through embodied self-dialogue by Slater et

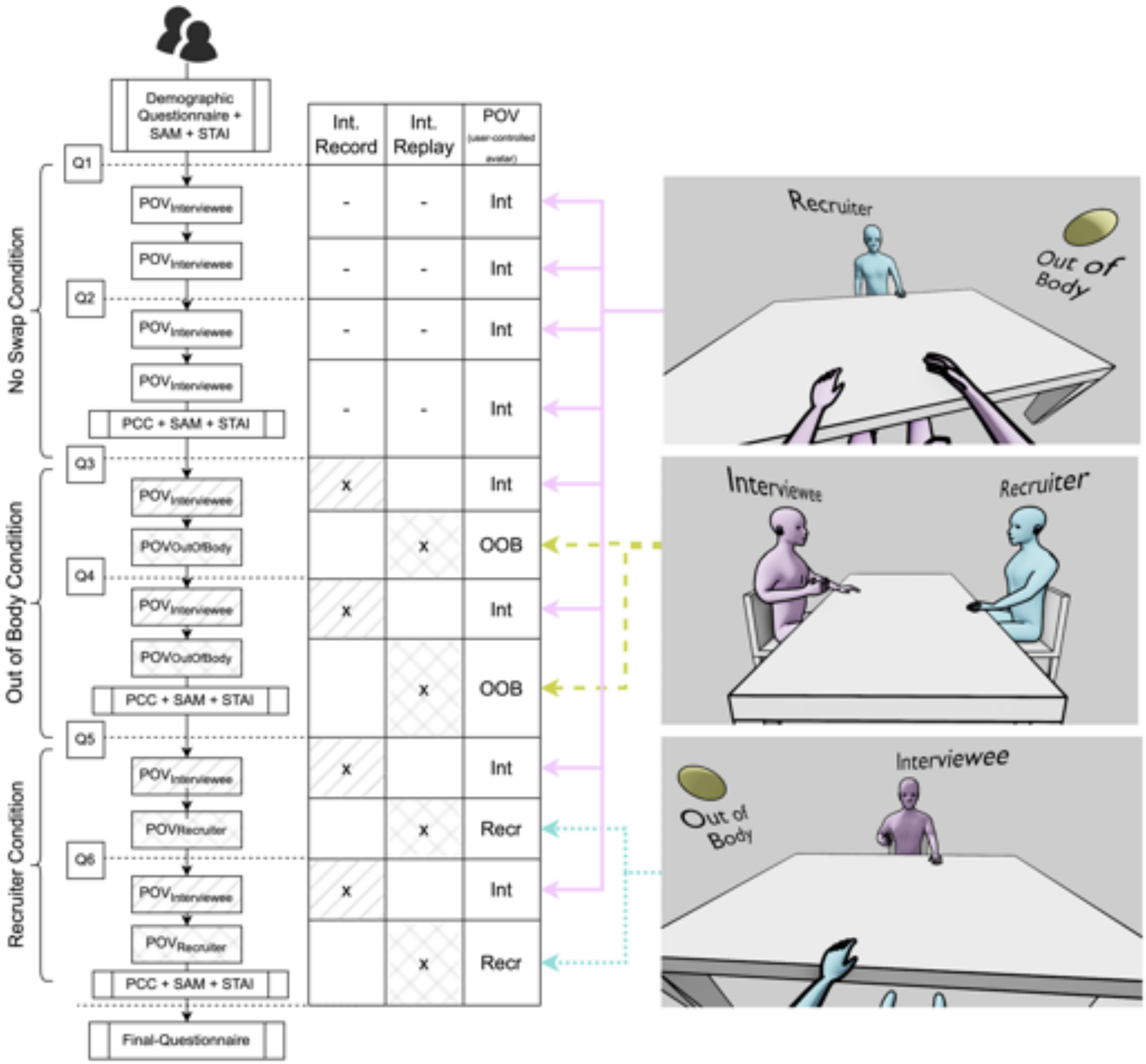


Fig. 5. Overview of the experiment design with conditions, POVs, and record/replay phases.

al. [45] and from a standardized questionnaire on avatar embodiment by Peck et al. [75]. More specifically, the questions evaluated three aspects that have been associated with the level of experienced embodiment towards an artificial body, namely experienced body ownership, agency, and self-location towards the virtual body of both the interviewee and the recruiter during the *Job interview practice mode* (i.e. answering job interview questions embodied in the interviewee) and *Replay mode* (i.e. watching the responses they provided in the different experimental conditions).

Questions related to presence and plausibility were also included and adapted from [45], and mainly focused on the extent to which the participants felt inside a job interview situation and the degree of perceived realism.

The usefulness of the experience was measured with ad-hoc questions that asked the participants to indicate to what extent they felt that the different experimental conditions could help them improve their job interview skills, as well as objectively self-assess their verbal and non-verbal responses to the job interview questions.

All the questions in this questionnaire (provided in the Appendix) had to be answered on a 7-point Likert scale.

3) Self-Assessment Manikin: The SAM [73] was selected as a quick and easy method to register the emotional state of the participants. It is a picture-oriented tool that enables to assess the valence, arousal, and dominance of a person's emotional reaction to a high range of stimuli. Valence indicates whether the elicited emotion is positive (high values) or negative (low values). Arousal measures how much the elicited emotion is “activating”, i.e., how much the person is excited (high values) or calm (low values). Finally, dominance indicates whether the person feels “in control” of the emotion (high values) or controlled by it (low values). The SAM tool can be administered using different scales: a 5-point scale, a 7-point scale, or a 9-point scale. For this experiment, the 9-point scale was chosen with the aim to give the participants as many options as possible to express their emotions. As said, the participants used the tool before the experiment and after going through each experimental condition.

4) State-Trait Anxiety Questionnaire: The STAI [76] consists of a validated instrument for measuring anxiety in adults. It differentiates the “state anxiety”, which is a temporary condition, from the “trait anxiety”, which, on the other hand, is a more general and long-standing condition. This questionnaire is used worldwide and it includes 40 questions with a range of four possible answers each. However, for this experiment, the shortened version validated by Marteau et al. [74] was used, including only six items. The purpose of including this questionnaire was to compare how much the participant was anxious before starting the experiment and after each condition, in order to understand whether the VR simulation affected this emotional state.

5) Final questionnaire: At the end of the experience, the participants were requested to answer some general, ad-hoc questions on job interviews in a real context. In particular, they were asked how many job interviews they had performed in the past, how good they considered themselves in answering job interview questions, and how anxious they normally felt in these situations.

Finally, four last questions requested the participants to establish a ranking indicating their order of preference (1st, 2nd, 3rd) for each of the experimental conditions in terms of the extent they could help them to improve their job interview skills in the future, as well as to assess their verbal and non-verbal responses more objectively. They were also asked to rank their overall preference for the different experimental conditions. These ranking questions were meant to complement the answers provided by the participants about the usefulness of the experience through the post-condition questionnaire which were given immediately after each experimental condition. However, in the case of these ranking questions, the participants were able to directly establish comparisons between the three conditions because they had already experienced them all.

Finally, they were asked to provide two positive comments and two negative comments on the experiment.

E. Statistical analysis approach

Due to the ordinal nature of the data, it was decided to use non-parametric tests for the statistical analysis. More specifically, Friedman tests [77] were used to compare the experimental conditions. In case there was a significant difference in an independent variable, post-hoc tests were based on Wilcoxon signed-rank tests [78] with Bonferroni corrections [79].

IV. Results

A. Demographic characteristics

The participants were aged between 21 and 31 ($M = 25.2$ y.o., $S.D. = 2.5$ y.o.); 14 were males, 10 females. All the participants were proficient in Italian. Most of them where BSc and MSc students (62,5%), though the sample also included PhD students (12,5%) and employees (20,83%). Most of the participants were studying or had an expertise in the field of engineering (20,83%). Therefore, many of them had good or at least some experience with programming; however, more than half of the sample had none to very little experience with VR (58,34%). In terms of social anxiety in everyday life, measured through SPIN, 37,5% of the participants reported not to suffer from social anxiety, 16,67% to suffer from mild anxiety, 37,50% from moderate anxiety, and only one participant reported to suffer from severe anxiety; hence, there was a rather varied range of experienced social anxiety represented in the sample. Overall, the participants had been exposed to an average number of 3.66 (SD=5.55) job interviews in the past. Spearman correlation analysis [80] indicated that SPIN scores and past job interview experience was not correlated with the outcomes of the post-condition questionnaire, SAM, and STAI measures.

B. Embodiment

Questions related to the degree of experienced embodiment either referred to the *Job interview practice mode*, in which first participants answered the job interview questions from the interviewee POV, or to the *Replay mode*, which consisted in re-experiencing the responses provided to the questions in the *Job interview practice mode* from either the out-of-body or the recruiter POVs. The results were analyzed taking into account these two modes.

1) Job interview practice mode: No significant differences in terms of embodiment were found in the first phase of each condition (Fig. 6). Overall, self-reported body ownership, agency, and self-location scores were high with respect to the virtual body of the interviewee regardless of the experimental condition, with medians greater or equal to five on a 7-point Likert scale. No embodiment was reported, in turn, for the virtual body of the recruiter. This result was largely expected, since in this mode the participants only experienced the POV of the interviewee when answering the questions and did not swap visual perspective.

2) Replay mode: No significant differences between the *Out of Body* and *Recruiter* conditions were found in the degree of experienced embodiment for the virtual body of the interviewee in the second phase of each condition, which consisted in re-experiencing one’s own job interview performance from an external POV. Median scores were relatively high for body ownership and agency, but low for self-location (Fig. 7). These findings confirm that, in agreement with what found by Galvan et al. [43], it is possible to feel embodied, to a certain extent, in a virtual body seen from a 3PP after having previously

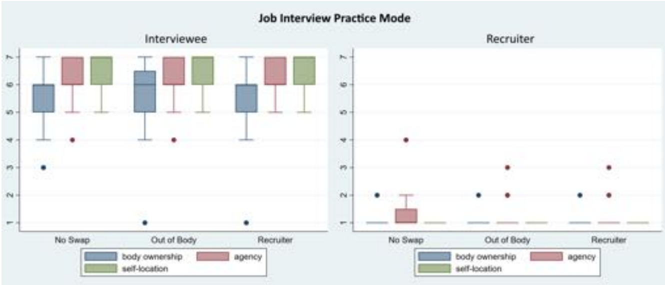


Fig. 6. Box plots with medians and interquartile ranges of body ownership, agency, and self-location for the Job interview practice mode. In this first phase of each condition, the participants answered the questions asked by the recruiter from the POV of the interviewee.

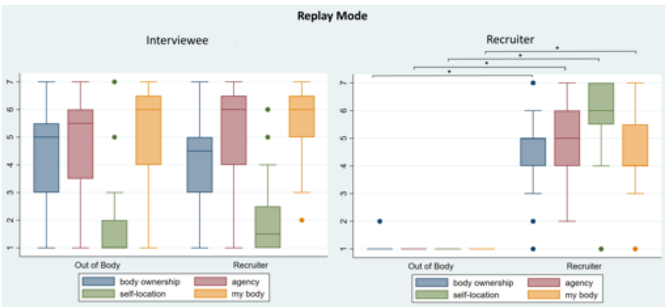


Fig. 7. Box plots with medians and interquartile ranges of body ownership, agency, and self-location for the Replay mode. In this second phase of each condition, the participants either practiced again their interview responses or watched a replay of their responses from an external POV or from the POV of the recruiter. The asterisk (*) marks statistical significant difference with 95% confidence interval.

experienced the same virtual body from a 1PP (i.e. that of the interviewee).

With respect to the virtual body of the recruiter, in this phase the participants reported significantly higher body ownership ($z = -4.28$, $p < 0.001$), agency ($z = -4.33, p < 0.001$), and self-location ($z = -4.30$, $p < 0.001$) in the *Recruiter* condition compared to the *Out of Body* condition (Fig. 7).

C. Presence and plausibility

No significant differences among conditions were found in terms of the extent to which the participants felt as if they were in a job interview (presence) having a conversation with a recruiter (co-presence). Scores assigned by the participants to questions regarding these dimensions were high, with a median of six on a 7-point Likert scale. This finding further corroborates the effectiveness of VR to simulate a situation with high ecological validity, where students can practice their interpersonal skills feeling like in the real situation. Similarly, the degree of perceived realism of the situation and the virtual recruiter did not differ among the conditions, and also had relatively high self-reported scores. Details are provided in Fig. 8.

D. Usefulness to support job interview skills training

No significant differences were found between the *No Swap*, *Out of Body*, and *Recruiter* conditions in terms of the extent

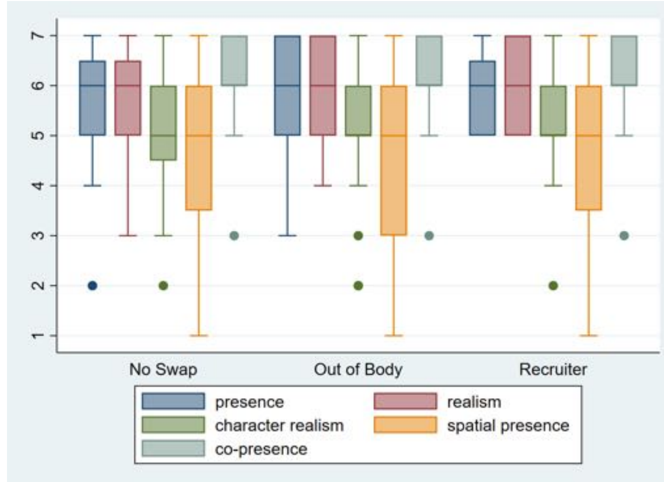


Fig. 8. Box plots with medians and interquartile ranges of questions related to presence and plausibility.

to which the participants considered them a useful tools to support the training of their job interview skills. In particular, scores were high for all the conditions (Fig. 9). However, there were significant differences for what it regards the extent to which the participants considered the three conditions an effective tool to potentially self-assess, in an objective way, their verbal ($x^2 = 9.59$, $p = 0.008$) and non-verbal language ($x^2 = 27.71$, $p < 0.001$) in a job interview. Post-hoc analysis with Bonferroni correction indicated that the *Recruiter* condition was more effective for self-assessing verbal ($z = -2.93$, $p = 0.003$) and non-verbal responses ($z = -4.28$, $p < 0.001$) when compared to the *No Swap* condition. The *Out of Body* condition was also more effective to self-assess non-verbal ($z = -3.70$, $p < 0.001$) language in a job interview than the *No Swap* condition, but this difference was not clear for the assessment of verbal aspects, where there was only a statistical trend ($z = -1.93$, $p = 0.053$). No significant differences were found between the *Out of Body* and *Recruiter* conditions for what it concerns the self-assessment of non-verbal language; there was, in fact, a significant difference showing that the *Recruiter* condition was a better tool than the *Out of Body* condition for self-assessing verbal responses, but this difference was not significant anymore after Bonferroni correction ($z = -2.14$, $p = 0.033$).

When the participants were asked to rank the conditions they judged as more effective to support the training of their job interview skills and to self-assess their verbal responses, non-verbal language, and overall (Table II), more than half put the *Recruiter* condition in the first place. As for the second place, most of the participants chose the *Out of Body* condition, putting the *No Swap* in the last position. These findings are in line with the results summarized in the previous paragraph based on Likert-type questions, and confirm that the most effective VR condition for the self-assessment of job interview skills is the *Recruiter* condition. Despite that, some participants still judged as more effective and useful the *Out of Body* and *No Swap* conditions.

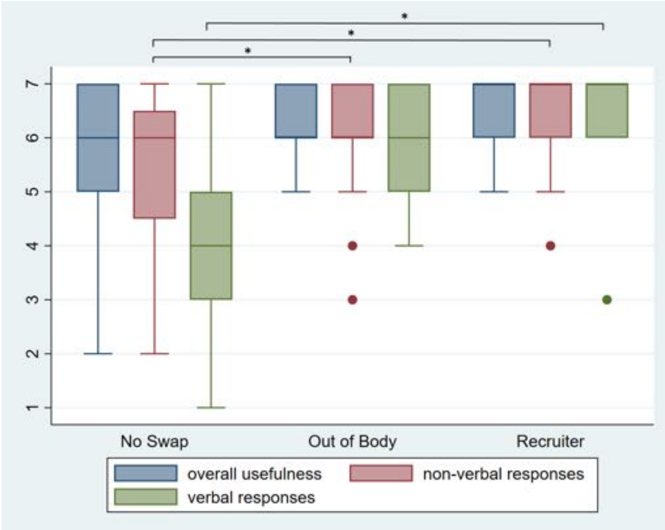


Fig. 9. Box plots on the usefulness of the experience. The asterisk (*) marks statistical significant difference with 95% Confidence Interval.

TABLE II
PERCENTAGE OF PARTICIPANTS (IN NUMERICAL VALUES) THAT RANKED EACH OF THE EXPERIMENTAL CONDITIONS (NO SWAP, OUT OF BODY AND RECRUITER) AS THE FIRST, SECOND, AND THIRD MOST EFFECTIVE TOOLS TO IMPROVE JOB INTERVIEW SKILLS, SELF-ASSESSMENT OF VERBAL AND NON-VERBAL SKILLS, AND THE MOST USEFUL OVERALL.

	Improve job interview skills	Self-assessment of verbal language	Self-assessment of non-verbal language	Overall usefulness
	1st			
No Swap ranked 1st	13% (3)	29% (7)	0% (0)	17% (4)
Out of Body ranked 1st	21% (5)	17% (4)	38% (9)	33% (8)
Recruiter ranked 1st	67% (16)	63% (15)	63% (15)	63% (15)
	2nd			
No Swap ranked 2nd	25% (6)	25% (6)	0% (0)	25% (6)
Out of Body Rank 2nd	50% (12)	46% (11)	63% (15)	38% (9)
Recruiter ranked 2nd	29% (7)	21% (5)	38% (9)	29% (7)
	3rd			
No Swap ranked 3rd	63% (15)	46% (11)	100% (24)	58% (14)
Out of Body ranked 3rd	29% (7)	38% (9)	0% (0)	29% (7)
Recruiter ranked 3rd	29% (7)	21% (5)	38% (9)	29% (7)

E. SAM and STAI

In terms of valence, arousal, and dominance, measured through the SAM tool, no significant differences were found. Russell's circumplex model of affect [81] applied to the obtained results shows that most of the emotions experienced in all the conditions were positive (Fig. 10). No relevant differences in terms of anxiety responses were observed among the three conditions, as evidenced through the STAI.

V. DISCUSSION AND CONCLUSIONS

The findings of the present study advance knowledge on how visual perspective changes and embodiment in VR might modulate the self-assessment of soft skills, specifically in the context of a job interview simulation. The results indicate that trainees were able to more effectively self-assess their verbal and non-verbal communication skills as well as their overall performance when they had the possibility to re-experience their responses to the job interview from either an external, non-embodied POV (*Out of Body* condition) or from the embodied POV of the virtual character that was "evaluating" them (*Recruiter* condition), compared to practicing twice from

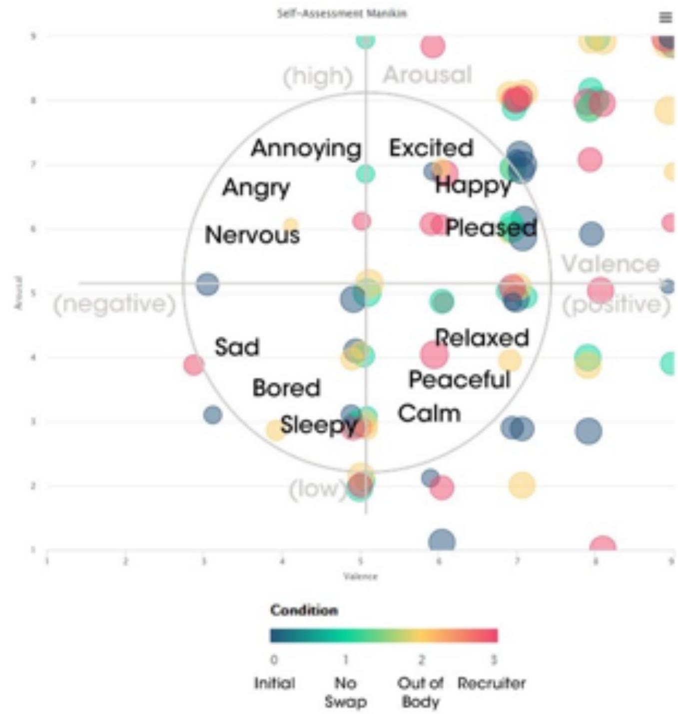


Fig. 10. Russell's circumplex model of affect applied to the results of the SAM tool.

the POV of the interviewee (*No Swap* condition). Based on the participants' ratings and preferences, the possibility to adopt the POV of the recruiter seems to present further benefits to self-assess job interview performance compared to a non-embodied, out-of-body POV. Interestingly, throughout the study, self-reported embodiment scores remained high despite swapping to different virtual bodies in real time (i.e. interviewee and recruiter). No differences in emotional state were found between the conditions, with most emotions reported during the study having a positive valence.

Self-assessment can be defined as a wide range of mechanisms and techniques through which individuals are able to assess and evaluate the qualities of their own learning processes and products, as well as their competences, with the main goal of using these evaluations to improve specific behaviors and cognitions in the future [17]. Self-assessment has proven to be a critical process for the development of meta-cognitive capabilities that can help learners to improve their acquisition of new skills, as well as to accomplish complex goals [82]. A body of research has shown that being able to effectively self-assess performance can positively impact self-monitoring, emotional regulation, and self-efficacy [83]. In the context of a job interview, having the opportunity of practicing and self-assessing performance is important since it can help learners to gain confidence, as well as to improve and develop skills that might be critical for accomplishing professional goals such as getting a desired job position [4]. In this regard, the findings of the present study advance knowledge on how visual perspective changes and virtual body swapping (i.e. embodiment) in VR might be leveraged to enhance self-assessment and self-observation in a realistic job simulation, with potential implications for the training of other soft skills that future studies should research, such as leadership, public speaking, assertiveness, or negotiation skills, among others.

A similar approach used to capture, review, and assess performance in the real world is based on video-recordings. Video-recordings have been successfully used, for instance, to help learners self-assess their oral presentations, with some studies by Ritchie [82] as well as Tailab and Marsh [16] showing positive effects. Mentally self-distancing from a situation is another strategy that can be used to better regulate emotions and gather deeper insights into how to cope with difficult circumstances [47], [48]. However, based on existing research, it is not clear whether the effectiveness of observing and reviewing oneself from a 3PP might be differently modulated by adopting a non-embodied or embodied perspective (e.g., external observer, audience, or recruiter). This is a question that the present study has partly addressed, since it was found that most of the participants felt they could better self-assess their verbal and non-verbal communication skills when they re-experience and evaluated the situation from an embodied perspective (i.e. that of the recruiter) compared to only having an external view of the situation. The positive effect of being embodied in the recruiter for self-assessment might be related to an effect documented by several past studies, where the semantic properties of a virtual environment (interview situation) and an embodied avatar where proved to have an effect on making the participants adapt their perceptions and behaviors to conform to the characteristics and the role played by the avatar they were embodying in the scene, the so-called Proteus effects [34]. Thus, it may be the case that, somehow, the participants of the present study assumed that they could better assess their performance when they adopt the role and exact POV of the person who was previously evaluating them, namely the recruiter.

A study similar to the present one is that by Zhou et al. [8], in which the participants could re-experience and self-evaluate their performance in an oral presentation carried out in VR from the audience's perspective. This condition was compared to virtual and real-world 2D video recordings. The experiments showed that participants with low self-confidence in their public speaking skills seemed to benefit the most from re-experiencing their talk in VR, gaining more confidence and more objectively self-assessing their performance, with no other personal characteristic or gender playing a role in the findings. However, to the best of the authors' knowledge, no similar studies have been carried out in the context of simulating a job interview, and also past studies have not compared how self-assessment might be modulated by having or not an embodied perspective when re-experiencing performance, aspects that the present study has addressed. Moreover, Zhou et al. did not simulate facial movements such as lip synchronization or used precise arm tracking, while in the present study special care was devoted to include more faithful body tracking systems in order to foster ownership and agency over the virtual bodies.

Regarding emotional aspects, it is common for individuals to experience high levels of anxiety, frustration and distress before facing a job interview due to its evaluative and competitive nature. Moreover, job interviews are normally carried out by strangers and it is impossible to know in advance what exact questions will be asked, factors that can exacerbate anxiety in some individuals [84]. Exposure to immersive VR scenarios has been shown to be an effective technique for overcoming situations related to social anxiety, with similar results to those obtained by in-vivo exposure therapy [56]. Previous studies have shown that job interview simulations in VR can lead to a strong sense of presence, which in turn can lead participants to experience high anxiety and learn how to better control it. The findings of the present study further corroborate some of these effects, with participants reporting high levels of experienced presence, realism, and co-presence. However, emotions reported in this study were mainly positive. It is possible that the participants did not experience strong distress or anxiety, or managed to control these feelings throughout the experiment, since all the conditions encompassed strategies that have been found to reduce anxiety in the real world, such as repeated exposure and self-distancing [56]–[58]. Moreover, as outpointed by social anxiety scores reported in the Appendix, the sample used in the present study included participants who experienced a diverse level of anxiety in social situations. Future studies should be carried out to better understand whether the impact of this type of VR-based training could differently impact and benefit participants with lower or higher levels of anxiety to job interviews.

In terms of embodiment, the present study confirmed that swapping to different virtual bodies in real time can still lead to experiencing strong ownership and agency for different virtual bodies, a result that is in line with existing literature [22], [44], [45]; moreover, individuals can also feel embodied to some extent in a body they see from a 3PP but that previously experienced from a 1PP, as found by Galvan et al. [43]. Importantly, in all these studies, the critical aspect that seems to favor embodiment is the presence of congruent multisensory feedback between a real body and an artificial one. Further research should be carried out in order to explore the possible use and benefits of virtual body swapping in the training of soft skills.

Despite these findings regarding the potential of visual perspective changes in VR to improve self-assessment in job interviews, it is important to highlight that self-assessment might involve several phases, ranging from identifying potential areas of improvement and personal abilities, to the development and execution of a plan to improve weaker areas and consolidate strengths [17]. In the case of the present study, in spite of the suitability of the developed VR application to also let trainees develop and execute a plan for enhance job interview performance through repeated practice (i.e. answering the job interview questions more than once), it was decided to only focus on the first phases of self-assessment. In particular, the aspects pertaining to emotions, embodiment and potential usefulness of the application to support the trainees in analyzing their verbal and non-verbal communication abilities were investigated.

However, it worth noting that the possibility to practice more than once the answers provided to the job interview questions was not included in the *Out of Body* and *Recruiter* conditions in order to avoid the potential confounding effect of familiarity and practice, since the main goal was to understand what type of visual perspective was more effective for self-assessing verbal and non-verbal behaviours, independently of training. With training in mind, future studies could consider allowing the trainees to provide feedback to themselves (spoken or written), and letting them use this information to improve their future performance.

Even though the present study was based on a fully counter-balanced repeated-measures design to carefully control for participants' inter-individual differences and order effects, there could still be carry-over effects related to how experiencing the different conditions might influence the judgments made over other conditions. Future studies should investigate whether these results still hold when using between-subjects or mixed design studies. Moreover, these designs would also allow to further investigate the actual learning performance of the three conditions included in the present study. The learning performance of the approaches analyzed could be assessed by including, e.g., a training phase in which trainees are instructed about the skills to learn (what exact things they should look for to improve) and potentially include an external evaluator in-the-loop to provide them with personalized feedback. This feedback could be used also as a measure to directly evaluate the learning performance of the tool.

Finally, the present study did not include a 2D control condition, due to two main reasons. First, it would be hard to obtain an embodied, 1PP on a 2D screen as in VR, since the use of a screen implies that the real body of the participant is always separated from the scene and not inside the seen [40]. Second, the main goal of the present study was not to compare VR to a 2D screen, but rather to compare different interaction modes within VR, i.e. *No Swap*, *Out of Body*, and *Recruiter*. Nonetheless, it is worth observing that when individuals self-assess performance from a traditional 2D video recording, the camera position is normally fixed and viewers can only see the parts of the scene that were recorded, with a limited field-of-view. Contrarily, in setups based on VR technologies, such as the one used in the present study, viewers are able to observe and evaluate any part of the virtual environment. This might result in further advantages related to a higher spatial presence and in more freedom to voluntarily direct attention to desired behaviors or actions happening in the scene [13]. Notwithstanding, caution should be taken to not overload a VR scene with task-unrelated information, since a recent meta-analysis shows that this can negatively impact memory recall [24].

References

[1] A. D. Kaplan, J. Cruit, M. Endsley, S. M. Beers, B. D. Sawyer, and P. A. Hancock, "The effects of virtual reality, augmented reality, and mixed reality as training enhancement methods: A meta-analysis," *Human Factors*, vol. 63, no. 4, pp. 706–726, 2021.

[2] F. Lamberti, F. De Lorenzis, F. G. Pratticò, and M. Migliorini, "An immersive virtual reality platform for training cbrn operators," in *2021 IEEE 45th Annual Computers, Software, and Applications Conference*, 2021, pp. 133–137.

[3] N. Pellas, A. Dengel, and A. Christopoulos, "A scoping review of immersive virtual reality in STEM education," *IEEE Transactions on Learning Technologies*, vol. 13, no. 4, pp. 748–761, 2020.

[4] A. Ferrari, P. Spoletini, M. Bano, and D. Zowghi, "Learning requirements elicitation interviews with role-playing, self-assessment and peer-review," in *2019 IEEE 27th International Requirements Engineering Conference*, 2019, pp. 28–39.

[5] T. L. Dahl, "A preliminary scoping review of immersive virtual soft skills learning and training of employees," in *2021 7th International Conference of the Immersive Learning Research Network*, 2021, pp. 1–5.

[6] M. C. Howard and M. B. Gutworth, "A meta-analysis of virtual reality training programs for social skill development," *Computers & Education*, vol. 144, p. 103707, 2020.

[7] M. North, S. M. North, and J. R. Coble, "Virtual reality therapy: An effective treatment for," *Virtual environments in clinical psychology and neuroscience: Methods and techniques in advanced patient-therapist interaction*, vol. 58, p. 112, 1998.

[8] H. Zhou, Y. Fujimoto, M. Kanbara, and H. Kato, "Virtual reality as a reflection technique for public speaking training," *Applied Sciences*, vol. 11, no. 9, p. 3988, 2021.

[9] J. R. Morgan, M. Price, S. K. Schmertz, S. B. Johnson, A. Masuda, M. Calamaras, and P. L. Anderson, "Cognitive processes as mediators of the relation between mindfulness and change in social anxiety symptoms following cognitive behavioral treatment," *Anxiety, Stress, & Coping*, vol. 27, no. 3, pp. 288–302, 2014.

[10] D. Villani, C. Repetto, P. Cipresso, and G. Riva, "May I experience more presence in doing the same thing in virtual reality than in reality? An answer from a simulated job interview," *Interacting with Computers*, vol. 24, no. 4, pp. 265–272, 2012.

[11] 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*, vol. 71, no. 10, pp. 978–987, 2013.

[12] T. Horigome, S. Kurokawa, K. Sawada, S. Kudo, K. Shiga, M. Mimura, and T. Kishimoto, "Virtual reality exposure therapy for social anxiety disorder: A systematic review and meta-analysis," *Psychological Medicine*, vol. 50, no. 15, pp. 2487–2497, 2020.

[13] Y. Shu, Y.-Z. Huang, S.-H. Chang, and M.-Y. Chen, "Do virtual reality head-mounted displays make a difference? A comparison of presence and self-efficacy between head-mounted displays and desktop computer-facilitated virtual environments," *Virtual Reality*, vol. 23, pp. 437–446, 2019.

[14] I. Stanica, M.-I. Dascalu, C. N. Bodea, and A. D. B. Moldoveanu, "VR job interview simulator: where virtual reality meets artificial intelligence for education," in *2018 Zooming Innovation in Consumer Technologies Conference*. IEEE, 2018, pp. 9–12.

[15] 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*, vol. 15, no. 1, pp. 1–28, 2022.

[16] M. Tailab and N. Marsh, "Use of self-assessment of video recording to raise students' awareness of development of their oral presentation skills," *Higher Education Studies*, vol. 10, no. 1, 2020.

[17] H. L. Andrade, "A critical review of research on student self-assessment," in *Frontiers in Education*, 2019, p. 87.

[18] A. Cannavò, A. Castiello, F. G. Pratticò, T. Mazali, and F. Lamberti, "Immersive movies: the effect of point of view on narrative engagement," *AI & Society*, pp. 1–15, 2023.

[19] M. Hoppe, A. Baumann, P. C. Tamunjoh, T.-K. Machulla, P. W. Woźniak, A. Schmidt, and R. Welsch, "There is no first-or third-person view in virtual reality: Understanding the perspective continuum," in *Conference on Human Factors in Computing Systems*, 2022, pp. 1–13.

[20] P. Salamin, T. Tadi, O. Blanke, F. Vexo, and D. Thalmann, "Quantifying effects of exposure to the third and first-person perspectives in virtual-reality-based training," *IEEE Transactions on Learning Technologies*, vol. 3, no. 3, pp. 272–276, 2010.

[21] J. C. Chan, H. Leung, J. K. Tang, and T. Komura, "A virtual reality dance training system using motion capture technology," *IEEE Transactions on Learning Technologies*, vol. 4, no. 2, pp. 187–195, 2011.

[22] C. J. Falconer, A. Rovira, J. A. King, P. Gilbert, A. Antley, P. Fearon, N. Ralph, M. Slater, and C. R. Brewin, "Embodying self-compassion within virtual reality and its effects on patients with depression," *BJPsych Open*, vol. 2, no. 1, pp. 74–80, 2016.

[23] M. Slater, "Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1535, pp. 3549–3557, 2009.

[24] S. J. G. Ahn, K. L. Nowak, and J. N. Bailenson, "Unintended consequences of spatial presence on learning in virtual reality," *Computers & Education*, vol. 186, p. 104532, 2022.

[25] K. Meyerbröker and P. M. Emmelkamp, "Virtual reality exposure therapy in anxiety disorders: a systematic review of process-and-outcome studies," *Depression and anxiety*, vol. 27, no. 10, pp. 933–944, 2010.

[26] W.-K. Liou and C.-Y. Chang, "Virtual reality classroom applied to science education," in *2018 23rd International Scientific-Professional Conference on Information Technology*, 2018, pp. 1–4.

[27] M. Slater, A. Rovira, R. Southern, D. Swapp, J. J. Zhang, C. Campbell, and M. Levine, "Bystander responses to a violent incident in an immersive virtual environment," *PloS One*, vol. 8, no. 1, p. e52766, 2013.

[28] L. Gamberini, L. Chittaro, A. Spagnolli, and C. Carlesso, "Psychological response to an emergency in virtual reality: Effects of victim ethnicity and emergency type on helping behavior and navigation," *Computers in Human Behavior*, vol. 48, pp. 104–113, 2015.

[29] I. L. Kampmann, P. M. Emmelkamp, D. Hartanto, W.-P. Brinkman, B. J. Zijlstra, and N. Morina, "Exposure to virtual social interactions in the treatment of social anxiety disorder: A randomized controlled trial," *Behaviour Research and Therapy*, vol. 77, pp. 147–156, 2016.

[30] M. Takac, J. Collett, K. J. Blom, R. Conduit, I. Rehm, and A. De Foe, "Public speaking anxiety decreases within repeated virtual reality training sessions," *PloS One*, vol. 14, no. 5, p. e0216288, 2019.

[31] E. Kokkinara, M. Slater, and J. López-Moliner, "The effects of visuomotor calibration to the perceived space and body, through embodiment in immersive virtual reality," *ACM Transactions on Applied Perception*, vol. 13, no. 1, pp. 1–22, 2015.

[32] M. Botvinick and J. Cohen, "Rubber hands 'feel' touch that eyes see," *Nature*, vol. 391, no. 6669, pp. 756–756, 1998.

[33] L. Maister, M. Slater, M. V. Sanchez-Vives, and M. Tsakiris, "Changing bodies changes minds: Owning another body affects social cognition," *Trends in Cognitive Sciences*, vol. 19, no. 1, pp. 6–12, 2015.

[34] D. Mal, E. Wolf, N. Döllinger, C. Wienrich, and M. E. Latoschik, "The impact of avatar and environment congruence on plausibility, embodiment, presence, and the proteus effect in virtual reality," *IEEE Transactions on Visualization and Computer Graphics*, vol. 29, no. 5, pp. 2358–2368, 2023.

[35] D. Banakou, A. Beacco, S. Neyret, M. Blasco-Oliver, S. Seinfeld, and M. Slater, "Virtual body ownership and its consequences for implicit racial bias are dependent on social context," *Royal Society Open Science*, vol. 7, no. 12, p. 201848, 2020.

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

[37] D. Banakou, P. D. Hanumanthu, and M. Slater, "Virtual embodiment of white people in a black virtual body leads to a sustained reduction in their implicit racial bias," *Frontiers in Human Neuroscience*, p. 601, 2016.

[38] K. Kilteni, I. Bergstrom, and M. Slater, "Drumming in immersive virtual reality: The body shapes the way we play," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 4, pp. 597–605, 2013.

[39] D. Banakou, S. Kishore, and M. Slater, "Virtually being einstein results in an improvement in cognitive task performance and a decrease in age bias," *Frontiers in Psychology*, vol. 9, p. 917, 2018.

[40] S. Seinfeld, T. Feuchtner, A. Maselli, and J. Müller, "User representations in human-computer interaction," *Human–Computer Interaction*, vol. 36, no. 5-6, pp. 400–438, 2021.

[41] B. Lenggenhager, T. Tadi, T. Metzinger, and O. Blanke, "Video ergo sum: Manipulating bodily self-consciousness," *Science*, vol. 317, no. 5841, pp. 1096–1099, 2007.

[42] P. Bourdin, I. Barberia, R. Oliva, and M. Slater, "A virtual out-of-body experience reduces fear of death," *PloS One*, vol. 12, no. 1, p. e0169343, 2017.

[43] H. Galvan Debarba, S. Bovet, R. Salomon, O. Blanke, B. Herbelin, and R. Boulic, "Characterizing first and third person viewpoints and their alternation for embodied interaction in virtual reality," *PloS one*, vol. 12, no. 12, p. e0190109, 2017.

[44] S. A. Osimo, R. Pizarro, B. Spanlang, and M. Slater, "Conversations between self and self as sigmund freud — A virtual body ownership paradigm for self counselling," *Scientific Reports*, vol. 5, no. 1, p. 13899, 2015.

[45] M. Slater, S. Neyret, T. Johnston, G. Iruretagoyena, M. Á. d. l. C. Crespo, M. Alabèrnia-Segura, B. Spanlang, and G. Feixas, "An experimental study of a virtual reality counselling paradigm using embodied self-dialogue," *Scientific Reports*, vol. 9, p. 10903, 2019.

[46] Z. Huang, A. Nagata, M. Kanai-Pak, J. Maeda, Y. Kitajima, M. Nakamura, K. Aida, N. Kuwahara, T. Ogata, and J. Ota, "Self-help training system for nursing students to learn patient transfer skills," *IEEE Transactions on Learning Technologies*, vol. 7, no. 4, pp. 319–332, 2014.

[47] E. Kross and I. Grossmann, "Boosting wisdom: Distance from the self enhances wise reasoning, attitudes, and behavior." *Journal of Experimental Psychology: General*, vol. 141, no. 1, p. 43, 2012.

[48] J. P. Powers and K. S. LaBar, "Regulating emotion through distancing: A taxonomy, neurocognitive model, and supporting meta-analysis," *Neuroscience & Biobehavioral Reviews*, vol. 96, pp. 155–173, 2019.

[49] M. D. Nazligul, M. Yilmaz, U. Gulec, M. A. Gozcu, R. V. O'Connor, and P. M. Clarke, "Overcoming public speaking anxiety of software engineers using virtual reality exposure therapy," in *24th European Conference on Systems, Software and Services Process Improvement*. Springer, 2017, pp. 191–202.

[50] F. Palmas, J. Cichor, D. A. Plecher, and G. Klinker, "Acceptance and effectiveness of a virtual reality public speaking training," in *2019 IEEE International Symposium on Mixed and Augmented Reality*, 2019, pp. 363–371.

[51] M. C. C. Tan, S. Y. L. Chye, and K. S. M. Teng, "In the shoes of another": Immersive technology for social and emotional learning," *Education and Information Technologies*, vol. 27, no. 6, pp. 8165–8188, 2022.

[52] M. Akdere, Y. Jiang, and K. Acheson, "To simulate or not to simulate? Comparing the effectiveness of video-based training versus virtual reality-based simulations on interpersonal skills development," *Human Resource Development Quarterly*, 2021.

[53] M. Schmid Mast, E. P. Kleinlogel, B. Tur, and M. Bachmann, "The future of interpersonal skills development: Immersive virtual reality training with virtual humans," *Human Resource Development Quarterly*, vol. 29, no. 2, pp. 125–141, 2018.

[54] D. M. Powell, D. J. Stanley, and K. N. Brown, "Meta-analysis of the relation between interview anxiety and interview performance." *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, vol. 50, no. 4, p. 195, 2018.

[55] S. A. Carless and A. Imber, "The influence of perceived interviewer and job and organizational characteristics on applicant attraction and job choice intentions: The role of applicant anxiety," *International Journal of Selection and Assessment*, vol. 15, no. 4, pp. 359–371, 2007.

[56] P. M. Emmelkamp, K. Meyerbröker, and N. Morina, "Virtual reality therapy in social anxiety disorder," *Current Psychiatry Reports*, vol. 22, pp. 1–9, 2020.

[57] E. Kross and O. Ayduk, "Self-distancing: Theory, research, and current directions," in *Advances in Experimental Social Psychology*, 2017, vol. 55, pp. 81–136.

[58] L. K. Libby and R. P. Eibach, "Visual perspective in mental imagery: A representational tool that functions in judgment, emotion, and self-insight," in *Advances in Experimental Social Psychology*, 2011, vol. 44, pp. 185–245.

[59] "Blender," https://www.blender.org, [accessed: 01-11-2023].

[60] "Reallusion's Character Creator," https://www.reallusion.com/character-creator/, [accessed: 01-11-2023].

[61] "Unity," https://unity.com, [accessed: 01-11-2023].

[62] "SteamVR Unity plugin," https://valvesoftware.github.io/steamvr_unity_plugin/, [accessed: 01-11-2023].

[63] "HTC VIVE Pro kit," https://www.vive.com/us/product/vive-pro-full-kit/, [accessed: 01-11-2023].

[64] "Valve Index Controller," https://www.valvesoftware.com/it/index/controllers, [accessed: 01-11-2023].

[65] "Vive Tracker 2.0," https://www.vive.com/nz/accessory/vive-tracker/, [accessed: 01-11-2023].

[66] "FinalIK," http://root-motion.com/#final-ik, [accessed: 01-11-2023].

[67] "Record and Play," https://recolude.gitlab.io/RecordAndPlay3D/, [accessed: 01-11-2023].

[68] J. Lazar, J. H. Feng, and H. Hochheiser, *Research methods in human-computer interaction*. Morgan Kaufmann, 2017.

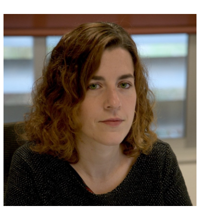
[69] H. Debarba, S. Bovet, R. Salomon, O. Blanke, B. Herbelin, and R. Boulic, "Characterizing first and third person viewpoints and their alternation for embodied interaction in virtual reality," *PLoS One*, vol. 12, no. 12, 2017.

[70] M. Nazligul, M. Yilmaz, U. Gulec, M. Gözcü, R. O'Connor, and P. Clarke, "Overcoming public speaking anxiety of software engineers using virtual reality exposure therapy," in *European Conference on Software Process Improvement*, 2017, p. 191–202.

[71] F. Palmas, J. Cichor, D. Plecher, and G. Klinker, "Acceptance and effectiveness of a virtual reality public speaking training," in *IEEE International Symposium on Mixed and Augmented Reality*, 2019.

[72] F. Faul, E. Erdfelder, A. Buchner, and A.-G. Lang, "Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses," *Behavior Research Methods*, vol. 41, no. 4, pp. 1149–1160, 2009.
[73] M. M. Bradley and P. J. Lang, "Measuring emotion: The self-assessment manikin and the semantic differential," *Journal of Behavior Therapy and Experimental Psychiatry*, vol. 25, no. 1, p. 49–59, 1994.
[74] T. M. Marteau and H. Bekker, "The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety inventory (STAI)," *British journal of clinical Psychology*, vol. 31, no. 3, pp. 301–306, 1992.
[75] T. Peck and M. Gonzalez-Franco, "Avatar embodiment: A standardized questionnaire," *Frontiers in Virtual Reality*, vol. 1, 2021.
[76] C. D. Spielberger, *State-Trait Anxiety Inventory for Adults. Sampler Set.*, 1983.
[77] M. Friedman, "The use of ranks to avoid the assumption of normality implicit in the analysis of variance," *Journal of the american statistical association*, vol. 32, no. 200, pp. 675–701, 1937.
[78] W. J. Conover, *Practical nonparametric statistics*. John Wiley & Sons, 1999, vol. 350.
[79] H. Abdi *et al.*, "Bonferroni and šidák corrections for multiple comparisons," *Encyclopedia of measurement and statistics*, vol. 3, no. 01, p. 2007, 2007.
[80] L. Myers and M. J. Sirois, "Spearman correlation coefficients, differences between," *Encyclopedia of statistical sciences*, vol. 12, 2004.
[81] J. A. Russell, "A circumplex model of affect." *Journal of personality and social psychology*, vol. 39, no. 6, p. 1161, 1980.
[82] S. M. Ritchie, "Self-assessment of video-recorded presentations: Does it improve skills?" *Active Learning in Higher Education*, vol. 17, no. 3, pp. 207–221, 2016.
[83] E. Mitsea, A. Drigas, and P. Mantas, "Soft skills & metacognition as inclusion amplifiers in the 21st century." *International Journal of Online & Biomedical Engineering*, vol. 17, no. 4, 2021.
[84] J. McCarthy and R. Goffin, "Measuring job interview anxiety: Beyond weak knees and sweaty palms," *Personnel Psychology*, vol. 57, no. 3, pp. 607–637, 2004.

Biography



Sofia Seinfeld is an associate professor at the Department of Psychology and Educational Sciences of the Universitat Oberta de Catalunya. Her research interests focus on using embodiment in Virtual Reality to improve psychological treatment and educational practices, as well as to leverage psychological and neuroscientific principles to improve interaction with technology.

Filippo Gabriele Pratticò received his PhD degree in Computer Engineering from Politecnico di Torino, Italy, in 2022, where he is currently serving as an assistant professor at the Department of Control and Computer Engineering. His research interests include eXtended Reality, human-machine interaction, vehicular technology and simulation.

Chiara De Giorgi received her MSc degree in Computer Engineering from Politecnico di Torino, Italy, in 2022. Her major interests regard eXtended Reality and human-machine interaction.

Fabrizio Lamberti is a full professor at the Department of Control and Computer Engineering of Politecnico di Torino, Italy, where he leads the VR@POLITO lab. His interests pertain computer graphics, human-machine interaction and computational intelligence. He serves as an Associate Editor for several journals, including IEEE Transactions on Computers, IEEE Transactions on Consumer Electronics and IEEE Transactions on Learning Technologies. He is a Member of the Board of Governors of IEEE Consumer Technology Society (Member at Large 2021–2023), for which he is also serving as Vice President of Technical Activities. He is a Senior Member of IEEE.