Avatars in Individualized Virtual Reality: The Impact of Avatar Design on Self-Compassion Therapy in VR

Thomas Elliott, Yanzhuo Yang, Jarrod Knibbe, Julie D. Henry, Nilufar Baghaei

Submitted to: JMIR XR and Spatial Computing (JMXR)
on: January 08, 2025

Table of Contents

Avatars in Individualized Virtual Reality: The Impact of Avatar Design on Self-Compassion Therapy in VR

Thomas Elliott[1*] BSc, MIDes; Yanzhuo Yang[1*] BEng, MIDes; Jarrod Knibbe[1] MEng, PhD; Julie D. Henry[2*] PhD; Nilufar Baghaei[1*] PhD

[1]School of Electrical Engineering and Computer Science Faculty of Engineering, Architecture and Information Technology University of Queensland Brisbane AU
[2]School of Psychology Faculty of Health and Behavioural Sciences University of Queensland Brisbane AU
[*]these authors contributed equally

Corresponding Author:
Thomas Elliott BSc, MIDes
School of Electrical Engineering and Computer Science
Faculty of Engineering, Architecture and Information Technology
University of Queensland
General Purpose South (78), University of Queensland, St Lucia QLD 4067 Level 4
Brisbane
AU

Abstract

Background: The role of avatars within virtual reality is increasingly important. Avatars have been introduced with increasing quality mainly due to historical technological limitations with respect to computer graphics and hardware. The VR community may assume that with the development of higher quality and therefore more 'lifelike avatars', this implies an advance in avatar performance, yielding ever better therapeutic outcomes in clinical applications, such as self-compassion therapy. But is that true? If not then what can be done to optimize human-avatar interactions for therapeutic benefit?

Objective: Considering the use of avatars in VR for self-compassion therapy, we conduct a study which focuses on improvements to user experience via avoidance of 'excessive realism' with the introduction of low-fidelity avatars, increased individualization and enhanced embodiment.

Methods: This study has acquired the original VR environment of Halim et al.'s [1] such that work can be done to continue an investigation into design changes and the effect on user experience and the therapeutic outcomes. This study used the User Experience Questionnaire (UEQ), Self-Compassion Scale (SCS), and Patient Health Questionnaire depression scale (PHQ-8) to measure experimental changes.

Results: The results show an improvement to self-compassion outcomes following introduction of greater avatar customisation, SCS $p = 0.03$. An increase in user experience and therapeutic outcomes for depression was observed after introducing mirror techniques to enhance body ownership over a personalized aesthetically low-fidelity avatar; PHQ-8 $p = 0.01$ and UEQ all categories $p < 0.04$ with exception 'Efficiency' category. Also, with respect to user experience, significant improvements in areas pertaining to feelings of dependability suggests avoidance of Uncanny Valley phenomenon.

Conclusions: Our work explores important questions to strengthen the community's knowledge of design decisions surrounding avatars that lead to optimal therapeutic outcomes when engaging VR therapy. This paper has demonstrated the harmlessness of low-fidelity avatars in VR therapy, reinforced the importance of individualizing aspects of virtual reality therapy and demonstrated the importance of virtual mirrors with respect to user experience in such VR therapy applications. Clinical Trial: This study is not RCT.

(JMIR Preprints 08/01/2025:71004)
DOI: https://doi.org/10.2196/preprints.71004

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.
2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?
✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Original Paper

Thomas C Elliott MIDes
Email: t.elliott1@uq.net.au
ORCID: 0009-0004-1836-8847
Affiliation: School of Electrical Engineering and Computer Science, The University of Queensland, Brisbane, Queensland, Australia.

Yanzhuo Yang MIDes
Email: yanzhuo.yang@uq.net.au
ORCID: 0009-0009-9331-4191
Affiliation: School of Electrical Engineering and Computer Science, The University of Queensland, Brisbane, Queensland, Australia.

Jarrod Knibbe MEng PhD
Email: j.knibbe@uq.edu.au
ORCID: 0000-0002-8844-8576
Affiliation: School of Electrical Engineering and Computer Science, The University of Queensland, Brisbane, Queensland, Australia.

Julie D Henry PhD
email: julie.henry@uq.edu.au
ORCID: 0000-0002-2081-3717
Affiliation: Affiliate Professor of Mater Research Institute-UQ & School of Psychology, The University of Queensland, Brisbane, Queensland, Australia.

Nilufar Baghaei PhD
email: n.baghaei@uq.edu.au
ORCID: 0000-0003-1776-7075
Affiliation: School of Electrical Engineering and Computer Science, The University of Queensland, Brisbane, Queensland, Australia.

Avatars in Individualized Virtual Reality: The Impact of Avatar Design on Self-Compassion Therapy in VR

Abstract

Background: The role of avatars within virtual reality is increasingly important. Avatars have been introduced with increasing quality mainly due to historical technological limitations with respect to computer graphics and hardware. The VR community may assume that with the development of higher quality and therefore more 'lifelike avatars', this implies an advance in avatar performance, yielding ever better therapeutic outcomes in clinical applications, such as self-compassion therapy. But is that true? If not then what can be done to optimize human-avatar interactions for therapeutic benefit?

Objective: Considering the use of avatars in VR for self-compassion therapy, we conduct a study which focuses on improvements to user experience via avoidance of 'excessive realism' with the introduction of low-fidelity avatars, increased individualization and enhanced embodiment.

Methods: This study has acquired the original VR environment of Halim et al.'s [1] such that work can be done to continue an investigation into design changes and the effect on user experience and the therapeutic outcomes. This study used the User Experience Questionnaire (UEQ), Self-Compassion Scale (SCS), and Patient Health Questionnaire depression scale (PHQ-8) to measure experimental changes.

Results: The results show an improvement to self-compassion outcomes following introduction of greater avatar customisation, SCS **p = 0.03**. An increase in user experience and therapeutic outcomes for depression was observed after introducing mirror techniques to enhance body ownership over a personalized aesthetically low-fidelity avatar; PHQ-8 **p = 0.01** and UEQ all categories p < **0.04** with exception 'Efficiency' category. Also, with respect to user experience, significant improvements in areas pertaining to feelings of dependability suggests avoidance of Uncanny Valley phenomenon.

Conclusions: Our work explores important questions to strengthen the community's knowledge of design decisions surrounding avatars that lead to optimal therapeutic outcomes when engaging VR therapy. This paper has demonstrated the harmlessness of low-fidelity avatars in VR therapy, reinforced the importance of individualizing aspects of virtual reality therapy and demonstrated the importance of virtual mirrors with respect to user experience in such VR therapy applications.

Trial Registration: This study is not RCT.

Keywords: Avatar Design; Virtual Reality; Self-Compassion; Depression; Mental Health; Individualised VR; User experience; Uncanny Valley

Figure 1. A participant viewing their customized low-fidelity avatar through a virtual mirror in the VR Self-Compassion Therapy application.

Introduction

Virtual Reality (VR) is proving a capable tool for therapeutic interventions for depression [2, 3]. In one application, for example, users embody avatars to both give and receive compassion - leading to positive self-compassion outcomes. Research has enhanced outcomes further by allowing users to personalize these experiences, through choices of avatars and environment features [1]. To date, however, the role of such factors has rarely been tested in a systematic manner, and as a consequence, the impact of different design decisions on therapeutic outcomes remains poorly understood.

In this paper, we examine the impact of avatar appearance and environmental designs which support avatar interactions, to improve user experience and self-compassion outcomes in VR therapy. On one hand, improving personalization of avatars has been shown to improve body ownership, agency, and immersion [4]. These are all factors that may improve the users experience with the iVR therapy. On the other hand, research has shown placement of virtual mirrors in VR self-counseling settings to be beneficial for both therapy and VR conditions which support user experience [5, 6]. The 'conditions' of interest to this study include; body ownership, agency, presence and immersion. To determine the feasibility of design changes to optimize therapeutic outcomes we ask the following questions:

- RQ1. Will avatar and environmental design features influence participant experience?
- RQ2. Will avatar and environmental design features influence benefits of therapy, operationalized here as self-rated self-compassion.

To better understand the impact of specific design changes on user experience and self-compassion outcomes in VR therapy, we conducted three studies, exploring the impact of (1) low-fidelity avatars, (2) customizable low-fidelity avatars, and (3) customizable, low-fidelity avatars with the presence of a virtual mirror. Our studies are based on prior work [1], where we varied only the features mentioned above. We recruited 104 participants in total to take part in three separate studies. To gain initial insights into whether specific design features do influence therapeutic outcomes, only healthy neurotypical individuals were eligible to participate. Our results show that changes in avatar realism have no impact on user experience or therapeutic outcomes . However, customization of avatars improved self-compassion, whilst placement of virtual mirrors improved depression scores and user experience.

Taken together, our work suggest that the relentless pursuit of realism may not be necessary for VR therapy, where leveraging simpler, low-fidelity avatars does not result in different outcomes to their more realistic counterparts. Enhancing VR conditions through customization and the presence of a mirror, however, does appear to have potential benefits, and it is theorized that this may be because participants can maintain a greater self-awareness throughout their interactions.

Background

Our work explores the impact of avatar and virtual environment design choices on outcomes in self-compassion therapy in VR. To situate our work, we discuss key theories that support virtual experiences, including impact on avatar perceptions and use of virtual mirrors. We also, describe the known benefits of individualized approaches to virtual reality and discuss prior work on compassion in avatar interactions.

Exploration of Embodiment Conditions and Mirrors

Our study should also consider the concepts of embodiment and body ownership, presence, and avatar agency to foster therapeutic potential, as seen in Falconer et al., who has previously explored the benefits of providing immersive virtual experiences for self compassion therapy. [7]. Therefore, this section will explore these concepts as they will be important for supporting the proposed iVR therapy.

This study will refer to Embodiment and Body Ownership, as it is defined by Kilteni et al. [8]. According to Kilteni et al. historically, the definition of Embodiment has been the source of confusion within the VR community due to its multidisciplinary use. However, it may be described as "the sense that emerges when [a body] properties are processed as if they were the properties of one's own biological body" [8]. Body ownership "refers to one's self-attribution of a body" [8] and can be achieved under the right conditions. These conditions include; self-location, agency and body ownership [8]. Importantly, if these conditions become strong enough then it is possible for a user to achieve embodiment of an avatar that is dissimilar to the user's body in the real world. As demonstrated by Slater et al. “The experiment had the unusual goal of attempting to generate a body ownership illusion where the virtual body did not visually resemble the real body of the participants and was not even the same gender.” [9]. This understanding will be useful for our study as we will require participants to embody low-fidelity avatars.

To achieve this the researchers of this study will need to consider support of these ’conditions’. Starting with Agency, which describes "global motor control, including the subjective experience of action, control, intention, motor selection and the conscious experience of will" [8, 10]. Whilst the user experiences VR a “sense of authorship of an action” [11] enables the experience of the body's action in the virtual world affording agency. Banakou and Slater also described agency as a participant's experience of choosing actions at freewill and so body ownership comes from the experience of that action [11]. With respect to providing healthy embodiment, the final condition which needs consideration is Self-Location, which according to Kilteni et al. can be enhanced via the use of first person perspective (1PP). It's important to note that ’self-location’ and ’presence’, though closely related, are distinct with respect to a relationship with ’the self’; the former is concerned with one's body and the later one's environment [8]. Therefore, presence also deserves consideration.

Presence has been described as "A human reaction to immersion" [12] which leads to a "sense of 'being there' in a virtual environment" [13]. Hence immersion and presence have also shared historical entanglement as according to Grassini et al. "immersion is a concept connected to—but not equivalent to—presence" [14]. Fortunately some of this 'entanglement' has been cleared by the work of Wirth et al. which brought perspectives from communication and psychology to infer that highly immersive technologies (such as VR) may enable an experience of presence for a user [15].

Immersion is described as "The technical capability of the system to deliver a surrounding and convincing environment with which the participant can interact." [16]. Therefore, in practice, this means that for a successful immersive experience the authors of this study have considered appropriate technology. This study has used the reputable Meta Quest 2 head mounted display (HMD) typically used for immersive gaming. Selection of an HMD is justifiable with regard to promoting immersion as HMDs provide high-immersive multi-sensory stimuli [17].

Facilitating a VR experience which supports these introduced VR concepts will provide support for the iVR therapy via user acceptance of the virtual environment. This has been seen in previous VR studies, where the power of body ownership illusion has been explored by Banakou and Slater [11], based on findings by Tajadura-Jiménez et al. [18], which revealed that under experimental VR conditions, adults could embody child avatars, even accepting a child's voice as their own. This powerful, though unusual, finding demonstrates it is possible to create an illusion of agency over a body even when it is markedly distinct from one's actual body, providing important proof-of-concept that VR body ownership can occur over virtually any avatar when conditions allow, making the use of low-fidelity avatars feasible. Also, Slater et al. was able to demonstrate that participant positioning of 1PP alone was sufficient to facilitate embodiment for avatars in VR [9], a welcome result for practicality, as visuo-motor synchrony can be employed to enhance sense of body ownership in VR effectively [19, 20] which is easily achieved by the chosen HMD.

In the interest of further enhancing embodiment, a virtual mirror [5] is a method of achieving greater visuo-motor synchrony, allowing the user to observe their own real-time movements within the VR environment [6] to provoke embodiment leading to increase immersion. In comparison to when only a 1PP view is provided, this virtual mirror setting allows users to observe all parts of their embodied avatars, including the face [5]. Furthermore, virtual mirrors can also enhance embodiment in VR through human's expectations about mirrors. Humans develop expectations about mirrors based on their experiences, starting from birth [21], enabling them to associate the reflected image of a virtual body with their own body, even if it differs from our biological body [9], [22]. Therefore, the perceptual cues provided by the virtual mirror, which allow the reflected image to be recognized as a mirror reflection [21], can also contribute to the embodied process. This includes mirror symmetry between the reflected image and user's position and movements, and physical appearance similar to real-world mirrors.

Individualized Virtual Reality

The previous subsection describes the conditions necessary to support a VR experience with respect to embodiment via; body ownership, agency, presence and immersion. However, there are more ways to support a virtual experience. Research supports the use of individualized virtual reality (iVR) as a means of improving user experience [1, 3]. iVR includes the customization of avatars, which has benefits for users with respect to identity [23]. In this paper 'customization of avatars' refers to allowing a user to modify an avatar's appearance (face, body, and clothing) to look like themselves. Use of customization has been credited for improvements in other projects where customization on avatars was permitted [24–26], but also customization and other aspects of the virtual experience, such as game difficulty [27, 28], audio [29], and objects from virtual environments [30]. Essentially, these papers are describing the benefit of ‘individualization’ which can be applied to many aspects of the virtual experience.

iVR, as explored by Halim et al. and Baghaei et al. provided preliminary evidence that enabling individualization of the VR experience allows for improvements to user experience and therapeutic outcomes [1, 3]. With respect to avatars, Freeman et al. suggested that users perceive avatars as extensions of themselves in VR, leading to a preference for customized avatars similar to themselves to achieve a "stronger sense of presence, embodiment, and attachment" [31]. Also, Waltemate et al. demonstrated the importance of individualization as their study allows participants to adjust a low-fidelity avatar to look as ‘similar’ to themselves as possible by modifying features such as; clothing, skin color, gender, and figure. For a healthy population, use of this approach led to significant improvements for avatar embodiment [4].

In summary, it remains to be established whether the iVR-related benefits identified by Halim et al. also emerge after systematic adjustments to avatar appearance, avatar customization and stronger embodiment via the introduction of a virtual mirrors. This is an important issue to address, since it speaks to how specific design features might influence user experience and therapeutic outcomes. However, the efficacy of iVR in treating depression remains poorly understood.
The next section discussed what is currently known about human-avatar interaction and self-compassion for treatment of depression.

Compassion & Human-Avatar Interactions

Self-compassion is the ability to “soothe oneself with kindness and non-judgmental understandings in times of difficulty” [32] Self-compassion therapy (CFT) is an intervention designed to assist individuals in developing self-compassion as a skill to combat depression [33] but also a useful predictor of depression [34].

For VR depression therapies Falconer et al. has already gone a long way towards advancing VR therapy interventions in the form of a CFT [2, 7]. Without VR, this therapy is performed with human-to-human interaction but in VR application the intervention is facilitated with non-human avatar characters, for example [1]. However, contemporary studies into forms of VR therapy tend to favor high quality avatars as agents as seen in [35, 36]. The use of such avatars with ’excessive realism’ risks undermining the therapeutic benefits for some clinical populations by adding unnecessary cognitive load [37, 38].

Depression is a debilitating disorder which, even in remission, can lead to "moderate impairments in the domains of executive functioning and attention" [39]. People living with depression therefore have reduced working memory capacity [40, 41]. CFT is a learning activity and like VR applications for learning outcomes they tend to be created detailed and realistic [42] to assist immersion.

This added immersion increases cognitive load [43], which risks diminishing the learning experience, even in the healthy population. Researchers in the field of VR learning discuss the importance of considering reductions to cognitive load where possible in line with "The Cognitive Theory of Multimedia Learning" [44] which is predicated on natural limitations in working memory capacity [45]. Designers of virtual experiences should therefore consider these risks prior to selection of virtual assets and reduce cognitive load where possible. This is why in our study we elected to trial the use of low-fidelity avatars, as reduction of avatar realism should reduce cognitive load.

"VR has extraordinary potential to help people overcome mental health problems if high levels of presence are achieved for situations that trouble them." Freeman et al. [46]. But for iVR "The gaps in meaningful applications to mental health are extensive" [46]. In addition to reducing 'excessive realism', this study will explore enhanced avatar customization and enhanced embodiment via virtual mirror placement to improve user experience in the context of self-compassion therapy. By introducing features that should theoretically promote a better user experience, it was hoped that this might lead to further benefits for participants engaging in iVR self-compassion therapy, and specifically, even greater self-compassion.

Methods

We replicate and extend the study by Halim et al. on individualized self-compassion therapy in VR [1]. Halim et al.'s study included an on-boarding lobby for avatar selection, scene selection, and companion selection. There were six avatars to choose from, three available environments, and two companion avatars with two behaviors; crying or upset. The virtual therapy consisted of two stages: Stage 1 involved delivery of compassion to an upset/crying avatar and stage 2 involved receiving of compassion.

Three studies were conducted to examine the impact of avatar design decisions on User experience (UX) and self-compassion outcomes. Study features that were manipulated were: (i) avatar appearance (Study 1), (ii) avatar personalization (Study 2), and (iii) inclusion of virtual mirrors (Study 3). These studies collectively built upon and meaningfully extend Halim et al.'s work [1]. and uses a between-subject study design for comparisons to Study 1 between each iteration. We obtained the original Unity project files [1] and made edits to include low-fidelity avatars, avatar customization and virtual mirrors.

Recruitment

Participants for all three studies were recruited via the community through posters, social media, and word-of-mouth, with most being students at The University of Queensland (UQ). To be eligible, participants needed to be at least 18 years old and capable of using VR. Ethical approval was obtained from The University of Queensland Human Research Ethics Committee, and all participants provided informed consent prior to participation.

Measures

Since our VR application is intended for self-compassion therapy, studies were compared via participants' responses on the Self-Compassion Scale (SCS) [47] and Patient Health Questionnaire depression scale (PHQ-8) [48]. SCCS evaluates users' responses in various scenarios on a 7-point Likert Scale, assessing both self-compassion and self-criticism levels, and providing a comprehensive measurement for this research. The PHQ-8 is an eight-item self-report measure of depression that asks participants how many days in the past 2 week time-frame that they experienced specific depressive symptoms. Higher scores indicate greater depression, with scores greater than or equal to 10 indicative of current depression.

UX data were also collected, through the User Experience Questionnaire (UEQ) [49]. UEQ provides results from averaged 7-point Likert scale scores and benchmarks these scores against normative data across: Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation and Novelty. As discussed previously, a users' sense of presence has a relationship with support of embodiment. Here, presence was measured as a part of the UX analysis in study 3, using the presence questionnaire developed by Slater, Usoh, and Steed (SUS) [13, 50], which includes a 7-point scale with 6 questions. SUS was chosen over a more direct measurements for embodiment because this type of questionnaire is more appropriate for detection of the virtual mirror stimuli as it pertains more to user experiences in a virtual environment.

In addition to collecting data for the above surveys/questionnaires, we included an additional questionnaire to collect qualitative feedback. To elicit responses related to virtual experience, dislikes, and impacts of avatars encounters, the questionnaire asked the following:

- What were the top three things that you liked about individualized VR?
- How do you think the next version can be improved?
- (Only in study 1 and 2) How do you feel about your experience with the avatars?
- (Only in study 3) How do you feel about your experience with the mirror? What aspects of the mirror did you like or dislike?

Studies 1 and 2 included avatar questions but did not include mirror questions as a mirror had not yet been introduced. Results of the questionnaires were analyzed using Thematic analysis [51].
Study 1 consisted of one experimental session, where Study 2 and Study 3 included 2 experimental sessions where the following assessments occurred:

Study 1, 2 and 3. Session 1:
1. Baseline SCS Assessment prior to iVR Exposure
2. For Session 2 & 3 only, Baseline PHQ-8 Assessment prior to iVR Exposure
3. iVR Exposure
4. Post Exposure UEQ Assessment
5. Post Exposure SCS Assessment
6. Post Exposure Qualitative questionnaire

Study 2 & 3. Session 2:
1. iVR Exposure
2. Post Exposure UEQ Assessment
3. Post Exposure SCS Assessment
4. Post Exposure PHQ-8 Assessment
5. Post Exposure SUS Assessment (for Study 3 only)
6. Post Exposure Qualitative questionnaire

Given the unknown consequences of using low-fidelity avatars, Study 1 was included as a pilot study to test feasibility, the following studies included to accommodate a 2 week interval between VR therapy exposures due to a condition of the PHQ-8 [48].

Procedure

iVR Experience: On-boarding the participant

Prior to engaging with the virtual environment, participants follow a sequence of menus through which they can individualize the experience. This process includes selecting an avatar, selecting a therapeutic environment (Figure 2), and selecting a companion Non-Player Character (NPC).

During the avatar personalization, participants are encouraged to produce a character that looks similar to themselves (within the scope of the tool). When the participant is in the therapy environment, they can observe their surroundings by looking around, however, they are unable to move from their fixed position. This is a deliberate design choice to keep the participant focusing on the task of comforting the avatar and allows the participant to believe they are in a large space.

iVR Experience Stage 1

Giving Compassion. Once the iVR therapy launches the participant begins the experience in their chosen environment. In that environment the participant was joined by their chosen companion which then acted out the mode selected in the on-boarding stage (crying or upset). To the immediate left of the participant, in their periphery, there is an instruction panel that guides the participant through three compassion delivery methods: 'validation', 'redirection of attention', and 'memory activation'. Each method was attempted in this stage, Stage 1. These three strategies of compassion delivery are known as the 'three-stage approach' for emotional situations [2], and they are defined as:

- "Validation: The aim of this stage is to acknowledge that the other person is upset that you do not judge them for this, and that it is perfectly acceptable for them to react in this way.
- Redirection of Attention: The aim of this stage is to direct the other person's attention towards something that is more positive, soothing, and comforting.
- Memory Activation: The aim of this stage is to suggest that the person could try to recall a memory of a person who loves or is kind to them. This memory is supposed to instill more positive feelings of warmth, comfort, and safety." [2].

The participant is expected to navigate through the scene as guided by the instructional panels. After attempting each strategy of compassion delivery, the participant then enters stage 2.

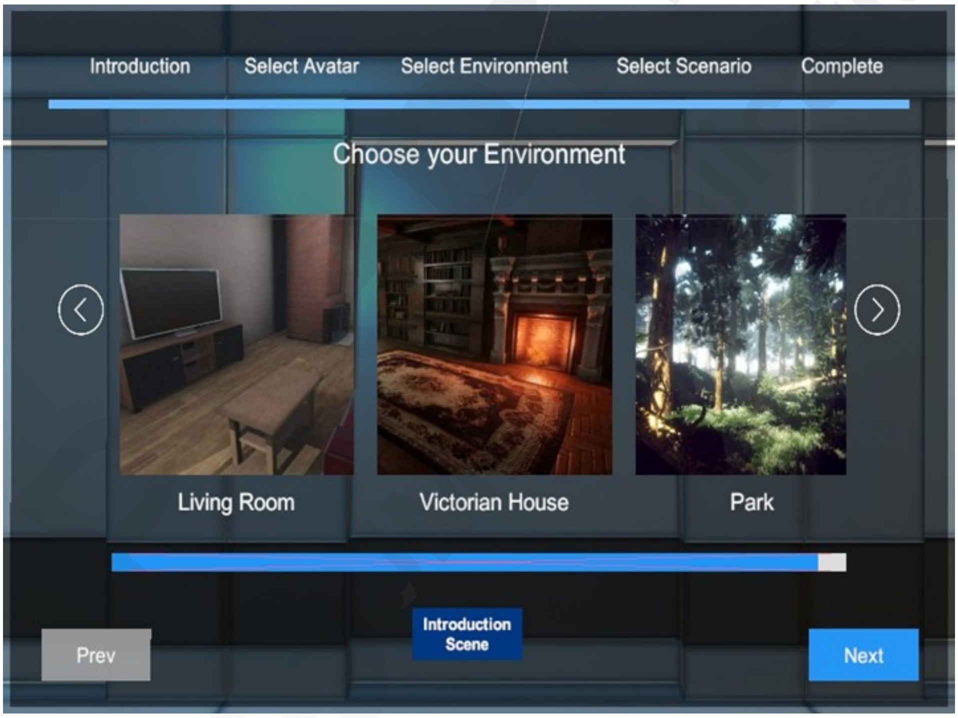


Figure 2. Individualization of virtual therapy via selection of preferred therapeutic environment.

iVR Experience Stage 2: Receiving Compassion

For the participant this final stage holds the potential for a meaningful therapeutic experience and forms the primary purpose of the iVR application. The participant enters the same environment as selected earlier but this time from a different perspective. The NPC from Stage 1 is no longer present, instead the participant is in the NPC's place, viewing their own character. From this perspective the participant then listens to the sound of their own voice as recorded in the previous stage. This voice will also be mimed by 'their' avatar. This exposes the participant to hearing themselves give compassion, which is a key feature of the therapy. This feature of the study aligns efforts of past VR compassion based therapy interventions [1, 7]. When the scene completes the participant can exit and will reappear in the on-boarding scene, thus completing the session. A second session[1] is then scheduled for each participant which is expected to occur approximately two weeks later.

Analyses

The present series of studies focused on two key comparisons: the internal comparison within each study itself (whether there were significant changes in levels of self-compassion, depression, and Usability after completion of the 2 sessions relative to baseline levels), and the comparison between the three studies with different conditions introduced (whether there were significant differences in levels of self-compassion, depression, and Usability across the studies at the time of completion).

All analyses were first checked for the normality of the data to be compared using the Shapiro-Wilk normality test. If the data were determined to be normal, they were further tested for significance using two-tailed t-test; and if the normality distribution was violated, data was examined using the Mann-Whitney U-test. For all analyses, the null hypothesis was that there were no significant differences between the data being compared.

Results

Study 1: Examining the Effects of low-fidelity Avatars

Figure 3. Low-fidelity (left) versus humanoid (right) avatar appearance.

A total of 20 participants completed Session 1, Session 1 had 12 self-identified Females, 6 self-identified Males, and 2 undisclosed. To explore RQ1 and RQ2; this study replaced all avatars from [1] with low-fidelity avatars purchased from SunBox Games [52]. The low-fidelity avatars were chosen because they would introduce less cognitive load given the reduced avatar realism of the avatar features, as can be seen in Figure 3. In this study, a selection of avatars will be offered for approximate resemblance of the participant, see Figure 4. The purpose of this study is a preliminary check that UX will not be severely impacted by the introduction of the chosen avatars, but also to establish preliminary data for further comparisons.

[1] Only Study 2 and 3 have 2 sessions.

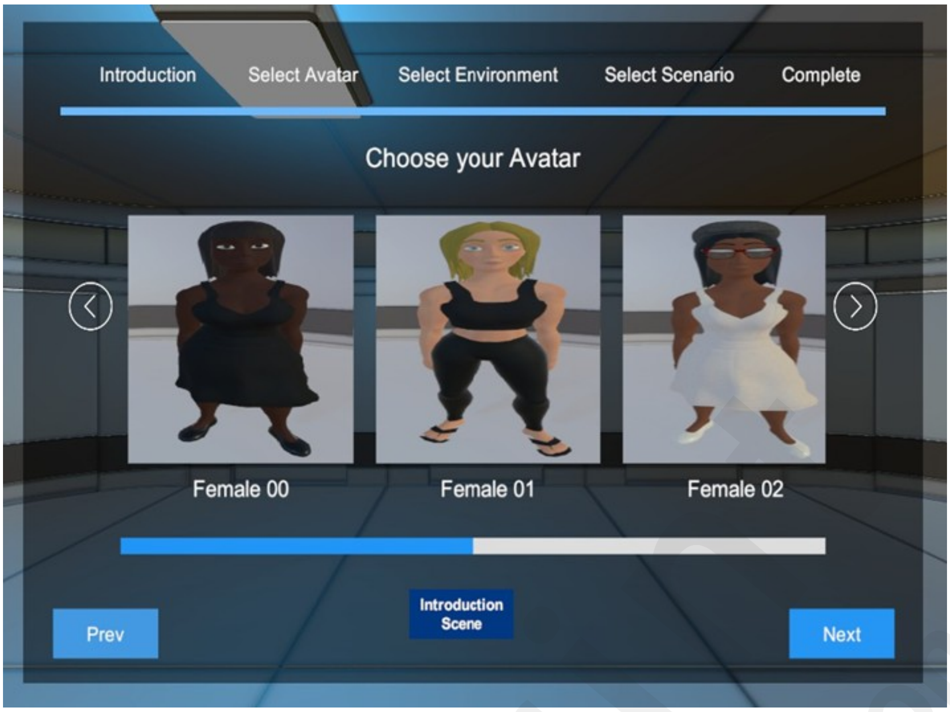


Figure 4. Regressed customization of low-fidelity Avatar, i.e. limited choices for user.

Session 1 of Study 1 had 9 Females and 11 Males, hence a total of 20 participants. The participants were recruited from our institution.

Impact of low-fidelity Avatars on UEQ

At the time of study 1 analysis, the UEQ comprised of a data set containing 21175 people from 468 studies evaluating technology such as software, websites, and social networks to give researchers measurable quality for new technology compared to known products [49]. As can be seen from Figure 5. UEQ increased in ‘Perspicuity’ by a category from their original positions in Halim et al. [1]. From this study, it appears that the use of low-fidelity avatars does not meaningfully negatively impact user experience as all categories are at least ’above average’. Therefore, with respect to RQ1, there is support for the use of low-fidelity avatars.

Impact of low-fidelity Avatars on SCS

This study has produced a mean score of approximately 3, as seen in Table 1. It’s important to note that "There has not been norms established for what constitutes low, medium, or high levels of self-compassion" [53], However according to Neff & Toth-Kiraly a score less than or equal to 2.4 may be considered low. A score between 2.4 and 3.6 is average, and a score greater or equal to 3.6 is high [53]. Hence the respondents of study 1 appear to have ’average’ amounts for self-compassion at the completion of one VR therapy exposure.

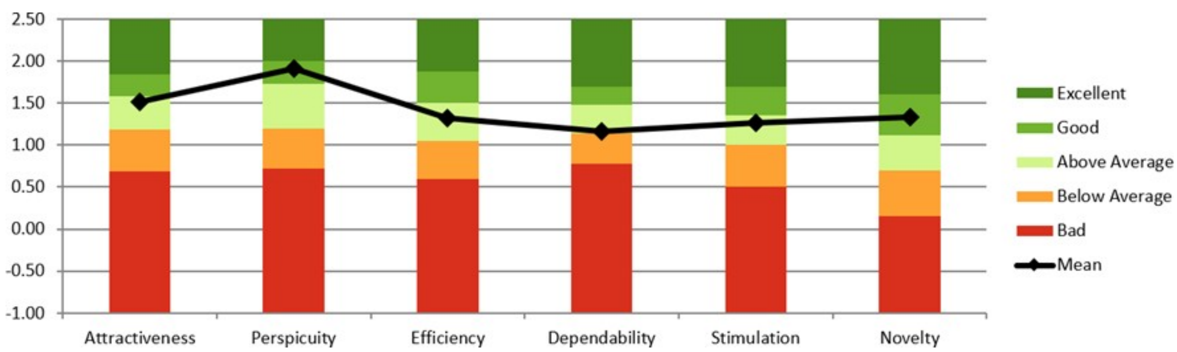


Figure 5. UEQ Study 1. Introduction of low-fidelity avatars.

Study 1: Summary Statistics

Table 1. Summary Statistics for Study 1 (Session 1 n = 20)

	UEQ Scale	Mean (SD)
SCS		
		3.11 (1.14)
UEQ		
	Attractiveness	1.51 (0.97)
	Perspicuity	1.91 (1.33)
	Efficiency	1.32 (1.11)
	Dependability	1.17 (0.86)
	Stimulation	1.26 (1.13)
	Novelty	1.33 (0.97)

Study 2: low-fidelity Avatar Customization

A total of 49 participants completed Session 1, with 42 of these returning to complete Session 2. Session 1 had 22 Female, 22 Male and undisclosed 5 and Session 2 had 18, Female, 21 Male and 3 undisclosed. Study 1 results suggests that UX has not been impacted by the introduction of cartoon avatars, therefore it is safe to retain the low-fidelity avatars. In Study 2, the researchers will use iVR theory by enhancing avatar customization features. The feedback from participants included the comment that "options for avatar styles could have more variations (e.g. hairstyle, colors, dress types)". Study 2 therefore included greater personalization of the participants low-fidelity avatar to help participants achieve greater embodiment (Figure 6). This change occurred in the on-boarding phase. Study 2 was therefore identical to Study 1 except that it included enhanced avatar customization and a second iVR exposure, session 2, 2 weeks after Session 1.

Figure 6. Enhanced customization of low-fidelity avatar

Impact of Avatar Customization on UEQ: Study 1 & Benchmark comparison

At the time of analysis for Study 2, the same UEQ tool was used as there had been no updates since Study 1. As can be seen in Figure 5 in comparison to Study 1, UEQ decreased in 'Perspicuity', 'Dependability' and 'Novelty' by one category with only dependency dropping 'below average' against the benchmark. Results of significance tests are reported in the 'cross study comparison' section of this paper, and reveals the UX comparisons were not significant, hence with respect to RQ1, the introduction of avatar customization has not impacted UX.

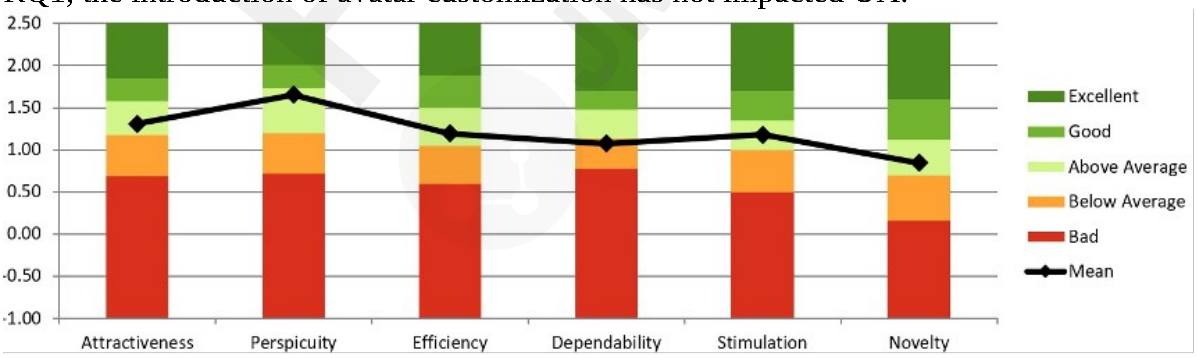


Figure 7. UEQ Study 2. Introduction of avatar customization

Impact of Avatar Customization on SCS: Internal Comparison

With respect to SCS scores, for RQ2 the researchers assume the null hypothesis $H_0 : S_{2B} = S_{2S2}$ Where the SCS results of Study 2 Baseline is S_{2B} and SCS results of Study 2 Session 2 is S_{2S2}. Then the alternative is $H_1 : S_{2B} \neq S_{2S2}$, there is a significant difference between

the two studies. A test for normality using the Shapiro-Wilk normality test via RStudio statistical software determined that the data sets are normal, hence t-tests is appropriate.
A two-tailed independent t-test between was used to compare SCS results from Study 1 and Study 2. As seen in Table 2, the null hypothesis was retained, with introduction of avatar customization producing no difference between the two studies. Since the introduction of low-fidelity avatars and avatar customization has had no impact on therapeutic results, this result validates further investigation into Study 3.

Impact of Avatar Customization on PHQ-8: Internal Comparison

With respect to PHQ-8 scores, for RQ2 the researchers assume the null hypothesis $H_0 : S2_B = S_{2S2}$ Where the PHQ-8 results of Study 2 Baseline is S_{2B} and PHQ-8 results of Study 2 Session 2 is S_{2S2}. Then the alternative is $H_1 : S_{2B} \neq S_{2S2}$, there is a significant difference between the Baseline and final intervention (session 2).
A test for normality using the Shapiro-Wilk normality test via RStudio statistical software determined that the data sets are not normal. Hence, a Mann-Whitney U-test between two independent groups was used to test whether there were significant differences between the PHQ-8 results of Study 2 Baseline and Session 2 results. As can be seen in Table 2, no significant difference was identified. Since introduction of low-fidelity avatars and avatar customization has had no impact on therapeutic results, this result validates further investigation into Study 3.

Study 2: Summary Statistics.

Table 2. Summary Statistics & Internal Comparisons for Study 2 (Session 1 n=49, Session 2 n=41)

	UEQ Scale	Mean (SD)			
UEQ					
	Attractiveness	1.31 (0.85)			
	Perspicuity	1.65 (0.96)			
	Efficiency	1.19 (0.79)			
	Dependability	1.07 (0.81)			
	Stimulation	1.18 (0.96)			
	Novelty	0.85 (1.01)			
		Mean (SD)		*t* test	P value
SCS					
	Baseline	3.05 (0.98)			
	Session 2	3.55 (1.16)			
	Comparison			$t_{89} = 2.219$	.03
		Mean Rank	Mann-Whitney U	*z*	P value (2-tailed)
PHQ-8					
	Baseline	46.16			
	Session 2	44.71			
	Comparison		*U* = 972	0.26	.79

Study 3: Introduction of Mirrors

A total of 38 participants completed Session 1, with 35 of these returning to complete Session 2.

Session 1 had 20 Females and 18 Males, Session 2 had 16 Females, and 19 Males. Study 3 included virtual mirrors because these should theoretically strengthen participant's attachment to their avatar via strengthened embodiment over the low-fidelity avatar, which may lead to better outcomes.

Figure 8. Mirror positioned for the participant (blue shirt) can see themselves interact with the upset low-fidelity avatar (red dress).

Virtual mirrors were placed on the right side of the two characters, as this position does not interfere with the participant's main task, i.e., comforting the crying companion avatar. From their vantage point, the participant is also able to observe the real-time reflected as they interact in Stage 1 (Figure 8). As discussed earlier, this was theorized to improve embodiment for the participant.

Impact of Virtual Mirrors on UEQ: Study 2 & Benchmark comparison

At the time of analysis for Study 3, there had been no updates to the UEQ tool since Study 2. As can be seen from Figure 7 and Figure 9, UEQ increased by at least one category in 'Attractiveness', 'Perspicuity', 'Dependability' and, 'Stimulation' in comparison to Study 2. Three of the six scales achieved 'Excellent' against the benchmark.

As noted previously, mirrors have been shown in other studies to be of benefit [2], [6], [5], and [21]. For user experience, a comparison between Studies 0 and 3 shows a clear improvement across four of the six scales. Therefore Study 3 suggests that with respect to RQ1, UX can be improved by including subtle virtual mirror placement. As noted, with respect to mechanisms, this is thought to be because mirrors should support greater embodiment and body ownership in this context of VR therapy.

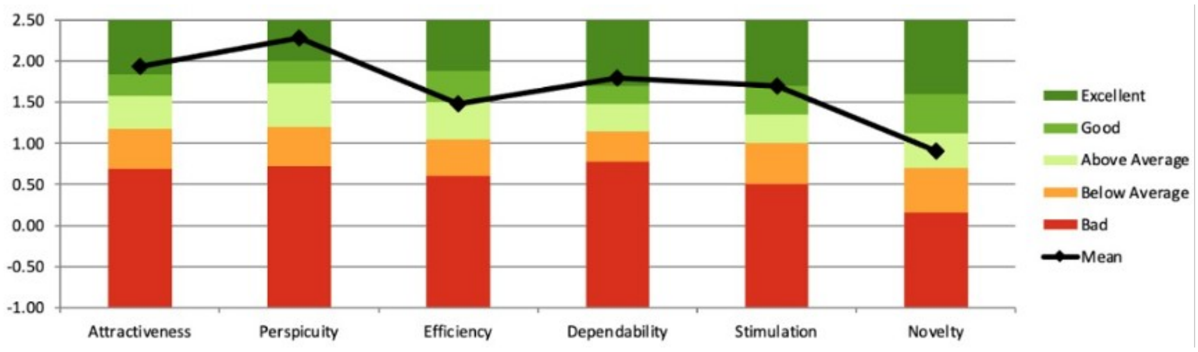


Figure 9. UEQ Study 3. Introduction of virtual mirrors.

Impact of Virtual Mirrors on SCS: Internal Comparison

Next, comparisons were made between Baseline and Session 2 of Study 3 following two exposures to the iVR treatment. After removing 5 outlying data points, a test for normality using the Shapiro-Wilk normality test via RStudio statistical software revealed the data sets are normal, hence analyses using t-tests is appropriate. $H_0 : S_{3B} = S_{3S2}$ Where the SCS results of Baseline is S_{3B} and SCS results of Session 2 are S
. Then the alternative is $H1 : S3B \neq S3S2$, where there is a significant difference between the two studies.

A two-tailed independent t-test comparison between two means was performed to compare the SCS results of Study 2 and Study 3. As seen in Table 3 this comparison revealed no significant difference, indicating that, with respect to RQ3, the introduction of virtual mirrors has not influenced SCS outcomes.

Impact of Virtual Mirrors on PHQ-8: Internal Comparison.

With respect to PHQ-8 scores, for RQ2 the null hypothesis being tested is again of no difference $H_0 : S_{3B} = S_{3S2}$ Where the PHQ-8 results of Study 3 Baseline is S_{3B} and PHQ-8 results of Study 3 Session 2 is S_{3S2}. The alternative hypothesis, $H_1 : S_{3B} \neq S_{3S2}$ is simply that there is a significant difference between the Baseline and final intervention (session 2).

A test for normality using the Shapiro-Wilk normality test via RStudio statistical software revealed the data sets are not normal, so a Mann-Whitney U-test was used to compare the PHQ-8 results of Study 3 Baseline and Study 3 Session 2. As can be seen in Table 3, the data fails to support the null hypothesis, as the introduction of virtual mirrors led to a difference between the two groups, such that pHQ-8 scores are lower following the introduction of mirrors.

Virtual Mirrors: Slater-Usoh-Steed Questionnaire

As can be seen in Table 3 the mean SUS score for Study 3 was 4.67. Because SUS has no ‘benchmark’ for a mean score comparison, often SUS scores are compared internally over multiple studies. Since Usoh et al. intended to support the validity of their questionnaire by showing that their virtual presence questionnaire results did not surpass scores produced by a real experience, here we can use Usoh et al.’s "reality test" [50] to explore our SUS results. Hence, ideally the results of this study should not meaningfully surpass mean ‘reality’ scores, but perform better the mean ’virtual’ scores of Usoh et al.

Table 3 presents comparisons between the SUS Mean of this study and the Usoh et al. ‘SUS Means’; ‘SUS Virtual’ and ’SUS Real’. Normal distribution of data is assumed given Usoh et al’s presentation of results, hence p-value testing via t-tests is used in this case. To formally test SUS results, the researchers assume the null hypothesis $H_0 : S_{SUS3} = S_{SUSR}$ Where the SUS Mean of Study 3 is S_{SUS3}

and ’SUS Real’ Mean is S_{SUSR}. Then the alternative is $H_1 : S_{SUS3} \neq S_{SUSR}$, where there is a significant difference between the two studies.

A two-tailed independent t-test was used to compare Study 3 SUS Mean and Usoh et al. ‘SUS Virtual’ Mean. As seen in Table 3 this comparison failed to reject the null hypothesis, and therefore the SUS Mean results of this study have been validated as they do not differ from the reality scores reported by Usoh et al.

For RQ1 the researchers assume the null hypothesis $H_0 : S_{SUS3} = S_{SUSV}$ Where the SUS Mean of Study 3 is S_{SUS3} and ’SUS Virtual’ Mean is S_{SUSV} . Then the alternative is $H_1 : S_{SUS3} \neq S_{SUSV}$,where there is a significant difference between the two studies. A two-tailed independent t-test was used to compare Study 3 SUS Mean and Usoh et al. ‘SUS Virtual’ Mean. As seen in Table 3 this comparison yields an insignificant result. The null hypothesis cannot be rejected, this comparison again failed to reject the null hypothesis, meaning that the presence of a virtual mirror does not influence SUS scores.

Study 3: Summary Statistics

Table 3. Summary Statistics & Internal Comparisons for Study 3 (Session 1 n = 35, Session 2 n = 35)

	UEQ Scale	Mean (SD)			
UEQ Scale					
	Attractiveness	1.88 (0.92)			
	Perspicuity	2.22 (0.65)			
	Efficiency	1.42 (1.04)			
	Dependability	1.76 (0.75)			
	Stimulation	1.65 (1.02)			
	Novelty	0.87 (1.04)			
		Mean (SD)		*t* test	P value
SCS					
	Baseline	3.88 (1.33)			
	Session 2	4.09 (1.05)			
	Comparison			$t_{63} = 0.71$	.47
		Mean Rank	Mann-Whitney U	*z*	P value (2-tailed)
PHQ-8					
	Baseline	44.93			
	Session 2	32.07			
	Comparison		*U* = 477.5	2.53	.01
		Mean (SD)		t-test	P value
SUS					
	Slater-Usho-Steed ’Virtual’	3.8 (1.3)			
	Virtual Mirror Intervention	4.56 (1.58)			
	Comparison			$t_{46} = 1.57$	.12
	Slater-Usho-Steed ’Real’	4.4 (1.5)			

	Virtual Mirror Intervention	4.56 (1.58)			
	Comparison			t_{46} = 0.30	.77
	Question Analysis	Mean (SD)		t-test	P value
SUS Question Breakdown					
	Slater-Usho-Steed 'Virtual' Q2	3.6 (1.3)			
	Virtual Mirror Intervention Q2	4.63 (1.48)			
	Comparison			t_{46} = 2.16	.04
	Slater-Usho-Steed 'Virtual' Q3	2.6 (1.6)			
	Virtual Mirror Intervention Q3	4.66 (1.56)			
	Comparison			t_{46} = 3.64	< .01

Cross Study Quantitative Comparisons

Significance of User Experience: UEQ

Given the UEQ data, a test for normality using the Shapiro-Wilk normality test via RStudio statistical software determined that most data sets are not of the normal distribution. Hence, a Mann-Whitney U-test was used to compare the UEQ results displayed in Table 4.

Table 4. Cross Study Comparison: UEQ (Study 1 n = 18, Study 2 n = 47, Study 3 = 34).

		Mean Rank	Mann-Whitney U	z	P value (2-tailed)
UEQ Study 1 Vs Study 2: Significance of Avatar Customization					
Attractiveness					
	Study 1	37.69			
	Study 2	31.20			
	Attractiveness Comparison		*U* = 338.5	1.24	.21
Perspicuity					
	Study 1	39.14			
	Study 2	30.65			
	Perspicuity Comparison		*U* = 312.5	1.63	.10
Efficiency					
	Study 1	37.33			

	Study 2	31.34			
	Efficiency Comparison		*U* = 345.0	1.15	.25
Dependability					
	Study 1	38.00			
	Study 2	31.90			
	Dependability Comparison		*U* = 333.0	1.33	.18
Stimulation					
	Study 1	34.86			
	Study 2	32.29			
	Stimulation Comparison		*U* = 389.50	0.49	.62
Novelty					
	Study 1	39.36			
	Study 2	30.56			
	Novelty Comparison		*U* = 308.50	1.69	.09
		Mean Rank	Mann-Whitney U	*z*	P value (2-tailed)
UEQ Study 2 Vs Study 3: Significance of Virtual Mirrors					
Attractiveness					
	Study 2	34.70			
	Study 3	49.71			
	Attractiveness Comparison		*U* = 1095.00	2.84	.01
Perspicuity					
	Study 2	34.98			
	Study 3	49.32			
	Perspicuity Comparison		*U* = 1082.00	2.72	.01
Efficiency					
	Study 2	37.88			
	Study 3	45.31			

	Efficiency Comparison		*U* = 945.50	1.41	.16
Dependability					
	Study 2	32.89			
	Study 3	52.21			
	Dependability Comparison		*U* = 1180.00	3.66	< .001
Stimulation					
	Study 2	36.33			
	Study 3	47.46			
	Stimulation Comparison		*U* = 1018.50	2.11	.04
Novelty					
	Study 2	41.28			
	Study 3	40.62			
	Novelty Comparison		*U* = 786.00	0.13	.90

Significance of low-fidelity Avatars on Self-Compassion: SCS Comparison to Halim et al.

In Study 1, comparisons are made using session 1 data, i.e. data taken after only one exposure to iVR treatment. For RQ2 the researchers assume the null hypothesis H_0: $S_0 = S_1$ Where the SCS results of Halim et al. [1] is S_0 and SCS results of Study 1 is S_1. Then the alternative is $H_1 : S_0 \neq S_1$, where there is a difference between studies. A two-tailed independent t-test was performed to used to compare the SCS results from Halim et al. [1] and Study 1. As seen in Table 5, this test revealed no significant difference. With respect to RQ2, low-fidelity avatars do not have a measurable impact on the SCS.

Table 5. Cross Study Comparison: SCS (Study 0 n = 35, Study 1 n = 20)

		Mean (SD)	*t* test	P value
SCS Study 0 Vs Study 1: Significance of low-fidelity Avatars				
	Study 0	3.07 (0.73)		
	Study 1	3.11 (1.14)		
	Session 1 Comparison		t_{53} = 0.14	.89

Cross Study Qualitative Results & Comparisons

This study reports both deductive and inductive thematic analysis and then combined codes to form themes that might provide insights into user experience between each study. Below there are two short lists of codes and corresponding excerpts derived from responses to the survey questions. The deductive analysis is using the items from the Schrepp et al. paper [49] as a framework for code referencing. Each of the item pairs form a scale map to an overarching aspect of user experience, hence use as a framework to analyze the data will produce relevant insights. The scales have been organized into the list below. Note that the UEQ has 6 scales, but this list only includes scales relevant to changes observed in the data:

- Attractiveness Codes: annoying, enjoyable, good, bad, unlikable, pleasing, unpleasant, pleasant, attractive, unattractive, friendly, unfriendly.
- Perspicuity Codes: not understandable, understandable, easy to learn, difficult to learn, complicated, easy, clear, confusing.
- Dependability Codes: unpredictable, predictable, obstructive, supportive, secure, not secure, meets expectations, does not meet expectations.
- Stimulation Codes: valuable, inferior, boring, exciting, not interesting, interesting, motivating, demotivating.
- Novelty Codes: creative, dull, inventive, conventional, usual, leading edge, conservative, innovative.

Themes have been identified from the codes above to be used in a deductive thematic analysis. The themes, codes, and corresponding excerpts may provide insight as to why differences on the UEQ scales occurred. First, a comparison of feedback from Study 1 and Study 2 is expected to reveal the impact of low-fidelity avatars and personalization. Using UEQ as a framework will give insights related to the results which may explain observed changes to the questionnaire:

Table 6. Deductive Analysis: A Cross-Study Performance

	Theme	Code	Study	Excerpts
Attractiveness				
	Engagement	Enjoyable	2 & 3	"I can design my own character, very cool", "the whole process. I quite like it, and quite creative." and "good experience, it’s my first time to use VR".
Perspicuity				
	Intuitive	Easy	All	"Simple and easy to go through" and "easy to follow instructions"
		Easy to learn	All	"easy to learn" and "The prompts to help you when trying to calm the person down."
	Experience and Reflection	Easy to learn	2	"allowed me to explore different ways that I could use to phrase my words in those situations"
		Understandable	2 & 3	"The experience taught me that reassurance

				and not talking about myself" and "each Step is easy to understand"
	Context	Difficult to learn	1	"Add some extra contexts pertinent to the user" and "Perhaps there could be a backstory as to why the avatar is upset".
		Confusing	All	"Perhaps there could be a backstory as to why the avatar is upset" and "hard to provide compassion when I didn't know what they were so upset a"
Dependability				
	Expectations of Application	Meets Expectations	All	"choice of environment and avatar and realistic crying" and "It was realistic".
	Accuracy	Not Meeting Expectations	1 & 2	"I think the degree of personalisation can be improved on", "[more avatar] emotions", and "it would be nice for the avatar to react to our dialogue"
	Uncanny Valley	Secure	2	"they're all cute"
	Therapy Expectations	Meets Expectations	3	"I think this can make me feel relaxed" and "[I] become more happy"
	Virtual Environment	Not Meeting Expectations	3	"[scenes] work in a stylised context, but in a [natural] environment, like the park, the immersion breaks".
Stimulation				
	Learning Experience	Valuable	3	"I see myself, I comfort myself, it really makes [feel] release", "I want to sit on the chair to talk with someone", and "the scenario is very

				good"
	Avatar Customisation	Interesting	3	"Personal[ising] your character [is] very interesting"
Novelty				
	Customisation	Creative	2 & 3	"the experience allowed for continuous creation and novel experiences through variations", "create your own avatar to look like you, with there being a multitude of options" and "I liked the ability to create a character that could seem like myself"
	New Experience	Innovative	All	"It was new experience" and "each scenario I engaged in was unique and it did slightly alter my interactions"

To summaries the findings in Table 4., Study 3 has outperformed on the ‘Attractiveness’ scale of the UEQ in comparison to all other studies. This could be attributed to such participant experiences pertaining to enjoyment of creating avatar characters. Study 2 saw the introduction of greater avatar penalization and it continues to get praise from the participants in Study 3. Excerpts such as; "They [scenes] work in a stylized context, but in a more real environment...the immersion breaks", provide insight about immersion. As noted previously, because the mirrors introduced in Study 3 are theorized to enhance immersion, it seems likely this may have then enhanced user experience.

For the ‘Perspicuity’ scale, Study 3 outperformed all other studies and received positive feedback from the participants, as participants expressed benefits in the form of a learning experience and understandability. Although, all studies suffered due to a lack of ’context’, which overall, could explain study 1 & 2’s underperformance on the ‘Perspicuity’ scale in comparison to the other studies. The theme ‘context’ arose as participants spoke of difficulty engaging in the scenarios provided as there was no briefing about why the avatars were upset, this made it difficult for some participants to play their roles.

The visual aesthetic of a low-fidelity avatar, if it was to cause concern for a participant and trigger the Uncanny Valley [54] would likely appear in the ‘Dependability’ scale, since dependability captures participant experience pertaining to unpredictability and insecurity. Yet participants were not deterred by the appearance of the avatar, some even finding them "cute". The theme ‘meeting expectations’ incurred comments such as; "I think this can make me feel relaxed" and "[I] become more happy". These statements imply that the participants must have been supported well enough by the system to inherit the therapeutic benefits.

Study 3 performed the most with respect to ‘Stimulation’ scale. That study included personalization but with the introduction of a mirror. Since the introduction of personalization for avatars there was interest peeked during the on-boarding and in the final scene. However, allowing participants to witness themselves embodied in their chosen avatar during the act of compassion delivery (in step 2) has brought benefits to the user experience.

With respect to novelty, all Studies received feedback which suggested that the iVR was "innovative, and unique". The impact of the introduction of personalization, which allowed for additional customization of avatar appearance, can be seen in the positive remarks concerning creativity for this scale. Curiously though, only Study 1 performed better than ’above average’. Feedback within this scale suggests that some participants felt restricted in the virtual environment, evidenced by excerpts such as "[better] if the users got a little more freedom during the simulation". The movement of the embodied avatar had disabled locomotion. However, since the embodied avatar is standing in a space with no obstacles, then this may have violated the affordability of movement, thus likely negatively impacting novelty.

Via the process of inductive thematic analysis, we identified some codes which seem to be consistent with experimental changes. Some of this data will be combined to form themes which will be found in the discussion section:

Discussion

User Experience

In this paper UX has been investigated via UEQ data because this research team is in the process of developing an iVR application suitable for a clinical population. According to Shahid et al. currently there is sparse research to link positive user experience within VR to positive therapeutic outcomes [55]. Though Shahid et al. concentrated on Social Anxiety Disorder (SAD) it relates to this paper by means of being a VR mental health application. Hence, the results of UEQ and SCS have been presented independently.

As noted, study 1 was treated as a pilot study, created to test low-fidelity avatar influence on user experience as the impact of such a radical design choice in this context had not previously been tested. In comparison to Halim et al. [1], though we identified UEQ results that are not significantly different, low-fidelity avatars still performed well against the benchmark, which appropriates the use further exploration of low-fidelity avatars in therapeutic applications. Study 2 therefore introduced avatars of the same design but with greater allowances for customization.

Following iVR theory, Study 2 tested the introduction of enhanced avatar personalization to boost user experience to support therapeutic outcomes. Hence, for RQ1, UEQ data did not differ from Study 1 with respect to questionnaire scores, as seen in Table 4. However, responses obtained from a thematic analysis of the qualitative data gave encouraging insights with relevance to ‘customization’ and ‘avatar’ experiences. The users called for greater customization features, in particular skin color, which to the researchers of this study emphasized how important such aspects are for user embodiment of virtual avatars.

‘Novelty’, remained in the ’Good’ category and produced feedback that reported a ’creative’ experience, as seen in Table 6. Hence, there is evidence of user enthusiasm credited to the added customization features uncovered in the thematic analysis. However, the insignificant UEQ score differences under comparison suggests that the inclusion of customizable low-fidelity avatars did not make the user experience worse, thus satisfying RQ1. Given the thematic analysis responses, it could be the case that UEQ scores might be significantly improved by allowing customization options that enable control over greater avatar detail.

For Study 3, UEQ scores in Figure 9 improved the most compared to all other studies within this paper as supported by the statistical comparisons of Table 4. The improvements attributed to an otherwise unseen increase in ‘attractiveness’, ‘perspicuity’, ‘dependability’ and ‘stimulation’. From the qualitative data, the rise in these scales could be explained by themes surrounding enjoyment and expectation as the participants seemed comfortable with the inclusion of mirrors.

But why did we see this spike in UEQ results? According to Schrepp et al. "Attractiveness is a pure valence dimension" [49] that is representative of the users general experience. Hence significantly positive performance here in comparison to the other studies suggest that the inclusion of virtual mirrors elevated the overall performance of the iVR experience. This jump in ‘attractiveness’ is likely due to the significant improvements to ‘perspicuity’ and ‘dependability’ which represent pragmatic quality aspects of the iVR treatment, but also ‘stimulation’; a hedonic quality which pertains to ’motivation’ of use.

Improvement to ‘perspicuity’ suggests that the iVR therapy was both easy to learn and "easy to get familiar with the product" [49]. ‘Dependability’ collects information pertaining to a user’s sense of control of interactions, prediction of system behavior and safety with the product. For this iVR experience, since the users spent a considerable amount of effort on customizing and interacting with low-fidelity avatars, it can be said that for the user the avatar experience is the ’product’. Though the overall intention is self-compassion therapy, a user extracts this therapy through the medium of human-avatar interaction. Interestingly, it took the introduction of virtual mirrors to ’unlock’ encouraging UEQ results.

This ’unlocking’ might be explained by the SUS results of this study. As seen in Table 3, comparisons of Questions 2 and 3 revealed significant differences, and these questions pertain to not only the ’realness’ of the experience but also how the user recalls the experience. Question 3 judges the perception of environment from a user’s memory of the experience, if the memory is recalled as "as somewhere that you visited?" [50] then the user’s sense of immersion and presence must have been high in the virtual environment because the user recalls the memory as though it actually happened ’somewhere’. This accompanied by significantly positive Question 2 results which pertain to a sense of realness, i.e. ’being there’, offer a potential explanation to the UEQ results of Study 3.

Given the results of the SUS questionnaire, it appears that the introduction of virtual mirrors has impacted the users virtual presence, such that it has boosted UEQ results. However this paper is an effort to improve the original iVR application [1]. Introduction of low-fidelity avatars and greater customization had been systematically introduced with no positive impact until the introduction of virtual mirrors. But is this outcome the product of all introduced experimental changes combined or could we just have the mirrors by themselves? Given that UEQ have had no significant changes prior to the introduction of mirrors then it could be inferred that the virtual mirrors is the source of meaningful change with respect to improvement of presence. However, it is yet to be seen if introduction of low-fidelity avatars and greater customization are irrelevant. With respect to low-fidelity avatars, which are alternatives to realistic avatars, this study may have provided some preliminary steps to demonstrate that the avoidance use of ’excessive realism’ can be achieved without significant damage to UX or incurring the uncanny valley, however further investigation is required.

Table 7. User Experience: Inductive Thematic Analysis

Theme	Code	Excerpts
Compassion	Therapeutic Benefit	Some participants mentioned that they liked listening to their own voice as it was played back to them via their avatar; "Listening to myself was an interesting experience but it worked" and, "it was nice being able to hear yourself comfort someone else, and I felt I benefited from that".
Avatar Individualisation	Customisation	Study 2 and 3 included a feature which allowed for customisation of the avatar, hence participants stated "I liked the ability to create a character like myself" but also stated "customising it to look like me did make the difference of how I perceived the audio replay"
Dialogue Context	Interactions	The participants made requests for avatar responses to help guide their experience; "have some kind of method for response, it's difficult to comfort a character you can't interact with".
Avatar Improvement	Avatars	Considering statements about avatar behaviour and appearance, the participants have been quoted as saying "I think the degree of personalisation can be improved" and "options for avatar styles could have more variations (e.g hairstyle, colours, dress types)" and, "there wasn't a skin colour that matched".
System Issues	Performance	This theme considered matters such as graphics, audio and game mechanics which includes feedback such as; "The graphics could be improved so that it feels less computer-like.", "More [NPC] actions can be implemented and audio can be synchronized with the body language of the avatar", and "there were some bugs in the current system, minor, but if that was resolved I reckon the entire user experience would be much smoother".
Self Recognition	Mirror	This theme considered codes such as realisation and influence the participants expressed; "aware [of] what I'm doing [in] that scene." and "[I] see it".
Plausibility	Mirror	The mirrors tend to have a bad influence if the placement of the mirror is not consistent with reality. A code named 'weird in the park' emerged from the excerpts where participants complained "it's a bit weird the look at the mirror in the park".
Compassion	Therapeutic Benefit	Some participants mentioned that they liked listening to their own voice as it was played back to them via their avatar; "Listening to myself was an interesting experience but it

		worked" and, "it was nice being able to hear yourself comfort someone else, and I felt I benefited from that".
Avatar Individualisation	Customisation	Study 2 and 3 included a feature which allowed for customisation of the avatar, hence participants stated "I liked the ability to create a character like myself" but also stated "customising it to look like me did make the difference of how I perceived the audio replay"
Dialogue Context	Interactions	The participants made requests for avatar responses to help guide their experience; "have some kind of method for response, it's difficult to comfort a character you can't interact with".
Avatar Improvement	Avatars	Considering statements about avatar behaviour and appearance, the participants have been quoted as saying "I think the degree of personalisation can be improved" and "options for avatar styles could have more variations (e.g hairstyle, colours, dress types)" and, "there wasn't a skin colour that matched".
System Issues	Performance	This theme considered matters such as graphics, audio and game mechanics which includes feedback such as; "The graphics could be improved so that it feels less computer-like.", "More [NPC] actions can be implemented and audio can be synchronized with the body language of the avatar", and "there were some bugs in the current system, minor, but if that was resolved I reckon the entire user experience would be much smoother".
Self Recognition	Mirror	This theme considered codes such as realisation and influence the participants expressed; "aware [of] what I'm doing [in] that scene." and "[I] see it".
Plausibility	Mirror	The mirrors tend to have a bad influence if the placement of the mirror is not consistent with reality. A code named 'weird in the park' emerged from the excerpts where participants complained "it's a bit weird the look at the mirror in the park".
Compassion	Therapeutic Benefit	Some participants mentioned that they liked listening to their own voice as it was played back to them via their avatar; "Listening to myself was an interesting experience but it worked" and, "it was nice being able to hear yourself comfort someone else, and I felt I benefited from that".
Avatar Individualisation	Customisation	Study 2 and 3 included a feature which allowed for customisation of the avatar, hence

		participants stated "I liked the ability to create a character like myself" but also stated "customising it to look like me did make the difference of how I perceived the audio replay"
Dialogue Context	Interactions	The participants made requests for avatar responses to help guide their experience; "have some kind of method for response, it's difficult to comfort a character you can't interact with".

As for customization of low-fidelity avatars, though the UEQ produced no significant results in Study 2, inclusion of such features has received encouraging feedback in the qualitative analysis, with participants expressing enthusiasm for this feature and requesting yet greater control over avatar design, as seen in Table 7. Prior work on iVR has shown the benefits of individualizing virtual experiences [3, 56], however the results of Study 2 and 3 suggest that with customization of avatar appearance, such tools for avatar design must be accompanied with visual reminders to reinforce user embodiment in the virtual environment. "customizing it to look like me did make the difference of how I perceived the audio replay" [Participant], this statement implies the importance of the user being able to view themselves whilst participating in the virtual therapy, tying together avatar appearance and virtual mirrors.

It's interesting to note, that unlike other studies, the placement of mirrors was subtle and not referred to, i.e. users were not explicitly directed to interact with the mirrors in their environment. Yet, virtual mirrors seemed to have had a significant impact on the experience regardless. There is a caveat though, mirror placement had to confirm with expectations of the real world, as seemingly incorrect or improbable mirror placement was noticed by the user, and when this occurred, had negative consequences for user immersion as it violated expectations. Participants reported that mirror misplacement in the virtual environment risked loss of immersion or eeriness, see Table 7. Regardless, introduction of virtual mirrors has not negatively impacted UX, satisfying RQ1.

Therapeutic Outcomes

Following Study 1, the qualitative data found in the Thematic Analysis tables, called for more 'options' and a greater degree of personalization such as; hairstyle, colors, dress types, etc. Since the iVR therapy being replicated already had some allowances for customization with justifications from previous studies, this seemed a plausible strategy in the interest of improving UX and SCS. Subsequently for RQ2, this paper produced a significant result for the internal comparison of Study 2, which involved two therapy exposures with low-fidelity avatars and greater customization, making possible the argument that inclusion of these design changes alone has merit, unfortunately this trend did not continue into Study 3 despite a dramatic improvement to UX.

However, the researchers of this paper believe that null results for SCS is in itself meaningful. It can be stated that the therapeutic performance, which was already at an acceptable standard [1], has not been negatively impacted by the addition of low-fidelity avatars, thus satisfying RQ2. There is no evidence to suggest that inclusion of low-fidelity avatars, greater customization and virtual mirrors had a negative impact on therapeutic outcomes. Furthermore, in the case of Study 2, there is evidence to suggest it improved iVR self-compassion therapy.

Uncanny Valley & Cognitive Load

The thematic analysis of this paper enabled us to investigate UX insights beyond that afforded by questionnaires and informed design justifications between each study in the pursuit of optimal therapeutic benefit. The qualitative data has been applied to the UEQ via deductive analysis, using the scale descriptions as a framework to find excerpts which might support an understanding of variance across the studies. Inductive codes were also developed, the combination of these processes gave rise to the theme of customization, which revealed participant acceptance of the avatars.
From Table 7 the theme of ‘Avatar Individualization’ contained excerpts such as; "I liked [to] changing my avatar, it looked like me" and “Even though the avatar is animated and doesn’t exactly look like me with the provided options available, it made the experience feel more relatable” For Study 2 and 3, the authors found that body ownership was experienced by the participants of this study, despite the appearance of the avatar as theorized. The researchers of this paper found it encouraging to discover in the qualitative data that the appearance of the newly introduced avatars was generally well received as such unprovoked quotes emerged from the participant feedback “While I do suggest more realistic environments, I don’t recommend realistic avatars as much. I think the more cartoonish appeal is better”.

Masahiro Mori’s Uncanny Valley (UV) model (Figure 10.) suggests that entities which encroach on human-likeness may invoke ‘eeriness’ in the human observer [54]. Such an entity can be described as “Something that has come to appear very close to human — like a robot, puppet, or prosthetic hand — could easily tumble down into the uncanny valley.” [54]. This paper, though a study into improving iVR therapy, has consequently also made some progress in the investigation of human-avatar interaction via the use of low-fidelity avatars. The rationale behind the choice of low-fidelity avatar design was to avoid cognitive overload, but demonstration of uncanny valley avoidance might be an additional bonus.

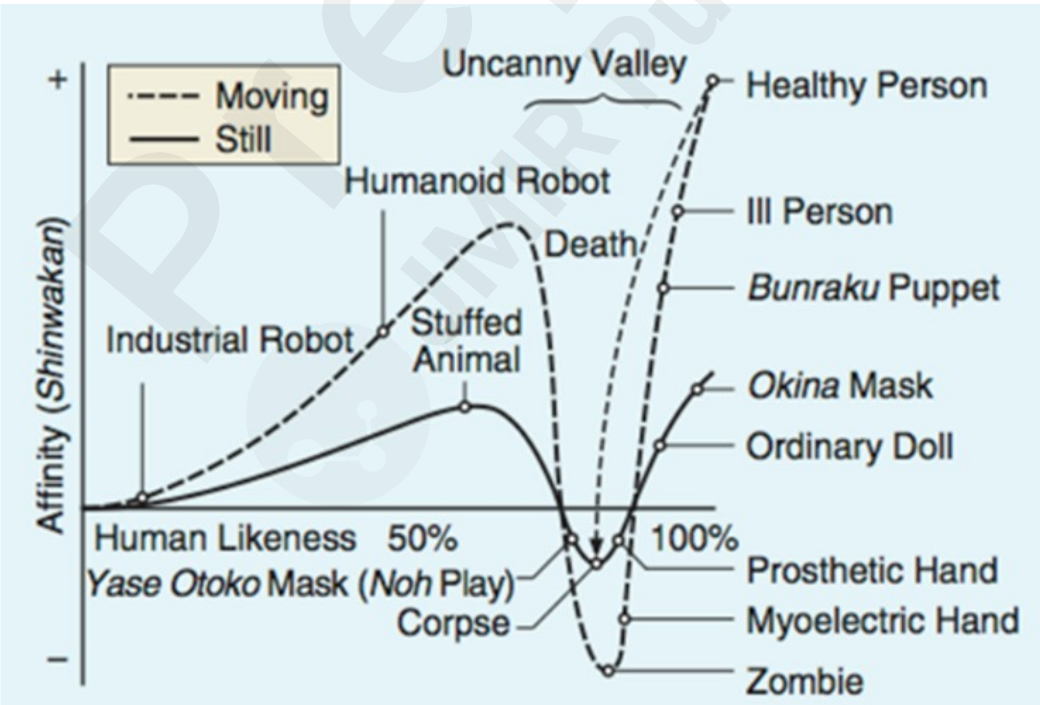


Figure 10. Hypothesized uncanny valley model.

“I recommend that designers instead take the first peak as their goal, which results in a moderate degree of human likeness and a considerable sense of affinity.” [54]. The choice of low-fidelity avatars aligns with this suggestion. In retrospect, Halim et al. [1] underperformed on the

‘Dependability’ scale of the UEQ, a scale where predictability, control, and safety of a participant’s experience are measured. Intrusions of ‘eeriness’ would likely violate responses sensitive to the dependability of the system, since such a UV encounter would undermine the experience. It’s possible that performance on the UEQ’s of this study occurred because the low-fidelity avatar appearance fall into the valley as seen in Figure 10.

Interestingly, it’s apparent that virtual mirrors have provided the most significant results for increased ‘Dependability’, hence with respect to securing against UV interference, virtual mirrors must play a role in addition to deliberately aiming for avatars in the realm of ’Humanoid Robot’, seen in Figure 10. As placed in the virtual scenes, the mirrors reflect the user’s conversations with the virtual confederate, the mirrors facilitate a subtle view of the users’ interactions within the virtual environment likely creating (consciously or subconsciously) an awareness of the current social situation and task (i.e., conveying empathy). UV research was not the intention of this paper, this discussion is provided merely as a reflection for the benefit of broader UX considerations.

However, avatars are not the only factor that can disrupt a user’s experience, there are virtual environment considerations as well. Noted in the thematic analysis of Study 3 which experimented with mirrors, some participants found it odd that a mirror would exist in a naturalistic environment, such as the park environment of this iVR. "[mirror] could be weird [to see] in the park environment". Violations of physics can also harm the user experience, e.g., "I see the mirror and then walk through it. I’m [a] ghost and that makes me a little bit scar[ed]". Though UX still performed well, it’s important to take such feedback into consideration in future design iterations.

This paper has shown that mitigating the risk of UV encounters via the selection of low-fidelity avatars could support therapeutic outcomes and good user experience, as supported by Studies 2 and 3 respectively. Reaching for ever more sophisticated avatars as a solution to achieving more successful human-avatar interactions in VR ignores the Sisyphus-like limitation that avatars, no matter how advanced, virtual avatars may never imitate human expression perfectly and therefore will always risk falling back into the uncanny valley [57]. Hence, the deliberate selection of ’Humanoid Robot’ caliber avatars may set user expectations such that there is minimized risk of negative UV encounters. UX-enhancing techniques, such as individualization and virtual mirrors, seem necessary in support of such avatars to ensure that application objectives can still be met.

Limitations

This study relied on SCS, which is a validated survey that measures self-compassion; however, there is a discussion of its limitations which must be acknowledged [58]. With respect to participants, for ethical reasons this study was restricted to trials on a healthy population. Therefore, this study is a preliminary step in a larger design process. Ideally the next iteration of participants will be from a clinical population for co-design activities such that iVR technology can be tested with its intended end-users. Studies 1 & 3 within this paper where underpowered, the researchers of this paper compensated by performing qualitative analysis to support insights.

Conclusion

If compassion means ‘To suffer together’ [59], then does engaging in self-compassion have to mean suffering alone? Within this iVR intervention participants were comforted by their own custom-made low-fidelity avatar. With these avatars the participants’ experienced self-compassion in a virtual environment which suited the participants’ preferences and included subtle virtual mirror placement. However, little is known about human-avatar interaction in this domain of VR-administered treatment of depression.

Prior literature reveals that high-quality avatars tax cognitive load and, since depression often compromises cognitive capacity, research into the suitability of such avatars in this domain was needed. The three studies reported here provide novel insights into the potential value of low-fidelity avatars in the context of iVR application in self-compassion for mental health therapy. The final investigation in the sequence revealed that, for a healthy population, UX and SCS therapeutic benefits are not reduced by the introduction of low-fidelity avatars, greater avatar customization, or virtual mirrors.

Demonstrating feasibility on a healthy population takes this research team closer to an appropriate iVR application for people with depression, and future iterations should seek to include a clinical population. Another insight of this study is a demonstration that avatar customization should be supported by virtual mirrors to promote embodiment for improved UX. This paper builds upon research into the benefits of individualizing virtual reality experience for users in the context of iVR self-compassion therapy [1, 3, 56, 60].

Studies 2 and 3, provided evidence which aligns well with iVR research. However, we have gone a step further in iVR by experimenting with the feasibility of developing low-fidelity avatars for VR mental health applications. We believe our contribution will lay the foundation for large-scale efficacy testing, clinical implementation, and cost-effective delivery of VR solutions for mental health therapy in the future. The next iteration intends to include participants from a clinical population for co-design activities such that this iVR therapy can be tested by the intended end-users.

Acknowledgments

The authors acknowledge Swaraj Vishwas Randhir and Vibhav Chitale for their contribution to the development of the iVR environment and provision of previous virtual environment. The authors would also like to acknowledge Lachlan Greig for providing consultation on psychological survey analysis and Chaitanya Dasi Luanzon for providing imagery.

Conflicts of Interest

None declared.

Abbreviations

1PP: First Person Perspective
CFT: compassion-focused therapy
HMD: Head Mounted Display
iVR: Individualized Virtual Reality
NPC: Non-Player Character
PHQ-8: 8-item Patient Health Questionnaire
RQ: Research Question
SCS: Self-Compassion Scale
SUS: Slater, Usoh, and Steed (Presence Questionnaire)
UEQ: User Experience Questionnaire
UQ: University of Queensland
UV: Uncanny Valley
UX: User Experience
VR: virtual reality

References

1. Halim I, Stemmet L, Hach S, et al. Individualized Virtual Reality for Increasing Self-Compassion: Evaluation Study. JMIR Mental Health. 2023;10:e47617.
2. Falconer CJ, Slater M, Rovira A, et al. Embodying compassion: a virtual reality paradigm for overcoming excessive self-criticism. PloS one. 2014;9(11):e111933.
3. Baghaei N, Lehan S, Khaliq I, et al. Designing Individualised Virtual Reality Applications for Supporting Depression: A Feasibility Study. Companion of the 2021 ACM SIGCHI Symposium on Engineering Interactive Computing Systems. 2021:6-11.
4. Waltemate T, Gall D, Roth D, Botsch M, Latoschik ME. The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. IEEE transactions on visualization and computer graphics. 2018;24(4):1643-1652.
5. Gonzalez-Franco M, Perez-Marcos D, Spanlang B, Slater M. The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment. in 2010 IEEE virtual reality conference (VR):111–114IEEE 2010.
6. Slater M, Neyret S, Johnston T, et al. An experimental study of a virtual reality counselling paradigm using embodied self-dialogue. Sci Rep. 2019;9(1):10903.
7. Falconer CJ, Rovira A, King JA, et al. Embodying self-compassion within virtual reality and its effects on patients with depression. BJPsych Open. 2016;2(1):74-80.
8. Kilteni K, Groten R, Slater M. The sense of embodiment in virtual reality. Presence: Teleoperators and Virtual Environments. 2012;21(4):373–387.
9. Slater M, Spanlang B, Sanchez-Vives MV, Blanke O. First person experience of body transfer in virtual reality. PLoS One. 2010;5(5):e10564.
10. Blanke O, Metzinger T. Full-body illusions and minimal phenomenal selfhood. Trends in cognitive sciences. 2009;13(1):7–13.
11. Banakou D, Slater M. Embodiment in a virtual body that speaks produces agency over the speaking but does not necessarily influence subsequent real speaking. Sci Rep. 2017;7(1):14227.
12. Slater M. A note on presence terminology. Presence connect. 2003;3(3):1-5.
13. Slater M, Usoh M, Steed A. Depth of presence in virtual environments. Presence: Teleoperators & Virtual Environments. 1994;3(2):130-144.
14. Grassini S, Laumann K, Rasmussen Skogstad M. The Use of Virtual Reality Alone Does Not Promote Training Performance (but Sense of Presence Does). Front Psychol. 2020;11:1743.
15. Wirth W, Hartmann T, Böcking S, et al. A process model of the formation of spatial presence experiences. Media psychology. 2007;9(3):493–525.
16. Sanchez-Vives MV, Slater M. From presence to consciousness through virtual reality. Nature Reviews Neuroscience. 2005;6(4):332-339.
17. Wu B, Yu X, Gu X. Effectiveness of immersive virtual reality using head-mounted displays on learning performance: A meta-analysis. British Journal of Educational Technology. 2020;51(6):1991-2005.
18. Tajadura-Jiménez A, Banakou D, Bianchi-Berthouze N, Slater M. Embodiment in a child-like talking virtual body influences object size perception, self-identification, and subsequent real speaking. Scientific reports. 2017;7(1):9637.
19. Sanchez-Vives MV, Spanlang B, Frisoli A, Bergamasco M, Slater M. Virtual hand illusion induced by visuomotor correlations. PloS one. 2010;5(4):e10381.
20. Tsakiris M, Haggard P. The rubber hand illusion revisited: visuotactile integration and self-attribution.. Journal of experimental psychology: Human perception and performance. 2005;31(1):80.
21. Inoue Y, Kitazaki M. Virtual mirror and beyond: The psychological basis for avatar embodiment via a mirror. Journal of Robotics and Mechatronics. 2021;33(5):1004–1012.

22. Banakou D, Groten R, Slater M. Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. Proceedings of the National Academy of Sciences. 2013;110(31):12846–12851.
23. McArthur V. The UX of Avatar Customization. in Proceedings of the 2017 CHI Conference on Human Factors in Computing SystemsCHI ’17(New York, NY, USA):5029–5033Association for Computing Machinery 2017.
24. Fitton I, Clarke C, Dalton J, Proulx MJ, Lutteroth C. Dancing with the Avatars: Minimal Avatar Customisation Enhances Learning in a Psychomotor Task. in Proceedings of the 2023 CHI Conference on Human Factors in Computing SystemsCHI ’23(New York, NY, USA)Association for Computing Machinery 2023.
25. Buck L, Young GW, McDonnell R. Avatar Customization, Personality, and the Perception of Work Group Inclusion in Immersive Virtual Reality. in Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social ComputingCSCW ’23 Companion(New York, NY, USA):27–32Association for Computing Machinery 2023.
26. Birk MV, Mandryk RL. Combating Attrition in Digital Self-Improvement Programs using Avatar Customization. in Proceedings of the 2018 CHI Conference on Human Factors in Computing SystemsCHI ’18(New York, NY, USA):1–15Association for Computing Machinery 2018.
27. Du Q, Song Z, Jiang H, Wei X, Weng D, Fan M. LightSword: A Customized Virtual Reality Exergame for Long-Term Cognitive Inhibition Training in Older Adults. in Proceedings of the CHI Conference on Human Factors in Computing SystemsCHI ’24(New York, NY, USA)Association for Computing Machinery 2024.
28. Cuthbert R, Turkay S, Brown R. The Effects of Customisation on Player Experiences and Motivation in a Virtual Reality Game. in Proceedings of the 31st Australian Conference on Human-Computer-InteractionOzCHI ’19(New York, NY, USA):221–232Association for Computing Machinery 2020.
29. Kao D, Ratan R, Mousas C, Joshi A, Melcer EF. Audio Matters Too: How Audial Avatar Customization Enhances Visual Avatar Customization. in Proceedings of the 2022 CHI Conference on Human Factors in Computing SystemsCHI ’22(New York, NY, USA)Association for Computing Machinery 2022.
30. Altarteer S, Vassilis C, Harrison D, Chan W. Product customisation: virtual reality and new opportunities for luxury brands online trading. in Proceedings of the 21st International Conference on Web3D TechnologyWeb3D ’16(New York, NY, USA):173–174Association for Computing Machinery 2016.
31. Freeman G, Zamanifard S, Maloney D, Adkins A. My Body, My Avatar: How People Perceive Their Avatars in Social Virtual Reality. in Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing SystemsCHI EA ’20(New York, NY, USA):1–8Association for Computing Machinery 2020.
32. Wilson AC, Mackintosh K, Power K, Chan SW. Effectiveness of self-compassion related therapies: A systematic review and meta-analysis. Mindfulness. 2019;10:979-995.
33. Germer CK, Neff KD. Self-compassion in clinical practice. Journal of clinical psychology. 2013;69(8):856-867.
34. Kurebayashi Y, Sugimoto H. Self-compassion and related factors in severe mental illness: A scoping review.. Perspectives in Psychiatric Care. 2022;58(4).
35. Döllinger N, Mal D, Keppler S, et al. Virtual Body Swapping: A VR-Based Approach to Embodied Third-Person Self-Processing in Mind-Body Therapy. in Proceedings of the CHI Conference on Human Factors in Computing Systems:1–18 2024.
36. Yamashita Y, Yamamoto T. Effect of virtual reality self-counseling with the intimate other avatar. Scientific Reports. 2024;14(1):15417.
37. Kim DY, Lee HK, Chung K. Avatar-mediated experience in the metaverse: The impact of

avatar realism on user-avatar relationship. Journal of Retailing and Consumer Services. 2023.
38. Rock PL, Roiser JP, Riedel WJ, Blackwell AD. Cognitive impairment in depression: a systematic review and meta-analysis. Psychological Medicine. 2014;44(10):2029–2040.
39. Gruber M, Mauritz M, Meinert S, et al. Cognitive performance and brain structural connectome alterations in major depressive disorder. Psychological Medicine. 2023;53(14):6611–6622.
40. Nikolin S, Tan YY, Schwaab A, Moffa A, Loo CK, Martin D. An investigation of working memory deficits in depression using the n-back task: A systematic review and meta-analysis. Journal of Affective Disorders. 2021;284:1–8.
41. Wang XL, Du MY, Chen TL, et al. Neural correlates during working memory processing in major depressive disorder. Progress in neuro- psychopharmacology and biological psychiatry. 2015;56:101–108.
42. Albus P, Seufert T. The modality effect reverses in a virtual reality learning environment and influences cognitive load. Instructional Science. 2023;51(4):545–570.
43. Han J, Zheng Q, Ding Y. Lost in virtual reality? Cognitive load in high immersive VR environments. Journal of Advances in Information Technology. 2021;12(4).
44. Mayer RE. Multimedia learning. in Psychology of learning and motivation;41:85–139Elsevier 2002.
45. Marraffino MD, Johnson CI, Garibaldi AE. Virtual reality is better than desktop for training a spatial knowledge task, but not for everyone. In International Conference on Human-Computer Interaction:212–223Springer 2022.
46. Freeman D, Reeve S, Robinson A, et al. Virtual reality in the assessment, understanding, and treatment of mental health disorders. Psychol Med. 2017;47(14):2393-2400.
47. Neff KD. The self-compassion scale is a valid and theoretically coherent measure of self-compassion. Mindfulness. 2016;7:264-274.
48. Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. Journal of affective disorders. 2009;114(1-3):163-173.
49. Schrepp M, Hinderks A, Thomaschewski J. Construction of a Benchmark for the User Experience Questionnaire (UEQ). International Journal of Interactive Multimedia and Artificial Intelligence. 2017;4(4).
50. Usoh M, Catena E, Arman S, Slater M. Using presence questionnaires in reality. Presence. 2000;9(5):497–503.
51. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative research in psychology. 2006;3(2):77-101.
52. Stylized Customizable Avatars. URL: https://sunbox.games/asset/stylized-customizable-avatars/ [accessed 2024-12-20]
53. Neff KD, Tóth-Király I. Self-compassion scale (SCS). in Handbook of assessment in mindfulness research:1–22Springer 2022.
54. Mori M. The uncanny valley: the original essay by Masahiro Mori. IEEE Spectrum. 1970.
55. Shahid S, Kelson J, Saliba A, others . Effectiveness and User Experience of Virtual Reality for Social Anxiety Disorder: Systematic Review. JMIR Mental Health. 2024;11(1):e48916.
56. Baghaei N, Lehan S, Hlasnik A, et al. Time to Get Personal: Individualised Virtual Reality for Mental Health. in Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing SystemsCHI EA ’20(New York, NY, USA):1–9Association for Computing Machinery 2020.
57. Elliott TC, Henry JD, Baghaei N. Designing Humanoid Avatars in Individualised Virtual Reality for Mental Health Applications. in 2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct):315-321 2023.

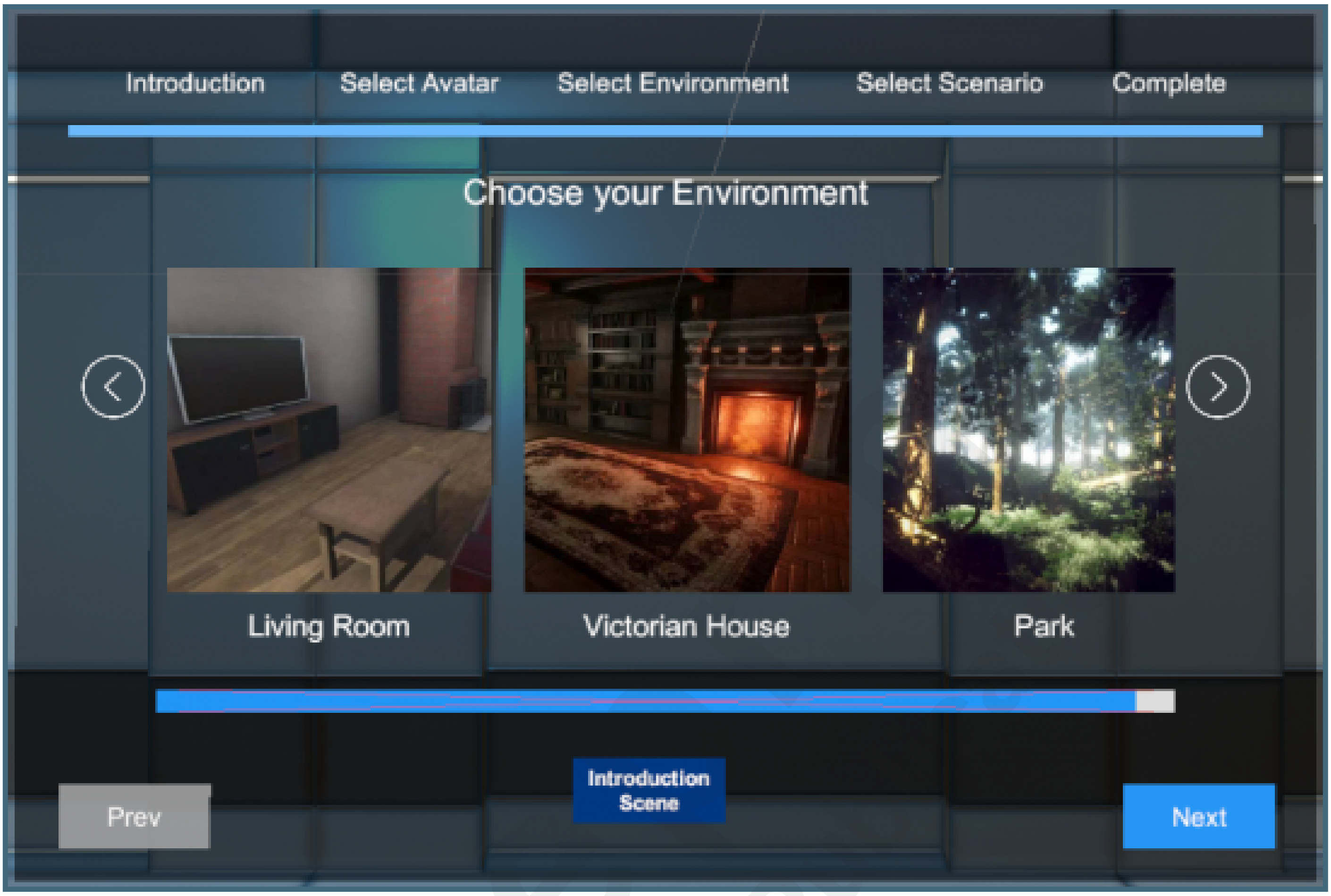
58. Muris P, Otgaar H. The process of science: A critical evaluation of more than 15 years of research on self-compassion with the Self-Compassion Scale. Mindfulness. 2020;11:1469-1482.
59. Lilius JM, Kanov J, Dutton JE, Worline MC, Maitlis S. Compassion revealed: What we know about compassion at work (and where we need to know more). 2011.
60. Baghaei N, Ahmadi A, Khaliq I, Liang HN. Individualised virtual reality for supporting depression: Feedback from mental health professionals. In 2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct):63-67IEEE.

Supplementary Files

Figures

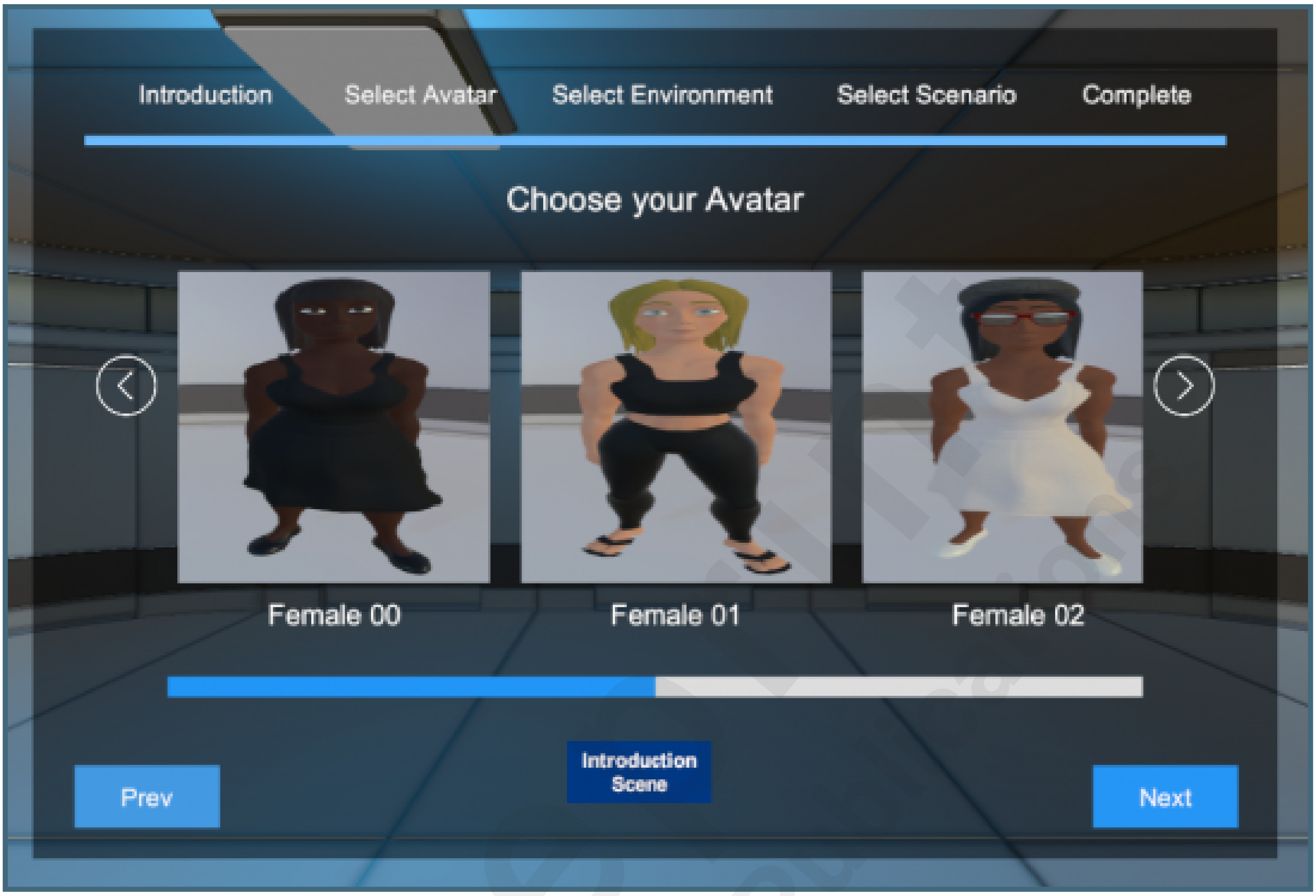
A participant viewing their customized low-fidelity avatar through a virtual mirror in the VR Self-Compassion Therapy application.

Individualization of virtual therapy via selection of preferred therapeutic environment.

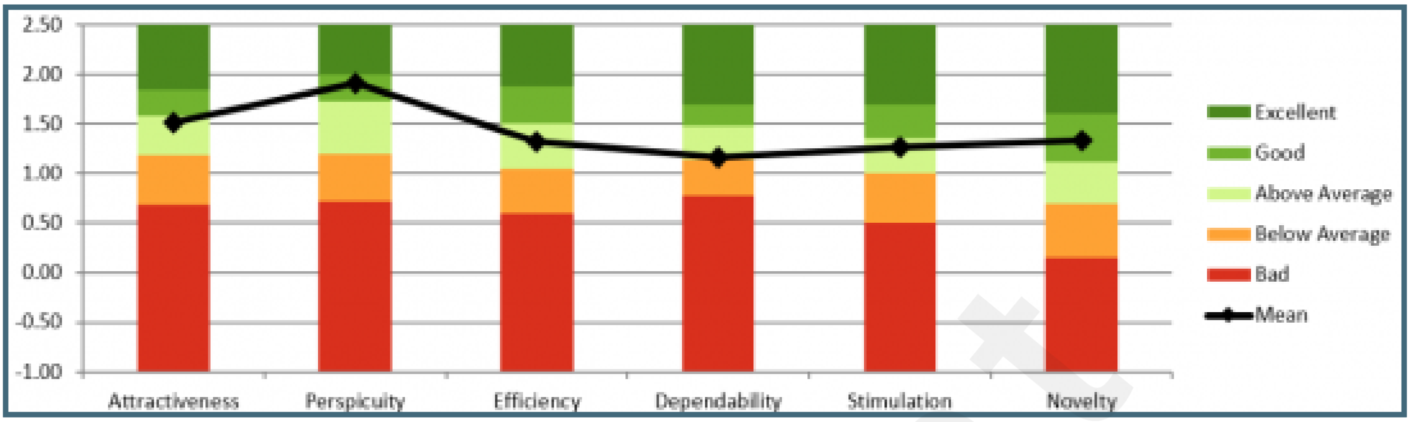


Low-fidelity (left) versus humanoid (right) avatar appearance.

Regressed customization of low-fidelity Avatar, i.e. limited choices for user.

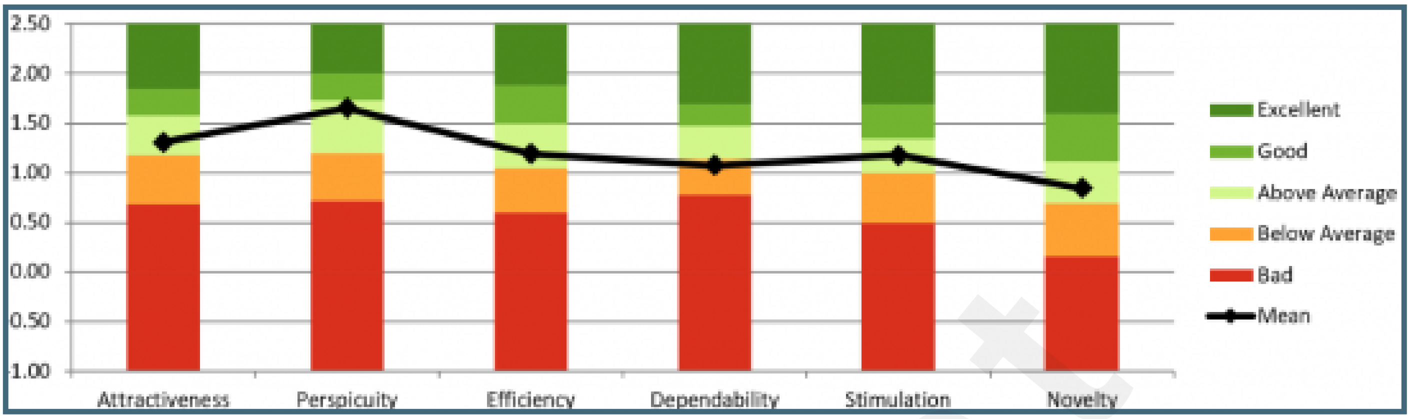


UEQ Study 1. Introduction of low-fidelity avatars.



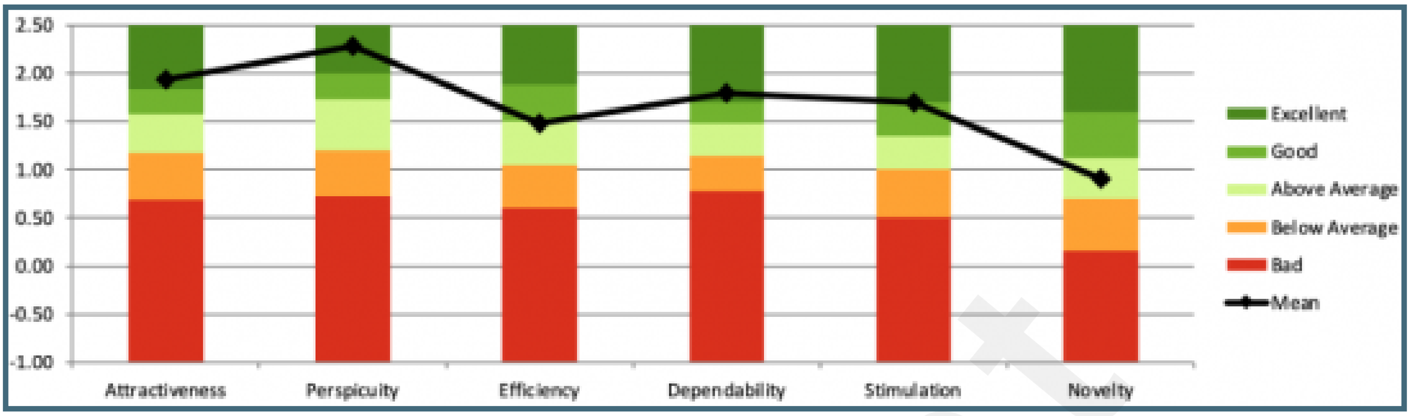
Enhanced customization of low-fidelity avatar.

UEQ Study 2. Introduction of avatar customization.



Mirror positioned for the participant (blue shirt) can see themselves interact with the upset low-fidelity avatar (red dress).

UEQ Study 3. Introduction of virtual mirrors.



Hypothesized uncanny valley model.