Article

User Experience of Virtual Human and Immersive Virtual Reality Role-Playing in Psychological Testing and Assessment: A Case Study of 'EmpathyVR'

Sunny Thapa Magar [ID], Haejung Suk *[ID] and Teemu H. Laine [ID]

Department of Digital Media, Ajou University, Suwon 16499, Republic of Korea; sunnythapa@ajou.ac.kr (S.T.M.); teemu@ubilife.net (T.H.L.)
*   Correspondence: dbdip@ajou.ac.kr

Abstract: Recent immersive virtual reality (IVR) technologies provide users with an enhanced sense of spatial and social presence by integrating various modern technologies into virtual spaces and virtual humans (VHs). Researchers and practitioners in psychology are attempting to understand the psychological processes underlying human behavior by allowing users to engage in realistic experiences within illusions enabled by IVR technologies. This study examined the user experience of role-playing with VHs in the context of IVR-based psychological testing and assessment (PTA) with a focus on EmpathyVR, an IVR-based empathy-type assessment tool developed in an interdisciplinary project. This study aimed to evaluate the advantages and disadvantages of integrating IVR-based role-playing with VHs into PTA by examining user immersion, embodiment, and satisfaction. A mixed-method approach was used to collect data from 99 Korean adolescents. While the participants reported high levels of immersion and satisfaction, the sense of embodiment varied across the correspondents, suggesting that users may have had disparate experiences in terms of their connection to the virtual body. This study highlights the potential of IVR-based role-playing with VHs to enhance PTA, particularly in empathy-related assessments, while underscoring areas for improvement in user adaptation and VH realism. The results suggest that IVR experiences based on role-playing with VHs may be feasible for PTA to advance user experience and engagement.

Keywords: user experience; virtual human; virtual reality; role-playing; psychological testing and assessment; empathy type

1. Introduction

The digital age is driving significant shifts in the field of psychological therapy [1,2], with psychological testing and assessment (PTA) playing an essential part in these transformations. One of the most significant changes introduced by technology is the use of immersive virtual reality (IVR) in PTA, which allows the user to experience places and situations that they otherwise may not have experienced cognitively or emotionally. Continuous use of IVR technology and tracking devices to collect and transmit vast amounts of behavioral data is anticipated to become a reality in the near future [3]. The interactivity and sensor-based behavioral tracking technology of IVR will have a major impact on PTA, which goes beyond the social phenomenon or industrial trend that technological breakthroughs inevitably increase people's desire to reflect on themselves or others. Many studies have found advantages in investigating social distance and human behavior through social presence using immersive virtual environments [4–6].

Traditional PTA methods rely on self-reporting, which raises concerns about the validity of causal conclusions owing to several factors, including social desirability bias [7], memory limitations, and recall bias [8], lack of self-awareness [9], and cultural and linguistic differences [7,10]. In addition, traditional PTA methods include systematic response bias, method variance, single-method bias, and the psychometric properties (reliability and validity) of questionnaire scales [11]. Simms (2019) demonstrated the poor psychometric accuracy of response scales containing two to five point options [12]. Despite the benefits of technology and concerns regarding traditional methods of conducting PTA, research shows that integrating IVR into mainstream PTA is hindered by diverse challenges such as cost, accessibility, and user adaptation [13,14]. IVR role-playing-based scales, however, are advantageous in their ability to provide real-time experiences and collect instant responses regarding scenario events [15]. Therefore, the potential remains if the issues of cost, accessibility, and adaptability are addressed through industry solutions. The authors and their interdisciplinary research team demonstrated the efficacy of utilizing scales based on user role-playing experiences as the foundation for self-determination in IVR scenarios [15].

The strength of IVR lies not only in its ability to provide spatial illusions for telepresence, but also in the user's ability to interact with virtual humans (VHs). IVR role-playing utilizes VHs unless the virtual avatars are played by real humans. VHs are embodied agents with a human-like appearance [5,16] and provide a representation of humans through computer graphics dummies in a virtual space that can simulate human movements and evaluate workloads [17]. They mimic the human skeleton, exhibit movement following human anthropometry and biomechanics, and their position and posture mirror that of humans. VHs can be used in any software in a virtual environment and must have the ability to interact with and manipulate other virtual objects [18]. Virtual human interaction (VHI) should be natural and intuitive, as people prefer to interact with VHs in the same way as they interact with real people [19]. VHIs are a rapidly developing area of human–computer interaction and have found applications in various fields, such as education, healthcare, entertainment, and customer service, where VHs act as interactive/simulated tutors, patients, and assistants [20,21]. VHIs aim to simulate social interactions, communication, and engagement with entities that VHs may have, their behaviors, and responses. Previous research has shown that VHs can be used for psychological therapy, such as cognitive behavioral therapy, via an imagery focus technique [22]. Moreover, VHs present an effective tool for inducing and assessing human psychological responses [23]. De Melo [24] demonstrated that VHs can accurately assess and respond to human emotions. In addition, studies have explored the opportunities and challenges of using AI and VHs in PTA, showing that VHs can provide more standardized methods of psychological assessment [25]. Studies have examined the impact of human interaction with VHs to serve as virtual avatars on psychological and emotional responses and how these responses can be applied to psychological testing [26].

This study aims to examine the advantages and disadvantages of IVR-based role-playing with VHs for PTA through a user experience evaluation of EmpathyVR, an IVR-based empathy-type testing tool created in an interdisciplinary project. The validity of EmpathyVR for testing empathy types was demonstrated in [15]. In this study, we evaluated its user experience based on prior IVR experience, immersion, satisfaction, and embodiment in IVR role-playing scenarios involving VHs. These results evaluate the use of VHs in IVR-based role-playing for PTAs.

2. Background

An immersive virtual environment (IVE) is a virtual space with seamlessly integrated human–computer interaction and where immersive technology directly stimulates the

user's sensory organs through visual, auditory, tactile, and haptic modalities [27]. IVEs have provided a new direction for psychological research, particularly for understanding social and cognitive behaviors through VH interactions. IVR is increasingly being explored for applications in PTA as it provides a more interactive and controlled experimental environment than traditional self-reporting methods.

Recent studies on IVR in psychological and clinical settings highlight its potential and challenges. Kvapil Varšová and Juřík [28] noted that IVR can be used as exposure therapy to treat acrophobia, but issues like realistic user experiences and biofeedback integration remain. Research has also suggested that therapist-led sessions are more effective than automated ones [29]. Araiza et al. [30] discussed the affordability of IVR and its use in pain distraction and neuropsychological assessments, and concluded that immersion and simulation remain challenging. Riva [31] studied the long-term effects of IVR on anxiety, pain management, and eating disorders. Elor and Kurniawan [32] explored IVR in physical rehabilitation and showed that it enhanced engagement and outcomes with biofeedback. Schöne et al. [33] demonstrated that photorealistic IVR can mimic real-life experiences by showing emotional responses similar to those of real-world experiences. Quah [34] showed that IVR enhances clinical training for health students, although design issues remained. Jespersen [35] introduced the CAVIR test, an IVR tool for assessing cognitive impairment in mood or psychosis disorders, and proved its effectiveness in real-world evaluations.

Some studies have demonstrated the advantages of social presence using VHs and IVR in IVE settings, which enables social presence to be used efficiently to understand human behavior. Studies that report on relational and social behaviors are particularly noteworthy. Felnhofer [36] investigated gender and age differences in physical and social presence in an IVE. Their results revealed that men experienced more physical presence than women, but no gender differences were identified with regard to social presence. Regarding assessing empathy in IVE, Shin [37] found that the cognitive processes through which users experience quality, presence, and flow determine how they empathize with and embody IVR stories. Furthermore, they demonstrate that users' personal traits correlate with their perceived immersion in IVE.

Recent studies have emphasized the role of VHs and IVR technologies in enhancing PTA. Lucas [38] highlighted the use of VH interviewers to reduce stigma in mental health reporting, particularly among military service members with PTSD, and suggested that VHs improve symptom disclosure through anonymity and rapport. Rizzo [39] discussed VH agents such as SimCoach and SimSensei, which enhance client interaction in healthcare by interpreting behavioral cues such as facial expressions and body gestures, thereby improving telehealth diagnostics. Gaggioli [20] explored autonomous VHs in VR therapy, emphasizing their potential to foster effective therapeutic relationships and improve patient interactions.

Previous studies investigated the use and effectiveness of role-playing in IVR in the context of psychology and training. Suárez [40] compared the effectiveness of role-playing in leadership training using VHs and real humans using three training methods: IVR, mixed reality, and conventional methods. The results revealed no significant differences between the conditions based on VHs and real humans, thus indicating the usefulness of role-playing with VHs in skill training. Sapkaroski [41] investigated the effectiveness of role-play-based communication training in a hospital setting using IVR-based role-playing with VHs and traditional role-playing conditions. Their results showed a more significant improvement in communication skills among participants using the IVR-based method compared with those using conventional methods. While these studies demonstrate the effectiveness of IVR-based role-playing with VHs for skill training, research on the user experience of IVR-based PTA solutions based on role-playing with VHs is lacking.

We assembled an interdisciplinary project team comprising psychologists, digital media creators, and computer scientists, who are domain experts and researchers. In this project, a multidimensional empathy scale for adolescents was designed by psychologists based on different empathy theories and instruments such as the Positive and Negative Empathy Scale (PaNES) [42] and the Cognitive, Affective, and Somatic Empathy Scale (CASES) [43]. Six elements along three dimensions of empathy (cognitive−affective, positive−negative, and majority−minority) were used to form eight unique types of empathy. The Multidimensional Empathy Scale for Adolescents (MESA) [44] was adapted for EmpathyVR, a scenario-based IVR role-play content for PTA, followed by a validation study [15]. However, user experiences focusing on VHI and IVR role-playing, which are important aspects of scenario-based PTA, have not been investigated. Therefore, this study examines the utility of IVR and role-playing with VHs in PTA through the user experience in EmpathyVR.

3. Research Question and Methodology

3.1. Research Questions

The mixed-method study aims to answer the following research questions:

- RQ 1: What aspects of IVR role-play and VHs for PTA provide positive and negative experiences for the participants?
- RQ 2: Do the participants' prior IVR experience affect their sense of embodiment during role-play with VHs in EmpathyVR?
- RQ 3. Does the participants' immersion level affect their satisfaction with the IVR experience during role-play with VHs in EmpathyVR?

3.2. Stimulus: EmpathyVR

EmpathyVR is a role-playing IVR experience containing VHs embedded in scenarios that are divided into three parts based on the number of VH friends: one friend, two friends, and several friends. The first and third parts are divided into two episodes containing opposite situations. The second part contains two friends in opposite situations. This trend is repeated for five episodes. This approach differs from the original PTA scale because the participants can be informed about the other scenarios by friends, who are played by the VHs in the IVR. The setting is a virtual high school environment with high-school characters in which the user role-plays as one of the student characters in the scenario, expressing and diagnosing the empathy type. The content starts with a title, warning, tutorial, and introduction using a virtual character. Moreover, the user's virtual avatar is introduced through a virtual mirror so that they can see themselves and check their motor synchrony [45] by waving their hands and making facial expressions. The scenarios in the IVR role-play replace questions from the MESA scale and have been developed by psychological researchers, writers, and content designers. The PTA questions examine what the user feels or thinks about what the VH would feel in the given situation (see Appendix A). After experiencing the episodes, participants rated the questions described earlier in this chapter and explained their experiences. Table 1 summarizes the episodes and Table 2 describes the design details of the virtual environment, VHs, and interaction design in EmpathyVR.

Table 1. IVR scenarios and episodes.

VHs	Process	Story	Scene Image
Avatar	Virtual self [46] check in a mirror (male and female)		 (button text: complete verification)
One VH	Episode1. Positive	A boy has a crush on a girl and she likes him too.	
	Episode5. Negative	A boy was rejected by one of the girls he liked.	
Two VHs	Episode2. Positive and negative	A boy who lost and a girl who won in the badminton tournament.	
Several VHs	Episode4. Positive for the majority vs. negative for the minority	A grumpy student who lost his phone, and a happy group for celebrating someone's birthday.	
	Episode3. Positive for the minority vs. negative for the majority	A group of students lost a soccer game; a boy won a medal in a running competition.	

Table 2. EmpathyVR world and interaction details.

Category	Details
World Design	
Environment	Classroom, school cafeteria
Virtual Humans (VH)	11 virtual humans (VHs) performing 14 roles in the scenario
VH Creation	3D digital humans created through image-based modeling in Character Creator v4.5 (Reallusion Inc., San Jose, CA, USA), using photographs of staff members in their early 20s
Animation	Character animations were recorded with actors using the Perception Neuron 3 motion capture system and AxisStudio 2.8 (Noitom Ltd., Beijing, China). The captured motion data was refined and processed in Maya 2021 (Autodesk Inc., San Francisco, CA, USA) and then applied to the characters in iClone 8 (Reallusion Inc., San Jose, CA, USA), with additional clean-up, polishing, and post-animation work. Missing animations were added, and unnecessary parts were removed to enhance realism and immersion.
Tools	
Character Creator 4.5 [47]	Character creation tool from Reallusion Inc. Version: 4.5 The software generates and customizes realistic 3D characters that are used across multiple 3D modeling platforms and game development engines. Scientific rationale: We tested Unreal Engine's MetaHuman with FaceAR, Photogrammetry, Face Builder for Blender, and Headshot in Character Creation to generate an image-based 3D face model. Headshot in Character Creation was selected owing to its ability to quickly create realistic characters and its compatibility across multiple platforms. It is suitable for efficient integration into diverse 3D environments, as well as being fast, affordable, and less dependent on the skill of the user.
Perception Neuron 3 [48]	Motion capture system from Noitom Ltd. (Beijing, China). Version: 3.0 A full-body motion tracking system with 23 wireless inertial measurement sensors that are mounted on a custom suit, which the actor must equip while recording motions. Scientific rationale: We compared the HTC VIVE Tracker 3.0, Rococo Vision, and Perception Neuron 3 as potential motion capture technologies. Perception Neuron 3 tool was selected because of its portability, compact design, and cost-effectiveness compared to the other motion capture technologies. Moreover, it allows seamless transfer of recorded motion data to digital character animations.
Interaction Design	
Object manipulation	The user interacts with the world and objects using the IVR controllers, which enable basic actions such as grabbing, rotating, pointing, and clicking.
Seated IVR and control	Although EmpathyVR is used in a seated position, the user can move their upper body, hands, and shoulders, and rotate their head in IVR based on their physical movements.
Device and specifications	HTC VIVE Pro Eye (HTC Corporation, Taoyuan, Taiwan) was used for the HMD IVR display and to track participants' head and eye movements. HTC VIVE Trackers (HTC Corporation, Taoyuan, Taiwan) were used to track the participants' upper body and arm movements.
Empathy-type evaluation	After exploring each episode, the user answers questions (based on MESA), and their empathy type result is computed based on their responses.

The overall content workflow is illustrated in Figure 1. The number of episodes matched the number of episodes listed in Table 1. Q1–Q5 refer to the in stimuli questionnaires used to determine the empathy type of the users.
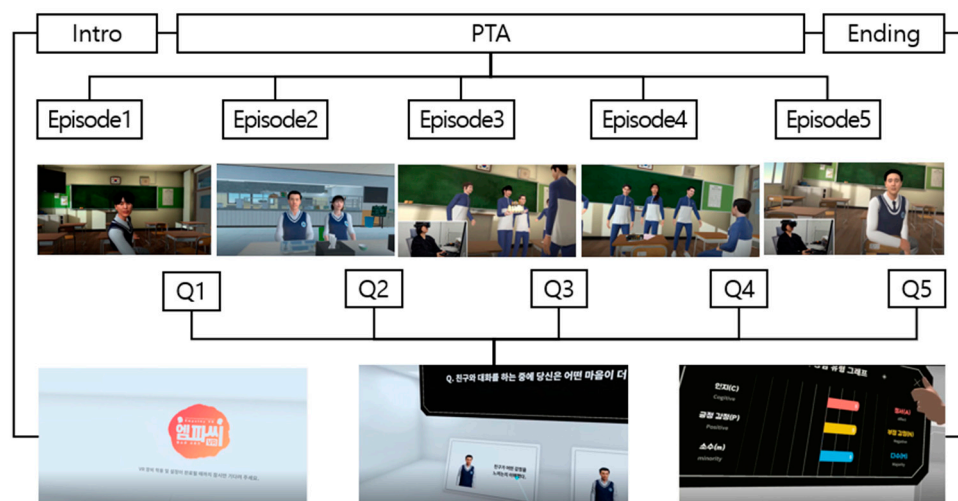
Figure 1. EmpathyVR content flow (bottom left: splash screen showing the EmpathyVR logo in Korean; bottom center: a partial question in Korean about what the character feels when they talk to the friend; bottom right: empathy type test results indicating that the user has tendency toward affective, negative, and majority empathy types).

3.3. Experimental Design

To answer the research questions, we designed a mixed-method experiment comprising pre- and post-test questionnaires to collect data on the participants' prior experience of IVR and their perceptions of the EmpathyVR tool for PTA in terms of embodiment, immersion, and overall satisfaction. The following sections describe the applied data collection instruments, participants, procedures, data analysis methods, and ethical considerations.

3.3.1. Instruments

In the experiment, participants responded to a series of questionnaires that assessed various aspects of the user experience, including prior IVR experience, embodiment, immersion, and satisfaction, as well as open-ended feedback.

The quantitative statements in the pre- and post-test questionnaires were divided into the following sections.

1. Prior experience with IVR (pre-test): The participants were asked to evaluate their prior IVR experience on a four-point Likert scale, with the following responses: never (0), rarely (1), occasionally (2), frequently (3). This was performed to ascertain the extent of the participants' experience with IVR before the experiment.
2. Embodiment (post-test): This section included three questions rated on a seven-point Likert scale to measure participants' sense of embodiment within a virtual environment [49]. The questions focused on the extent to which the participants felt that their virtual body was their own, their ability to control it as if it were real, and their perception of their body's existence in the virtual space.
3. Immersion (post-test): Three questions assessed the level of immersion on a seven-point Likert scale [50]. These questions evaluated the immersion of the visual aspects of the virtual environment, whether the EmpathyVR experience matched real-life experiences, and how easily the participants adapted to the virtual environment.
4. Satisfaction with Empathy Diagnosis Content (post-test): A 10-point question asked participants how likely they were to recommend the IVR content to other middle and high-school students [51] to measure their overall satisfaction with the content. Open-ended questions were used to explore the elements of user satisfaction.

In addition to the quantitative measures, open-ended questions were included to gather qualitative data on the participants' experiences. One positive question asked what they liked about EmpathyVR, and four negative questions explored what they found unfavorable, difficult, or uncomfortable; if they noticed any issues; and any changes they would suggest for EmpathyVR.

3.3.2. Participants

The target users of EmpathyVR were Korean adolescents in secondary education. Both the MESA and IVR role-playing scenarios featured virtual school friends as VHs. The experiment was conducted at a high school in Suwon, Republic of Korea, where 99 adolescents were recruited as test participants. The participants were 15–16 years old, attending their first year of high school, and consisted of 47 females, 49 males, and 2 who identified as other.

3.3.3. Procedure

The participants completed the experiment individually. A researcher greeted the participants upon arrival in the test room and explained the purpose of the experiment. The participants then completed the pretest questionnaire, following which the researcher aided them in putting on the HTC VIVE Pro Eye head-mounted display (HMD). The participant played through the EmpathyVR content and interacted with the VHs. The participants' empathy type was determined by completing a questionnaire after each episode of the scenario. The empathy-type test result was shown to the participant at the end of EmpathyVR. The researcher assisted the participants in removing the HMD and asked them to complete the post-test questionnaire. The entire experiment lasted approximately 35 min.

3.3.4. Data Analysis

We analyzed the quantitative questionnaire data using descriptive statistics (e.g., frequency, mean, standard deviation (variability), median, mode, and range). Furthermore, the chi-square test of independence was used to determine whether the key variables were dependent. The first chi-square test was conducted in the time frame of the questionnaire following the prior IVR experience (four-point frequency) and before the questionnaire on the sense of embodiment (seven-point Likert scale). The second chi-square test was conducted in the time frame between the levels of immersion (seven-point Likert scale) and satisfaction (10-point numeric scale) questionnaire. As chi-square tests of independence were conducted on categorical data, we first converted the numerical embodiment, immersion, and satisfaction data into categorical data by assigning each participant's results to low or high groups. The division threshold for each scale was the mean score calculated from the participants' responses. Chi-square test results were then obtained by calculating the observed and expected frequencies for each combination of categories, calculating the chi-square test statistic based on the frequencies, and checking the chi-square test result against the chi-square distribution table to determine whether the tested variables were independent.

The factors contributing to the corresponding advantages were further examined by analyzing participants' responses to open-ended questions. Additionally, we investigated whether any aspects were identified as disadvantages or potential limitations of IVR, VHs, or role-playing. To analyze the qualitative data from the open-ended questions, we assigned a score to each key term based on the participants' responses. First, keywords were identified, and their connotations were classified as positive or negative. Subsequently, the frequency of the appearance of each keyword within each element was calculated to generate a score.

3.3.5. Ethical Considerations

The experimental plan was reviewed by the Institutional Review Board of Ajou University (202207-HS-001) before data collection. Participants and their guardians provided informed consent prior to the experiment. The collected data were handled anonymously.

4. Result

4.1. Prior Experience with IVR

The pretest questionnaire included questions about the participants' gender, age, and prior experience with IVR. The results indicated that most participants (68%) never or rarely used IVR before the experiment, while those who occasionally or frequently used IVR represented 32% of the participants (Figure 2).
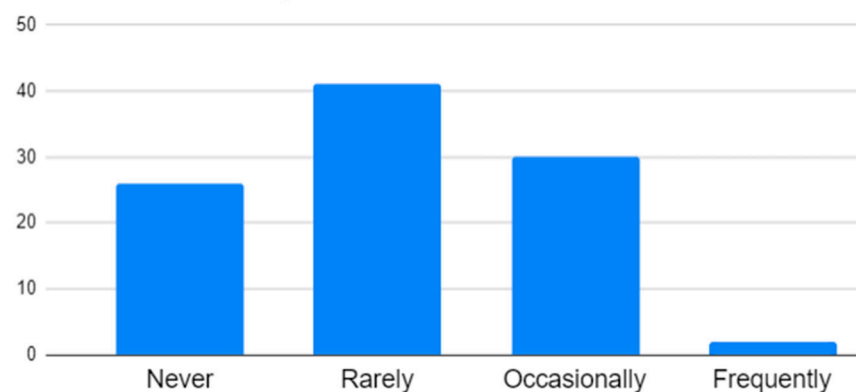
Figure 2. Pre-survey: Q3. How often have you used VR before this evaluation?

4.2. Embodiment and Immersion

Figure 3 shows the responses to the embodiment statement. The scores for embodiment were moderate, with most responses ranging from 4 to 5. The statement "I felt as if the virtual body was my body" (Q1) received the lowest mean ($\mu = 4.09$, $\sigma = 1.45$) among the three statements. The statement "I felt like I could control the virtual body as if it were my own body" (Q2) had the highest mean ($\mu = 4.75$, $\sigma = 1.37$), indicating that the participants felt a relatively strong sense of control over their virtual body. The results also suggest a wider variability in how much the participants felt that the location of their physical body corresponded to the observed location of the virtual body (Q3) ($\mu = 4.11$, $\sigma = 1.64$). Moreover, the results for Q4 showed a mixed experience among the users, as many participants provided higher scores (5 or 6), while some provided lower ratings.
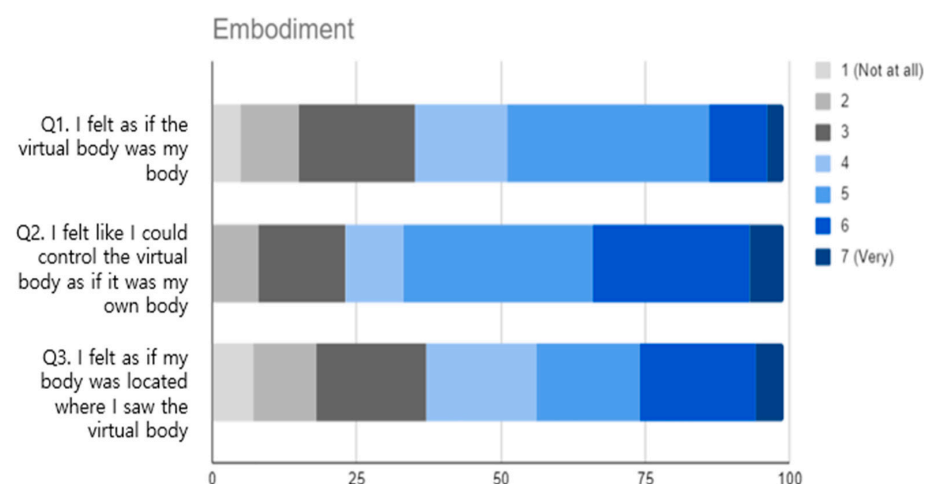
Figure 3. Results of the embodiment statements in the post-test questionnaire.

Regarding the immersion statements (Figure 4), the participants rated the ease of adaptation to the virtual environment (Q6) the highest, with a mean of 5.76 and the lowest variability (σ = 1.11). This result suggests that the participants generally found it easy to adjust to the virtual environment. The visual aspects of the virtual environment (Q4) were also rated highly (μ = 4.71, σ = 1.24), although slightly lower than the ease of adaptation (Q6). The immersion experience appeared to be consistent with the real-world experience (Q5), but this statement received a slightly lower mean (μ = 4.59, σ = 1.43) than the other statements.
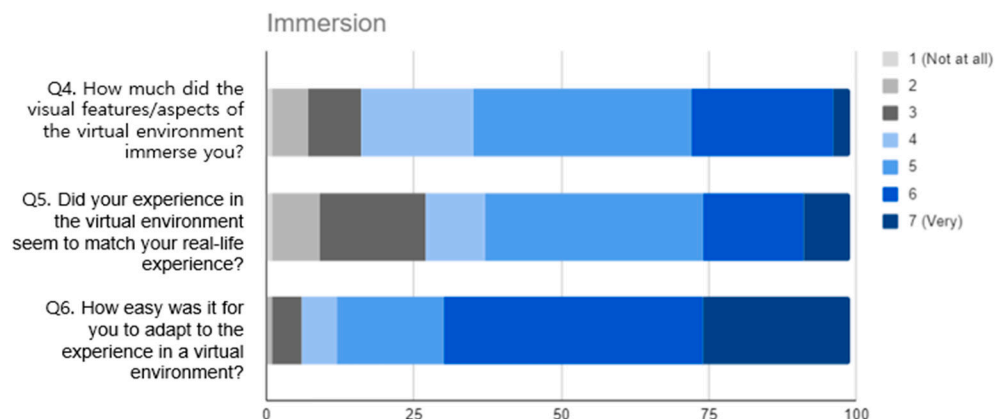

Figure 4. Results of the immersion statements in the post-test questionnaire.

The participants reported a sense of embodiment, although their connection to the virtual body was moderate. In Q2, the participants demonstrated the highest sense of control over the virtual body. Participants rated their immersive experiences highly, particularly their ability to adjust to the virtual environment, as evidenced by their responses to Q6. The visual aspects also significantly contributed to their sense of immersion. The responses to statements about embodiment exhibited more variability than the statements about immersion, particularly to Q3, which suggests that the experience of embodiment may be a more subjective or context-dependent phenomenon for the participants.

4.3. Satisfaction with EmpathyVR

The question "How likely do you recommend this VR content to middle and high school students?" measured the overall satisfaction of participants with EmpathyVR in its use as IVR role-playing with VHI. The results presented in Table 3 indicate a mean score of 8.31 out of 10 (σ = 1.24), thus the majority of the participants found EmpathyVR to be moderately to highly recommendable. This result suggests that EmpathyVR aligns well with the needs and expectations of the target audience.

Table 3. Results of the user satisfaction statement ("How much do you recommend this VR content to middle school and high school students?").

	Mean	Median	Mode	Standard Deviation	Min	Max	Range	25th Percentile	50th Percentile	75th Percentile
Value	8.31	8.00	8.00	1.24	4.00	10.00	7.00	7.50	8.00	9.00

We analyzed the participants' responses to open-ended questions to gain further insight into their satisfaction and dissatisfaction with EmpathyVR, specifically focusing on VHI through IVR role-playing. As the qualitative data do not include specific numerical satisfaction scores but rather the frequencies of the participants who expressed likes or dislikes in various categories, we focus on analyzing the distribution and implications of these feedback categories (see Appendix B).

Participants' satisfaction levels were determined based on the categories identified from positive feedback, whereas their dissatisfaction levels were extracted from negative feedback. The feedback covered various aspects including immersion, realism of virtual humans and their environment, usability of hardware and software, interest based on enjoyment and educational value, and opportunities for quality content to provide empathy-type detection.

Relevant comments were categorized based on their meaning and appropriate keywords were extracted. The final categories and their explanations are as follows.

- Immersive: Phrases indicating a feeling of presence.
- Realistic Virtual Humans and Environment: Phrases describing an accurate representation similar to the real world.
- Usability—Hardware: Phrases signifying that the devices were useful.
- Usability—Software: Phrases indicating that the content fulfilled its intended purpose.
- Interesting and Enjoyable: Phrases suggesting that the content was entertaining.
- Interesting and Educational: Phrases referring to engaging and valuable content for empathy detection and learning.
- Opportunity Providing: Phrases indicating that the content facilitated knowledge growth and served as a platform for development and empathy detection.

To ascertain the prevalence of positive and negative experiences, one point was allocated for each positive and negative connotation of the keyword, respectively. This approach enabled the systematic quantification of qualitative feedback from the open-ended questions.

Figure 5 illustrates the frequencies of the aspects in EmpathyVR on which the participants expressed positive and negative feedback on the identified categories. Figure 6 provides additional detailed information on the positive and negative responses related to VHs and the environment.
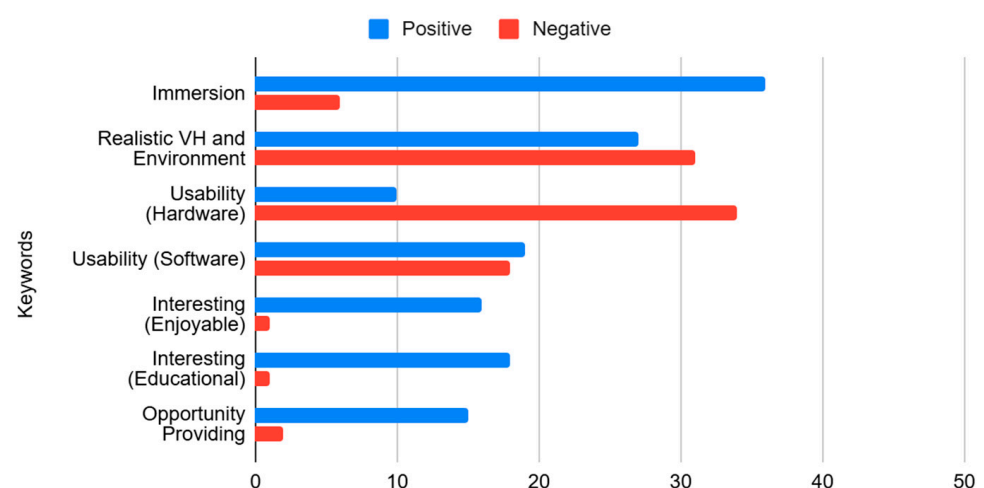

Figure 5. Categorized scoring of participants' responses to the post-test questions: "What did you like about the VR content?" and "What was bad about the VR content?".

Notably, participants highlighted the realistic environment in positive responses, while they expressed reservations about the realism of VHs in negative responses.
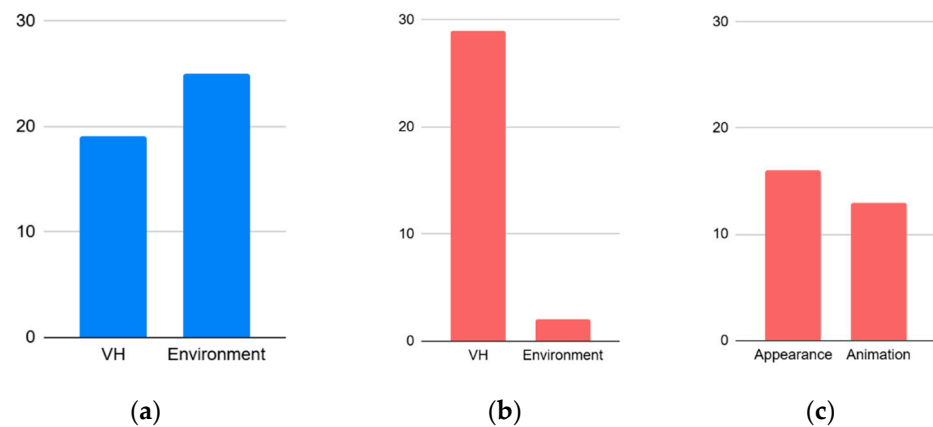
Figure 6. Results of responses: (**a**) Positive responses of participants for VH and environment post-test questions: "What did you like about the VR content?"; (**b**) Negative responses of participants for VH and environment "What was bad about the VR content?"; (**c**) Negative responses of participants on VH regarding appearance and animation.

4.4. Relationship Between Familiarity with IVR and Embodiment

For further investigation, a chi-square test was conducted to examine the following hypothesis for answering Research Question 2:

The participants' prior IVR experience is independent of their embodiment level during role-play with VHs in EmpathyVR.

To examine the relationship between prior experience with IVR and the level of embodiment during EmpathyVR, each element of the frequency of experience was scored on a scale of 1 to 4, with higher numbers indicating more frequent use and relative familiarity with IVR. We assigned scores of 1 and 2 to the low-experience group, and scores of 3 and 4 to the high-experience group. Similarly, we assigned the embodiment results (three statements on a seven-point Likert scale) to the low- and high-embodiment groups by comparing the sum of the statement scores to the mean of all participants' embodiment scores. Thirty-two participants reported having occasional IVR, while 67 participants reported having rarely experienced IVR. Of the participants who occasionally-frequently used IVR, 19 reported a high level of embodiment and 13 reported a low level of embodiment. Conversely, of the participants who rarely or never used IVR, 44 reported a high level of embodiment and 23 reported a low level of embodiment. Table 4 lists these results along with the results of the chi-square test. Because the *p*-value was 0.700, we failed to reject the null hypothesis. This result indicates that there is no statistically significant relationship between the participants' prior IVR experience and their embodiment levels; thus, these variables are independent.

Table 4. Results of the chi-square test of independence between familiarity with IVR level and satisfaction level (significance level = 0.05).

Experience/Embodiment	High Embodiment	Low Embodiment	Total
High Experience	19	13	32
Low Experience	44	23	67
Total	63	36	99

Chi Square Test Result	Chi Square Statistic	*p*-Value	Degrees of freedom	Expected frequencies
	0.149	0.700	1.0	[[20.36, 11.64], [42.64, 24.36]]

4.5. Relationship Between Immersion and Satisfaction

Another chi-square test was conducted with the following hypothesis to answer Research Question 3:

The participants' immersion and satisfaction levels are independent of each other during role-play with VHs in EmpathyVR.

To examine the relationship between immersion and satisfaction, we first converted the numerical data on immersion and satisfaction into categorical data by comparing each participant's score to the mean of all the participants' scores. Scores below or above the mean were assigned to the low and high groups, respectively. We then performed a chi-square test on the two variables to determine whether the immersion and satisfaction levels were independent (Table 5). The chi-square test of independence results in Table 5 suggest that we can reject the null hypothesis because the *p*-value (0.0036) is less than 0.05, which is indicative of a statistically significant relationship between immersion and satisfaction.

Table 5. Results of the chi-square test of independence between the immersion level and satisfaction level (significance level = 0.05).

Immersion/Satisfaction	High Satisfaction	Low Satisfaction	Total
High Immersion	50	6	56
Low Immersion	24	19	43
Total	74	25	99

Chi Square Test Result	Chi Square Statistic	*p*-Value	Degrees of freedom	Expected frequencies
	12.72	0.000362	1.0	[[41.86, 14.14], [32.14, 10.86]]

5. Discussion

Previous research has demonstrated the effectiveness of role-playing with VHs for skill training; however, a significant gap remains in understanding the user experience of IVR-based PTA solutions that incorporate role-playing with VHs [41,52]. To address this gap, an interdisciplinary team of psychologists, digital media creators, and computer scientists collectively examined the effectiveness of IVR-based role-playing with VHs in PTA through an analysis of user experience in a scenario-based EmpathyVR IVR experience. Quantitative data were analyzed using descriptive statistics and chi-square tests to examine the relationships among IVR immersion, VH experience, prior VR experience, embodiment, and content satisfaction. Open-ended responses were evaluated by identifying key terms, classifying them as positive or negative, and calculating their frequencies to assess the advantages, limitations, and user perceptions of IVR, VHs, and role-playing in PTA. The following sections present our interpretations of the results along with the implications and limitations of the study.

5.1. Advantages and Disadvantages of Role-Playing with VHs in IVR-Based PTA

Participants demonstrated a mixed sense of embodiment. While users reported a relatively strong ability to control the virtual body (Q2) with a mean score of 4.75, their overall connection to the virtual body was moderate, as evidenced by the lower scores in other embodiment-related statements. The variability in the embodiment scores, especially in Q3, with a mean of 4.11 and a standard deviation of 1.65, suggests that the feeling of embodiment is subjective. Some participants felt more embodied than others, perhaps owing to differences in their prior IVR experiences, personal preferences, or engagement with the scenario.

In contrast, the participants reported higher levels of immersion more consistently, particularly in terms of how quickly they adapted to the IVR environment. The mean score (5.76) and low variability ($\sigma = 1.11$) for the adjustment statement highlight that the participants generally found it easy to immerse themselves in the IVE. This response aligns with the high ratings for the visual aspects of the environment (Q4) ($\mu = 4.71$, $\sigma = 1.25$). Notably, the participants rated their immersion as consistent with real-world experiences (Q5), although this rating was slightly lower ($\mu = 4.59$, $\sigma = 1.44$) than the adaptation and their visual aspects of the virtual environment. This trend indicates that, while the virtual environment was immersive, some participants may still have felt certain limitations in how closely it mimicked real-world scenarios.

Furthermore, Figures 5 and 6 illustrate that the lack of realism of VHs can present a negative experience in IVR role-playing. This can be attributed to quality issues in the VH animations and the lack of realism in the real-time rendering of the VHs caused by the limitations of IVR hardware, where the uncanny valley phenomenon is evident. However, this did not significantly impact the validity of PTA or satisfaction with the content (Table 3 and Figure 5).

In conclusion, role-playing with VHs could provide advantages for PTA, and the nature of immersive situations and interesting stories might introduce great potential for future PTA. Empathy can appeal to a user's social desirability and allow the user to find its usefulness. However, the embodiment of the user's virtual body may be subjective, depending on the user's perspective, thus suggesting the need for a personalized embodiment experience.

Furthermore, this study showed that the lack of realism in VHs and the quality of the VHI could diminish users' interest, as indicated by the high scores shown in Figures 5 and 6. Despite the negative evaluation of the VHs and hardware-related complications (Figure 5), no significant disadvantages were reported regarding overall satisfaction.

5.2. Prior IVR Experience and Embodiment

The chi-square test results between prior IVR experience and embodiment did not reveal a significant dependency between the variables. This finding suggests that no statistically significant correlation exists between prior experience with IVR and the embodiment in a dataset. This result suggests that familiarity with IVR technology does not affect the users' sense of connection with their virtual bodies. It can be reasonably assumed that users' prior experience aids them in navigating and interacting within a virtual environment. Nevertheless, the findings indicate that even participants with limited experience demonstrated moderate to high levels of embodiment, especially regarding control over the virtual body, thereby suggesting that the design and content of EmpathyVR effectively evoked a sense of embodiment, even among those with minimal experience with IVR.

5.3. Level of Immersion and Satisfaction

The high mean satisfaction score of 8.31 out of 10 and the results of the qualitative data analysis indicated that most participants found IVR-based role-playing with VHs to be immersive, interesting, enjoyable, and educational. Furthermore, the chi-square test results for the levels of immersion and satisfaction revealed a significant dependency between the variables. Participants who exhibited higher immersion scores consistently reported higher satisfaction with the IVR role-playing content, whereas those exhibiting lower immersion scores correlated with lower satisfaction. This trend indicates that the immersive quality of role-playing with VHs in the IVR environment significantly shaped the overall satisfaction of the user.

5.4. Implications

This study demonstrated that IVR-based role-playing with VHs can provide a positive user experience in PTA. Using scenarios that reflect real-world social situations through the VHI likely contributed to the high levels of immersion and satisfaction reported by the participants. This positive response suggests that combining IVR, role-playing, and VHs could offer a new frontier in PTA by providing more dynamic and interactive methods for assessing psychological states. We expect that the quality and realism of VHs can be enhanced alongside the development of new IVR technology to enable higher resolutions and frame rates, which are essential for producing realistic IVEs rendered in real time. Furthermore, recent advances in artificial intelligence, especially in terms of large language models capable of reasoning and mimicking emotional intelligence, have enabled the development of VHs capable of personalizing interactive PTA experiences to match user expectations, preferences, and needs.

In the user study, we used a questionnaire to measure the sense of embodiment. However, Guy et al. [53] pointed out in their literature review on embodiment and its assessment methods that questionnaires may not be optimal for measuring embodiment. Instead, researchers should utilize physiological sensors, particularly electroencephalography, which enables real-time measurement of embodiment in virtual environments [53]. Moreover, electroencephalography can be used to implement brain–computer interfaces [54], thus opening avenues for innovation in human–computer interactions and VHI.

5.5. Limitations

This study is a derivative of the effectiveness study of EmphathyVR and was not designed to fully test the feasibility of IVR and role-playing with VHs in PTAs. Accordingly, a somewhat redundant evaluation methodology was employed to assess the user experience of the content. The data were processed by converting the scores to the contextual positivity of the keywords that correspond to IVR and role-playing with VHs in the subjective satisfaction questionnaire. Multiple researchers crosschecked the scores to reduce the likelihood of errors during score conversion. However, it must be acknowledged that this method is not optimal. Additionally, given that EmpathyVR is designed to assess empathy, it is plausible that role-playing with VHs was more effectively utilized and rated higher in terms of satisfaction because it prompts users to consider the feelings of others. To address this potential bias, it would be beneficial to expose users to situations, such as the Myers–Briggs-Type Indicator (MBTI), to explore users' perceptions and attitudes toward PTAs in a similar manner.

6. Conclusions

The results of this study shed light on the potential of IVR and VHI in PTA. The participants showed high levels of immersion, particularly in adapting to the virtual environment, whereas their experiences with embodiment exhibited more variation. Satisfaction was strongly linked to immersion, and participants with prior IVR experience overall reported higher levels of embodiment. These findings suggest a significant application of IVR-based PTA, including role-playing with VHs. However, further work is required to optimize the sense of embodiment to fully realize the potential of these immersive environments.

Future work should address the suboptimal appearance and animations of VHs in EmpathyVR (Figures 5 and 6), which were rated highly, but may not be sufficient to overcome the uncanny valley effect [55]. A realistic alternative is to balance quality and optimization, such as making VHs completely realistic or using more stylized characters. Although the shortcomings of VHs pose a potential threat to immersion, this study and

previous research have shown that these issues do not significantly affect the immersion levels, usability, or effectiveness of PTA [15].

The answer to RQ2 reveals the following insights into the future of IVR-based PTA. The growth in IVR users will provide a greater scope for observing changes in user psychology and behavior in an immersive environment conducive to PTA. In addition, it would be beneficial to determine whether the accuracy of diagnosis can be improved and predicted over time through objective measurement techniques, such as physiological sensors, and whether in-depth analysis beyond diagnosis by scale is possible. An important concern regarding the implementation of a PTA through IVR-based role-playing with VHs is the monitoring of user behavior using sensors.

Abbreviations

The following abbreviations are used in this study:

IVR	Immersive virtual reality
IVE	Immersive virtual environment
VH	Virtual Human
PTA	Psychological testing and assessment
VHI	Virtual Human Interaction
PaNES	Positive and Negative Empathy Scale
CASES	Cognitive, Affective, and Somatic Empathy Scale
MESA	Multidimensional Empathy Scale for Adolescents
HMD	Head-mounted display
MBTI	Myers-Briggs type indicator
MSIT	Ministry of Science and ICT
ITRC	Information Technology Research Center
IITP	Institute for Information & Communications Technology Planning & Evaluation

Appendix A

Table A1 presents the questions used in the scenarios presented in EmpathyVR to assess the user's empathy type.

Table A1. Questions for the empathy-type test in EmpathyVR.

Scenario	Questions	Possible Answers	Scoring
#1	Positive Situation		
	1. Which one of the following did you experience more while talking with Sangmin?	(a) I easily understood what Sangmin was feeling	(a) = 1 point for C (b) = 1 point for A
	2. From the list of provided emotion words, choose the emotion you felt after listening to Sangmin's story? (up to five words)	Delighted, upset, happy, depressed, sad, excited, worried, proud, excited, impressed, blessed, uncertain, embarrassed, interested, pitiful, admired	Five relevant words = 1 point for A
	3. How much do you agree with "I was excited and happy as if I confessed and was accepted"?	5-rating scale (1 = I strongly disagree, 5 = I strongly agree)	4&5 = 1 point for A
	4. How much do you agree with "I tried to imagine what Sangmin was feeling as much as possible when I was hearing his story"?	5-rating scale (1 = I strongly disagree, 5 = I strongly agree)	4&5 = 1 point for C

Scenario	Questions	Possible answers	Scoring
	Negative situation		
#1	5. Which one of the following did you experience more while talking with Sihwan?	(a) I easily understood what Sihwan was feeling (b) I felt what Sihwan was feeling	(a) = 1 point for C (b) = 1 point for A
	6. From the list of emotion words, choose the emotion you felt after listening to Sihwan's story? (up to five words)	Delighted, upset, happy, depressed, sad, excited, worried, proud, excited, impressed, blessed, uncertain, embarrassed, interested, pitiful, admired	Five relevant words = 1 point for A
	7. How much do you agree with "I felt down and sad as if I confessed and was rejected"?	5-rating scale (1 = I strongly disagree, 5 = I strongly agree)	4&5 = 1 point for A
	8 How much do you agree with "I tried to imagine what Sihwan was feeling as much as possible when I was hearing his story"?	5-rating scale (1 = I strongly disagree, 5 = I strongly agree)	4&5 = 1 point for C
	9. There are gift icons that you can give to your friends. Choose the person you want to give the gift icon to: Sangmin or Sihwan.	Sangmin or Sihwan	Sangmin = 1 point for P Sihwan = 1 point for N
	10. When you heard the story of Sangmin who confessed successfully and a friend Sihwan whose confession was rejected, whose story did you want to hear more?	Sangmin or Sihwan	Sangmin = 1 point for P Sihwan = 1 point for N
	11. Whose feelings influenced you more—Sangmin whose confession got accepted or Sihwan whose confession got rejected?	Sangmin or Sihwan	Sangmin = 1 point for P Sihwan = 1 point for N
	12. When there are many people who are happy like Sangmin who made a successful confession, and one person who is having a hard time like Sihwan whose confession was rejected; whose story do you want to hear more, from (Sangmin) group or (Sihwan)? [Show images of several Sangmins and one Sihwan.]	Sangmin or Sihwan	Sangmin = 0.5 point for M Sihwan = 0.5 point for m
	13. When there is one person who is happy like Sangmin who confessed successfully, and there are many people who are struggling like Sihwan whose confession was rejected, which one do you want to hear more from (Sihwan) group or (Sangmin)? [Show images of several Sihwans and one Sangmin]	Sangmin or Sihwan	Sangmin = 0.5 point for m Sihwan = 0.5 point for M

Table A1. *Cont.*

Scenario	Questions	Possible answers	Scoring
	Positive situation		
	1. Which one of the following did you experience more while talking with Jaehong?	(a) I easily understood what Jaehong was feeling (b) I felt what Jaehong was feeling	(a) = 1 point for C (b) = 1 point for A
	2. From the list of provided emotion words, choose the emotion you felt after listening to Jaehon's story? (up to five words)	Delighted, upset, happy, depressed, sad, excited, worried, proud, excited, impressed, blessed, uncertain, embarrassed, interested, pitiful, admired	Five relevant words = 1 point for A
	3. How much do you agree with "I was excited and happy as if I won first place in the competition"?	5-rating scale (1 = I strongly disagree, 5 = I strongly agree)	4&5 = 1 point for A
#2	Negative situation		
	5. Which one of the following did you experience more while talking with Juyeon?	(a) I easily understood what B was feeling (b) I felt what B was feeling	(a) = 1 point for C (b) = 1 point for A
	6. From the list of emotion words, choose the emotion you felt after listening to Juyeon's story? (up to five words)	Delighted, upset, happy, depressed, sad, excited, worried, proud, excited, impressed, blessed, uncertain, embarrassed, interested, pitiful, admired	Five relevant words = 1 point for A
	7. How much do you agree with "I felt down and sad as if I lost in the first round"?	5-rating scale (1 = I strongly disagree, 5 = I strongly agree)	4&5 = 1 point for A
	8 How much do you agree with "I tried to imagine what Juyeon was feeling as much as possible when I was hearing his story"?	5-rating scale (1 = I strongly disagree, 5 = I strongly agree)	4&5 = 1 point for C
	9. There is a glass of water that you can give to your friends. Choose the person you want to give a glass of water to: Jaehong or Juyeon	Jaehong or Juyeon	Jaehong = 1 point for P Juyeon = 1 point for N
	10. When you heard the story of Jaehong who won first place in the competition and Juyeon who lost in the first round, whose story did you want to hear more?	Jaehong or Juyeon	Jaehong = 1 point for P Juyeon = 1 point for N
	11. Among the stories of Jaehong who won first place in the competition or a friend Juyeon who lost in the first round, whose feelings influenced you more?	Jaehong or Juyeon	Jaehong = 1 point for P Juyeon = 1 point for N
	12. When there are many people who are happy like Jaehong who won first place in the competition and one person who is having a hard time like Juyeon who lost in the first round, whose story do you want to hear more from (Jaehong) group or (Juyeon)? [Show images of several Jaehongs and one Juyeon]	Jaehong or Juyeon	Jaehong = 0.5 point for M Juyeon = 0.5 point for m
	13. When there is one person who is happy like Jaehong who won the first place in the competition, and there are many people who are struggling like Juyeon who lost in the first round, which one do you want to hear more from (Juyeon) group or (Jaehong)? [Show images of several Juyeons and one Jaehong]	Jaehong or Juyeon	Jaehong = 0.5 point for m Juyeon = 0.5 point for M

Scenario	Questions	Possible answers	Scoring
	Negative situation for minority and positive situation for majority		
#3	1. Which one of the following did you experience more?	(a) I can better understand the positive feelings of celebrating the birthday (majority) than the negative feeling of losing a cell phone (Eunwoo) (b) I can better understand the negative feeling of losing a cell phone (Eunwoo) than the positive feelings of celebrating the birthday (majority)	(a) = 1 point for M (b) = 1 point for m

Table A1. *Cont.*

Scenario	Questions	Possible answers	Scoring
	2. Which one of the following did you want to do more?	(a) I want to help Eunwoo to feel better (b) I want to celebrate the birthday with them	(a) = 1 point for m (b) = 1 point for M
	3. Which one of the following did you feel more?	(a) I feel bad for Eunwoo (b) I feel good for students who are celebrating the birthday	(a) = 1 point for m (b) = 1 point for M
	4. Which one of the following did you want to do more?	(a) I want to do what Eunwoo wants to do with me (b) I want to do what students who are celebrating the birthday want to do with me	(a) = 1 point for m (b) = 1 point for M
	Positive situation for minority and negative situation for majority		
	5. Which one of the following did you experience more?	(a) I can better understand the negative feelings of losing in a soccer competition (majority) than the positive feeling of winning in a running competition (Woohyun) (b) I can better understand the positive feeling of winning in a running competition (Woohyun) than the negative feelings of losing in a soccer competition (majority)	5. Which one of the following did you experience more?
	6. Which one of the following did you want to do more?	(a) I want to congratulate Woohyun (b) I want to help students who lost in a soccer competition	(a) = 1 point for M (b) = 1 point for m
	7. Which one of the following did you feel more?	(a) I feel good for Woohyun (b) I feel bad for the students who lost in a soccer competition	(a) = 1 point for M (b) = 1 point for m
	8. Which one of the following did you want to do more?	(c) I want to do what students who lost in a soccer competition want to do with me (d) I want to do what Woohyun wants to do	(a) = 1 point for M (b) = 1 point for m

C = cognitive empathy; A = affective empathy; P = positive empathy; N = negative empathy; M = majority empathy; m = minority empathy

Appendix B

Table A2 lists the keywords identified for each category that were used to classify and score the responses to the open-ended questions.

Table A2. Keywords for scoring open-ended questions.

Category	Keywords/Phrases
Immersion	"Immersive", "I felt like real situation", "Sense of immersion quite high", "Easy to immerse in the situation that could have existed in reality", "Well immersed", "Allowing people to immerse themselves much more quickly", "Good to think as another world and immerse", "Immersed", " Immersing in situation", "Like facing real situation", "Able to experience actual situation in virtual reality", "People appeared seem to be real and easily immersed in them", "Immersive due to real human voice", "See and feel it with own eyes", "Making it immersed", "Allows as if I am experiencing the situation myself"
Realistic VH and Environment	"Experiencing real school", "Made based on things that can be experienced in reality", "Quite real", "Students similar to reality", "Like reality", "Felt like real", Characters felt alive", "Virtual space felt real", "Act in real situation", "Experiencing real school", "Things that could happen in real situation", "Felt like a real situation", "Expressed realistic and well", "Felt realistic", "Made based on things that can be experienced in reality", "More realistic", "Liked compositions of situation that could actually occur", "Realistic interaction", "The situation seemed real", "Like facing real situation", "Characters", "Quality", "Alive characters", "Friends", "People", "Human", "Students", "Character movements", "Facial expressions", "Appearances", "People appeared seemed real", "Disappointing"

Table A2. *Cont.*

Category	Keywords/Phrases
Opportunity Providing	"Could easily find out my empathy type", "Beneficial to learn empathy through simulation", "Able to experience technology that never encountered before", "Able to know empath type", "Able to know emotional type", "Able to pick objects felt like reality", "Various situation and options gave opportunity to think about how to act in real situation", "Was good to do it in a situation that could happen around people", "Was able to answer with more immersion than by understanding the situation in writing", "Variety of options of choosing emotions", "Able to conduct test", "Decide options", "Opportunity providing", "Being able to choose multiple emotions", "Empathy Increases", "Being able to sympathize and understand emotions"
Interesting	"Fun", "Nice to know empathy type", "Amazing to enter inside VR", "Was great being able to make decisions across multiple situations", "Liked compositions of situation that could actually occur", "Being able to know empathy was fun" "New and fun", "More realistic and interesting than regular one", "Less boring", "Nice and fun content", "Enjoy different type of test"
Usability	"Useful", "Good and easy to use", "Able to get used to it quickly", "Easy interface", "Able to contrate well and move freely", "Easy operation", "Simple and easy operation", "Heavy device", "Device", "VR device", "Felt dizzy", "Limited hand motion", "Headset", "Poor screen quality", "VR machine"
Role-Play	"Experiencing the actual situation", "As a student myself", "Able to sympathize more", "Simulation", "Asking opinion after the situation presented made more immersed", "Experience the situation and judge", "Feel as if experiencing the situation myself", "Faced a real situation rather than reading"
Negative	"Frustrating", "Awkward", "Uncomfortable", "Unnatural", "Scared", "Bad", "Inconvenient", "Not realistic", "Difficult", "Bad quality", "Heavy", "Poor", "Low", "Disappointing", "Hurt", "Ugly"
Positive	"Pretty", "Good", "Realistic", "Good quality", "Real", "Interesting", "Fun", "Like", "High sense", "Nice", "Like reality", "Amazing", "Easy", "Alive", "Able to enjoy", "Seemed real", "Less boring", "Familiar", "Refreshing"

References

1. Armaou, M.; Konstantinidis, S.; Blake, H. The Effectiveness of Digital Interventions for Psychological Well-Being in the Workplace: A Systematic Review Protocol. *Int. J. Environ. Res. Public Health* **2019**, *17*, 255. [CrossRef]

2. Fu, Z.; Burger, H.; Arjadi, R.; Bockting, C.L.H. Effectiveness of digital psychological interventions for mental health problems in low-income and middle-income countries: A systematic review and meta-analysis. *Lancet Psychiatry* **2020**, *7*, 851–864. [CrossRef]

3. Serrano, B.; Botella, C.; Wiederhold, B.K.; Baños, R.M. Virtual Reality and Anxiety Disorders Treatment: Evolution and Future Perspectives. In *Virtual Reality for Psychological and Neurocognitive Interventions*; Rizzo "Skip", A., Bouchard, S., Eds.; Virtual Reality Technologies for Health and Clinical Applications; Springer: New York, NY, USA, 2019; pp. 47–84, ISBN 978-1-4939-9480-9.

4. Guimarães, M.; Prada, R.; Santos, P.A.; Dias, J.; Jhala, A.; Mascarenhas, S. The Impact of Virtual Reality in the Social Presence of a Virtual Agent. In Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents, Virtual Event, 20–22 October 2020; ACM: New York, NY, USA, 2020; pp. 1–8.

5. Hoppe, M.; Rossmy, B.; Neumann, D.P.; Streuber, S.; Schmidt, A.; Machulla, T.-K. A Human Touch: Social Touch Increases the Perceived Human-likeness of Agents in Virtual Reality. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; pp. 1–11.

6. Rubin, R.B.; McHugh, M.P. Development of parasocial interaction relationships. *J. Broadcast. Electron. Media* **1987**, *31*, 279–292. [CrossRef]

7. West, M.R. The Limitations of Self-Report Measures of Non-Cognitive Skills. Available online: https://www.brookings.edu/articles/the-limitations-of-self-report-measures-of-non-cognitive-skills/ (accessed on 22 April 2025).

8. NeuroLaunch Self-Report Measures in Psychology: Advantages, Limitations, and Applications. Available online: https://neurolaunch.com/self-report-measures-psychology/ (accessed on 22 April 2025).

9. Tourangeau, R.; Rips, L.J.; Rasinski, K. *The Psychology of Survey Response*, 1st ed.; Cambridge University Press: Cambridge, UK, 2000; ISBN 978-0-521-57246-0.

10. Garcia, J.; Gustavson, A.R. The Science of Self-Report. *APS Observer* 1997. Available online: https://www.psychologicalscience.org/observer/the-science-of-self-report (accessed on 22 April 2025).

11. Razavi, T. *Self-Report Measures: An Overview*; Discussion Papers in Accounting and Management Science; University of Southampton: Southampton, UK, 2001.

12. Simms, L.J.; Zelazny, K.; Williams, T.F.; Bernstein, L. Does the number of response options matter? Psychometric perspectives using personality questionnaire data. *Psychol. Assess.* **2019**, *31*, 557–566. [CrossRef]

13. Elosua, P.; Aguado, D.; Fonseca-Pedrero, E.; Abad, F.; Santamaría, P. New Trends in Digital Technology-Based Psychological and Educational Assessment. *Psicothema* **2023**, *1*, 50–57. [CrossRef] [PubMed]

14. Goldhammer, F.; Scherer, R.; Greiff, S. Editorial: Advancements in Technology-Based Assessment: Emerging Item Formats, Test Designs, and Data Sources. *Front. Psychol.* **2020**, *10*, 3047. [CrossRef]

15. Kim, E.; Laine, T.H.; Suk, H.J.; Jo, Y.W. Using immersive virtual reality in testing empathy type for adolescents. *Curr. Psychol.* **2024**, *43*, 16183–16197. [CrossRef]

16. Ericsson, E.; Sofkova Hashemi, S.; Lundin, J. Fun and frustrating: Students' perspectives on practising speaking English with virtual humans. *Cogent Educ.* **2023**, *10*, 2170088. [CrossRef]

17. Mochimaru, M. Digital Human Models for Human-Centered Design. *J. Robot. Mechatron.* **2017**, *29*, 783–789. [CrossRef]

18. Badler, N.I.; Phillips, C.B.; Webber, B.L. *Simulating Humans: Computer Graphics, Animation, and Control*; Oxford University Press: Oxford, UK, 1993; ISBN 978-1-60129-867-866.

19. Gratch, J.; Wang, N.; Okhmatovskaia, A.; Lamothe, F.; Morales, M.; Van Der Werf, R.J.; Morency, L.-P. Can Virtual Humans Be More Engaging Than Real Ones. In *Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environments*; Jacko, J.A., Ed.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2007; Volume 4552, pp. 286–297, ISBN 978-3-540-73108-5.

20. Gaggioli, A.; Mantovani, F.; Castelnuovo, G.; Wiederhold, B.; Riva, G. Avatars in Clinical Psychology: A Framework for the Clinical Use of Virtual Humans. *Cyberpsychol. Behav.* **2003**, *6*, 117–125. [CrossRef]

21. Rizzo, A.; Shilling, R.; Forbell, E.; Scherer, S.; Gratch, J.; Morency, L.-P. Autonomous Virtual Human Agents for Healthcare Information Support and Clinical Interviewing. In *Artificial Intelligence in Behavioral and Mental Health Care*; Elsevier: Amsterdam, The Netherlands, 2016; pp. 53–79, ISBN 978-0-12-420248-1.

22. Paulik, G.; Taylor, C.D.J. Imagery-Focused Therapy for Visual Hallucinations: A Case Series. *Clin. Psychol. Psychother.* **2024**, *31*, e2993. [CrossRef] [PubMed]

23. Rizzo, A.; Bouchard, S. (Eds.) Virtual Reality for Psychological and Neurocognitive Interventions. In *Virtual Reality Technologies for Health and Clinical Applications*; Springer: New York, NY, USA, 2019; ISBN 978-1-4939-9480-9.

24. De Melo, C.M.; Gratch, J.; Marsella, S.; Pelachaud, C. Social Functions of Machine Emotional Expressions. *Proc. IEEE* **2023**, *111*, 1382–1397. [CrossRef]

25. Freiwald, J.P.; Schenke, J.; Lehmann-Willenbrock, N.; Steinicke, F. Effects of Avatar Appearance and Locomotion on Co-Presence in Virtual Reality Collaborations. In Proceedings of the Mensch und Computer 2021, Ingolstadt, Germany, 5–8 September 2021; pp. 393–401.

26. Blascovich, J.; Bailenson, J.N. *Infinite Reality: Avatars, Eternal Life, New Worlds, and the Dawn of the Virtual Revolution*; William Morrow: New York, NY, USA, 2011.

27. Slater, M.; Usoh, M. Presence in Immersive Virtual Environments. In Proceedings of the IEEE Virtual Reality Annual International Symposium, Seattle, WA, USA, 18–22 September 1993; pp. 90–96.

28. Kvapil Varšová, K.; Juřík, V. Using iVR to deliver optimal psychotherapy experience—Current perspectives on VRET for acrophobia. *Front. Psychol.* **2024**, *15*, 1491622. [CrossRef] [PubMed]

29. Iftikhar, Z.; Ransom, S.; Xiao, A.; Huang, J. Therapy as an NLP Task: Psychologists' Comparison of LLMs and Human Peers in CBT. *arXiv* **2024**, arXiv:2409.02244. [CrossRef]

30. Araiza, P.; Keane, T.L.; Beaudry, J.; Kaufman, J. Immersive Virtual Reality Implementations in Developmental Psychology. *Int. J. Virtual Real.* **2020**, *20*, 1–35. [CrossRef]

31. Riva, G. Virtual Reality in Clinical Psychology. In *Comprehensive Clinical Psychology*; Elsevier: Amsterdam, The Netherlands, 2022; pp. 91–105, ISBN 978-0-12-822232-4.

32. Elor, A.; Kurniawan, S. The Ultimate Display for Physical Rehabilitation: A Bridging Review on Immersive Virtual Reality. *Front. Virtual Real.* **2020**, *1*, 585993. [CrossRef]

33. Schöne, B.; Kisker, J.; Lange, L.; Gruber, T.; Sylvester, S.; Osinsky, R. The reality of virtual reality. *Front. Psychol.* **2023**, *14*, 1093014. [CrossRef] [PubMed]

34. Quah, T.C.S.; Lau, Y.; Ang, W.W.; Lau, S.T. Experiences of immersive virtual reality in healthcare clinical training for nursing and allied health students: A mixed studies systematic review. *Nurse Educ. Today* **2025**, *148*, 106625. [CrossRef]

35. Jespersen, A.E.; Lumbye, A.; Schandorff, J.; Damgaard, V.; Glenthøj, L.B.; Nordentoft, M.; Mikkelsen, C.; Didriksen, M.; Ostrowski, S.R.; Vinberg, M.; et al. Cognition Assessment in Virtual Reality (CAVIR): Associations with neuropsychological performance and activities of daily living in patients with mood or psychosis spectrum disorders. *J. Affect. Disord.* **2025**, *369*, 1053–1063. [CrossRef]

36. Felnhofer, A.; Kothgassner, O.D.; Hauk, N.; Beutl, L.; Hlavacs, H.; Kryspin-Exner, I. Physical and social presence in collaborative virtual environments: Exploring age and gender differences with respect to empathy. *Comput. Hum. Behav.* **2014**, *31*, 272–279. [CrossRef]

37. Shin, D. Empathy and embodied experience in virtual environment: To what extent can virtual reality stimulate empathy and embodied experience? *Comput. Hum. Behav.* **2018**, *78*, 64–73. [CrossRef]

38. Lucas, G.M.; Rizzo, A.; Gratch, J.; Scherer, S.; Stratou, G.; Boberg, J.; Morency, L.-P. Reporting Mental Health Symptoms: Breaking Down Barriers to Care with Virtual Human Interviewers. *Front. Robot. AI* **2017**, *4*, 51. [CrossRef]

39. Rizzo, A.; Scherer, S.; DeVault, D.; Gratch, J.; Artstein, R.; Hartholt, A.; Lucas, G.; Marsella, S.; Nazarian, A.; Stratou, G.; et al. Detection and computational analysis of psychological signals using a virtual human interviewing agent. *Virtual Real.* **2014**, *9*, 311–321.

40. Suárez, G.; Jung, S.; Lindeman, R.W. Evaluating Virtual Human Role-Players for the Practice and Development of Leadership Skills. *Front. Virtual Real.* **2021**, *2*, 658561. [CrossRef]

41. Sapkaroski, D.; Mundy, M.; Dimmock, M.R. Immersive virtual reality simulated learning environment versus role-play for empathic clinical communication training. *J. Med. Radiat. Sci.* **2022**, *69*, 56–65. [CrossRef] [PubMed]

42. Jabbi, M.; Swart, M.; Keysers, C. Empathy for positive and negative emotions in the gustatory cortex. *NeuroImage* **2007**, *34*, 1744–1753. [CrossRef]

43. Reniers, R.L.E.P.; Corcoran, R.; Drake, R.; Shryane, N.M.; Völlm, B.A. The QCAE: A Questionnaire of Cognitive and Affective Empathy. *J. Pers. Assess.* **2011**, *93*, 84–95. [CrossRef]

44. Jung, S.; Kim, E. Development and initial validation of the Multidimensional Empathy Scale for Adolescents. *Asian J. Soc. Psychol.* **2023**, *26*, 488–503. [CrossRef]

45. Kim, C.-S.; Jung, M.; Kim, S.-Y.; Kim, K. Controlling the Sense of Embodiment for Virtual Avatar Applications: Methods and Empirical Study. *JMIR Serious Games* **2020**, *8*, e21879. [CrossRef]

46. Dewe, H.; Gottwald, J.M.; Bird, L.-A.; Brenton, H.; Gillies, M.; Cowie, D. My Virtual Self: The Role of Movement in Children's Sense of Embodiment. *IEEE Trans. Vis. Comput. Graph.* **2022**, *28*, 4061–4072. [CrossRef]

47. Reallusion Character Creator 4. Available online: https://www.reallusion.com/character-creator/ (accessed on 10 April 2025).

48. Noitom Perception Neuron 3. Available online: https://www.noitom.com/perception-neuron-3 (accessed on 10 April 2025).

49. Roth, D.; Lugrin, J.-L.; Latoschik, M.E.; Huber, S. Alpha IVBO—Construction of a Scale to Measure the Illusion of Virtual Body Ownership. In Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017; pp. 2875–2883.

50. Witmer, B.G.; Singer, M.J. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence Teleoper. Virtual Environ.* **1998**, *7*, 225–240. [CrossRef]

51. Baehre, S.; O'Dwyer, M.; O'Malley, L.; Lee, N. The use of Net Promoter Score (NPS) to predict sales growth: Insights from an empirical investigation. *J. Acad. Mark. Sci.* **2022**, *50*, 67–84. [CrossRef]

52. Suarez, J.; Murphy, R.R. Hand gesture recognition with depth images: A review. In Proceedings of the 2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication, Paris, France, 9–13 September 2012; pp. 411–417.

53. Guy, M.; Normand, J.-M.; Jeunet-Kelway, C.; Moreau, G. The sense of embodiment in Virtual Reality and its assessment methods. *Front. Virtual Real.* **2023**, *4*, 1141683. [CrossRef]

54. Nwagu, C.; AlSlaity, A.; Orji, R. EEG-Based Brain-Computer Interactions in Immersive Virtual and Augmented Reality: A Systematic Review. *Proc. ACM Hum.-Comput. Interact.* **2023**, *7*, 1–33. [CrossRef]

55. Mori, M.; MacDorman, K.; Kageki, N. The Uncanny Valley [From the Field]. *IEEE Robot. Autom. Mag.* **2012**, *19*, 98–100. [CrossRef]