Facial Expressions on Virtual Humans: Optimising Virtual Reality Applications for Mental Health

Dissertation
submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy
in Psychiatry

by

Shu Wei

Department of Psychiatry | Magdalen College
University of the Oxford

DECLARATION

The work reported in this thesis was completed between October 2020 and September 2024, supervised by Dr Aitor Rovira and Professor Daniel Freeman. The work in this thesis is my own.

The pronoun 'we' is used throughout the thesis to acknowledge the research I conducted under the supervision of my advisors. My contributions span all aspects of the studies, including the design, creation of 3D animations, programming of the VR experience, data collection, statistical analysis, and publication of results. My advisors provided guidance and oversight throughout these stages.

I declare that the work presented has not been submitted for any other degree, in this or any other university or learning institute.

ChatGPT 4.0[1] was used to proofread some original paragraphs, with selected suggestions adopted after careful review.

[1] https://chat.openai.com/

ACKNOWLEDGEMENTS

First, to my supervisors Aitor and Daniel. Thank you for your generous support, patience, and the bold bravery in taking on someone from a completely different industry for this DPhil journey. I have learnt a lot about research and beyond. I hope this is just the beginning of a path toward creating positive facial expressions in real-world scenarios. And to Aitor, I hope that having me as your first DPhil student doesn't make me your last.

To the wider OCAP team, thank you for having me. Tackling C# bugs and animation issues wasn't always the most enjoyable task, but the Dev team—Andre, Ruppert, Matt, and Jack—made it much less painful. To Ginny, my studies wouldn't have been possible without your help in navigating all the administrative and logistical hurdles. And to Alice, Memoona, Felicity and the many future and current clinical psychologists who I have been grateful to work alongside, I hope we can work together in the future, even if not side by side.

I was lucky to have a 3-month research internship with CWI Amsterdam, and I'd like to thank Abdo and Pablo for making that happen. It was a compact and valuable experience, during which I learnt about bio-signal processing and archery targeting. I must also extend my gratitude to all my DIS friends, who made me feel there was no need to study Dutch to prepare for my life in the Netherlands.

Oxford becomes oxford for me because of the people and memories. There are too many personal Thank you to sum up—to Lina, for helping with the cross-city move; to all my Mathematician friends, for making a DPhil start during COVID more fun than it should have been; to Tom and his golden retriever Wilson, for the chicken nuggets and furry support; to Junhee, for all the Zoom chats and self-organised hackathons; to Linda and Finaly, for the therapeutic trips and hilarious punts; to Nikita and Rishi, for the refreshing 'home visit' in London; and to Desmond, for sharing all my ups and downs, and excessive Asian food choices.

A special thank you to Lin, my write-up buddy for the thesis sprint. I've truly enjoyed our write-and-drink sessions over the past weekends, even though we always ended up starting the drinking a bit too early. Also, a quirky thank you to the restaurant Zhangji, without any commercial partnership, but purely appreciation.

I must also extend my gratitude to myself, my parents, OCAP, the Oxford China Scholarship Fund, Magdalen College, and the Rabin Ezra Scholarship, for the collaborative support of my DPhil funding.

Finally, 感谢我的父母与家人。It doesn't matter if you understand all the words and abbreviations in this thesis—you are the ones who made this entire journey possible.

THESIS ABSTRACT

Virtual reality (VR) is increasingly being used to assess, understand, and treat mental health difficulties. A key aspect within many of these applications is virtual humans (VHs), for example as guides through the programme or as stimuli to elicit the mental health difficulties. However, there has been very little study on the impact of VHs and their characteristics in the use of VR for mental health. The aim of this thesis is to investigate the potential of the VH design, especially with an emphasis on their facial expressions, to enhance the success of VR mental health applications. The research is guided by three primary objectives. First, to understand how VHs have been used in VR mental health applications. Second, to evaluate in randomised controlled studies the effects of the detailed design of VHs. Third, to examine not only participants' subjective experiences (i.e. perceptions of the VHs) but also their physiological and behavioural responses.

Chapter 1 introduces the concepts around VR mental health research, VHs, and their facial expressions. In Chapter 2, a systematic review was conducted using the key search terms related to VR, VHs and mental health, which identified 2494 papers. Seventy-three research studies on the use of VHs in VR mental health research were analysed. The review revealed that while VHs play diverse roles in mental health applications, very few studies explicitly tested the characteristics of VHs, and the research generally did not report sufficient details about how they were designed. To address this gap, two experimental studies were conducted, involving the development and testing of two types of VHs: a VR coach within a therapy programme and the characters used as stimuli to elicit a particular mental health difficulty.

In Chapter 3 a study is reported that investigated how the positive attributes of a virtual coach might influence the therapeutic alliance and expectancy of therapy outcome during the introductory stage of an automated VR therapy aimed at treating fear of

heights. Out of 705 screened, 120 individuals with a fear of heights participated. In a two-by-two factor, between-groups, randomised design, participants met a virtual coach that varied in warm facial expression (with/without) and affirmative nods (with/without). Both warm facial expressions ($p = 0.001, \eta_p^2 = 0.10$) and affirmative nods ($p = 0.040, \eta_p^2 = 0.04$) by the virtual coach increased the therapeutic alliance. Affirmative nods increased the treatment credibility ($p = 0.015,\ \eta_p^2 = 0.05$) and expectancy ($p = 0.015,\ \eta_p^2 = 0.05$). The results indicated that these positive attributes of the virtual coach enhanced therapy engagement and increased confidence in the treatment. Chapter 4 analyses participants' physiological responses, including cardiac and electrodermal data, to the VR fear of heights scenario and the emotional attributes of the VR coach. The results found that the coach's warm facial expressions ($p = 0.043, \eta_p^2 = 0.040$) increased the skin conductance responses and that affirmative head nods ($p = 0.059, \eta_p^2 = 0.034$) showed a trend to increase the tonic skin conductance level during the consultation, demonstrating that the coach could effectively modulate these physiological reactions.

In Chapter 5 a study is reported that examined the impact on paranoid thinking of the facial expressions of VHs in a VR social situation. Out of 1581 people screened, 122 individuals with elevated levels of paranoia participated. In a factorial between-groups design they rated their perceptions of VHs that varied in facial animation (static or animated) and expression (neutral or positive). Both facial animation ($p < 0.001, \eta_p^2 = 0.125$) and positive expressions ($p = 0.049, \eta_p^2 = 0.033$) led to less triggering of paranoid thoughts about the VHs. Positive expressions led to less visual attention to VHs when their faces were static ($estimate = -0.169, SE = 0.064, p = 0.010$). The findings revealed that the detailed design of VH facial expressions significantly influenced the occurrence of paranoid thoughts and visual attention in VR.

Overall, the studies reported in this thesis provide initial evidence of the psychological, behavioural, and physiological impact of VH design in relation to aspects of facial characteristics. The evidence demonstrates the potential benefits of considering the emotional attributes of virtual coaches in automated VR therapies for mental health difficulties. It also highlights the potential importance of considering the detailed design of VH faces when eliciting interpretations of the characters' potential intentions. The thesis presents work on the detailed design and implementation of VHs for two conditions (fear of heights and paranoia). Generalisability of the results to other mental health conditions is unknown. Future work is recommended to systematically examine the impacts of VHs across mental health conditions. A focus on the detailed design and implementation of VHs may increase the successful use of VR applications in mental health.

CONTENTS

FIGURES

TABLES

ABBREVIATIONS

VR: virtual reality

VH(s): virtual human(s)

AI: artificial intelligence

PTSD: post-traumatic stress disorder

SAD: social anxiety disorder

CBT: cognitive behavioural therapies

VA(s): virtual agent(s)

EDA: electrodermal activity

BVP: blood volume pulse

IBI: inter-beat interval

HR: heart rate

HRV: heart rate variability

LF/HF: low-frequency to high-frequency

RMSSD: root mean square of successive differences

ANS: autonomic nervous system

SNS: sympathetic nervous system

PNS: parasympathetic nervous system

SCL: skin conductance levels

SCR: skin conductance responses

SWT: stationary wavelet transform

ROIs: regions of interest

VAS: visual analogue scales

PPI: patients and public involvement

LLM(s): large language model(s)

1 Introduction

1.1 VIRTUAL REALITY FOR MENTAL HEALTH

Virtual reality (VR) has shown great promise in enhancing the understanding, assessment, and development of innovative interventions for mental health difficulties (Freeman et al., 2018, 2022; Bell et al., 2020). VR immerses individuals in 3D simulations that closely resemble real-life situations they fear, such as a public speaking scene where they practice addressing virtual audiences (Slater, Pertaub and Steed, 1999; Lindner et al., 2019) or a simulated combat zone to assist war veterans in managing post-traumatic stress disorder (PTSD) (Rizzo et al., 2010; Rothbaum et al., 2014). Central to the effectiveness of VR is the sense of presence, which refers to the degree to which users feel as though they are truly "there" in the virtual environment (Witmer and Singer, 1998; Sanchez-Vives and Slater, 2005). This immersive capability creates a realistic and controlled environment for exposure therapy – a psychological treatment where individuals confront their fears in a safe setting to reduce anxiety and avoidance (APA, 2020). VR allows patients to face their fears without being exposed to real-life situations, which could otherwise be dangerous. Moreover, VR offers high ecological validity, meaning that experimental outcomes obtained in a lab setting can help to understand how people would respond in a similar situation in the real world (Schmuckler, 2001).

VR is also an effective medium to deliver cognitive behavioural therapies (CBT) –the psychological treatment that helps individuals identify and change negative thinking and behaviour patterns (APA, 2020). In VR-based CBT, therapists can provide targeted guidance as patients face fearful scenarios and gradually learn to reduce defence mechanisms or safety-seeking behaviours (Wu et al., 2021). Without VR, this process may be more difficult due to the challenges of creating controlled, safe, repeatable, and plausible scenarios in real life. For instance, Freeman et al. (2016) tested the effects of a VR CBT for persecutory delusions, where participants were encouraged to drop their safety-seeking behaviours (e.g. avoiding eye contact) to engage with the neutral characters in the VR environments. The study found that VR CBT produced a significant reduction in delusional beliefs, with stronger effects than VR exposure alone. Additionally, the use of VR can enhance patient engagement. As a relatively novel experience in mental health treatment, VR makes patients feel valued by giving them access to an innovative approach (Brown et al., 2022; Freeman et al., 2023). The immersive, engaging experience of VR also made patients more willing to participate, with generally high levels of satisfaction and low side effects (Guillén, Baños and Botella, 2018; Freeman et al., 2023; Lundin, Yeap and Menkes, 2023).

VR has been instrumental in research on mental health disorders like paranoia, the unfounded belief that others intend to cause harm (Freeman et al., 2008, 2022; Pot-Kolder et al., 2018; Brown et al., 2020). By providing highly controlled environments designed specifically for the research aim, VR allows researchers and clinicians to simulate situations that trigger paranoia and closely monitor participants' responses. Freeman et al. (2008) explored individuals' paranoid reactions to neutral characters on a virtual train, demonstrating that paranoid ideation is present even in the general population and VR can be a valuable tool for understanding paranoid thoughts and related cognitive factors. Brown et al. (2020) later leveraged the same VR environment to investigate the effects of self-compassion and compassion for others on reducing paranoia. They found that targeted compassion techniques effectively reduced negative beliefs about oneself and others, thereby alleviating paranoia. Furthermore, Pot-Kolder et al. (2018) developed a VR-based therapy that exposed patients to simulated environments likely to trigger paranoid thoughts, such as crowded spaces or social interactions. Their findings suggested that VR exposure could reduce patients' paranoid ideation and momentary anxiety. More recently, Freeman and colleagues (2022) created an automated VR therapy called gameChange, targeting agoraphobic avoidance and distress in patients with psychosis. This program, led by a virtual coach providing guidance through various social scenarios and offered personalised support and strategies to manage anxiety, significantly reduced patients' avoidance behaviours and distress levels.

Researchers can collect different types of data in VR – including psychological responses through self-reported measures (Freeman et al., 2015, 2022; Atherton et al., 2016), physiological data like heart rate, skin conductance and electroencephalogram (EEG) signals (Martens et al., 2019; Bolinski et al., 2021; Luong and Holz, 2022; Gupta et al., 2024), and behavioural data such as head pose, eye gaze behaviour and proximity to the other VHs (Bönsch et al., 2020; Geraets et al., 2021; Rubin et al., 2022). Collecting different types of data helps identify subtle behavioural and physiological markers that subjective measures alone could overlook. For example, in VR environments where individuals can freely explore their surroundings, eye-gaze patterns can reveal attentional biases, social engagement levels, and avoidance behaviours, providing insights into cognitive and emotional processing (Günther et al., 2021; Wolf et al., 2021). The increasing integration of wearable devices and advanced sensors with immersive VR experiences also enhances the ability to capture physiological and behavioural responses, offering complementary and objective insights into individuals' mental health states (Halbig & Latoschik, 2021).

Despite VR's versatility, ethical considerations are crucial when integrating it into mental health applications. Virtual scenarios should be designed sensitively to avoid reinforcing negative stereotypes or exacerbating mental health difficulties. In exposure environments, where individuals confront fears, ensuring user safety is essential, including allowing users to pause or exit the experience if needed (Marloth et al., 2020). Participants should also be screened for medical or psychiatric conditions, including ongoing treatments or medications that may affect their emotional and physiological responses. During VR delivery, obtaining informed consent is vital, particularly regarding the collection and use of sensitive behavioural and physiological data. Additionally, risks such as motion sickness and unintended emotional triggers should be mitigated through tailored pre-screening, structured onboarding, and adherence to safety protocols.

1.2 Virtual Humans in VR Mental Health Research

Virtual humans (VHs) play a central role in studying social interactions in virtual environments (Schroeder, 2002; Pan and Hamilton, 2018). They can simulate social situations commonly encountered in daily life (Slater, Pertaub and Steed, 1999; Freeman et al., 2008; Gürerk et al., 2019), create social interactions under special circumstances such as conflict or danger (Rovira et al., 2009, 2021; Kane et al., 2012), and serve as avatars (digital representation of the participant), creating the illusion of having a virtual body and being able to control it (Fisher et al., 2020). Compared to traditional research methods that rely on trained actors to test people's responses in contextual scenarios, VHs offer better experimental control while still maintaining high ecological validity. Additionally, they provide the advantage of being easily reproducible (Pan and Hamilton, 2018).

Researchers have used VHs to assess people's mental health states by analysing their thoughts about VHs (Freeman et al., 2008; Shiban et al., 2015; Reichenberger et al., 2019) and observing their behaviours (e.g. distance to the VHs, eye gaze) during virtual encounters (Wieser et al., 2010; Han et al., 2014; Fusaro, Fanti and Chakrabarti, 2023). Additionally, VHs have been designed to interact with individuals to help develop specific skills, such as expressing compassion (Falconer et al., 2016; Kim et al., 2020; Hidding et al., 2024) and overcoming substance abuse (Cho et al., 2008; Shin et al., 2018; Guldager et al., 2022). One of the most exciting VH features, as demonstrated in the gameChange program (Freeman et al., 2022), is the use of virtual coaches. A virtual coach guides people through the experience and give feedback, facilitating the automation of VR therapies. This approach allows high-quality treatments to be delivered without the presence of a clinical psychologist during the VR session, thus helping mitigate the problems related to the shortage of clinical psychologists and helping overcome time and location constraints. Automated CBTs are still in their early days with limited applications (Freeman et al., 2018, 2022, 2024), but they open up opportunities to scale up the accessibility of evidence-based psychological treatments to a broader range of people who might need them.

Despite their growing use and potential, VHs have been barely studied as a primary focus in mental health research. This gap is partly due to the inherently multidisciplinary nature of the field, which requires expertise in psychology, immersive technology, and character animation to design VHs that are both clinically relevant and technically feasible. Additionally, real-world resource constraints make it impractical to conduct full-scale clinical trials solely to evaluate VH design within patient groups. In contrast, the computer science community has extensively explored VHs to optimise their visual and animation quality (Tsoli, Mahmood and Black, 2014; Zibrek, Kokkinara and Mcdonnell, 2018; Ponton et al., 2022), develop realistic and adaptive behaviours (Wienrich and Latoschik, 2021; Rovira and Slater, 2022; Llanes-Jurado et al., 2024), and create open-source character libraries (Gonzalez-Franco, Egan, et al., 2020; Gonzalez-Franco, Ofek, et al., 2020; Volonte et al., 2022). These efforts highlight the value of VHs in broader research domains. Understanding the role and impact of VHs in mental health research is essential to optimising current applications and expanding future uses. As VR technology advances, accessible and customisable VHs have the potential to play a greater role in mental health applications, offering personalised therapeutic experiences with finely tuned emotional attributes.

1.3 FACIAL EXPRESSIONS ON VIRTUAL HUMANS

Facial expressions are one of the most important channels of emotions and non-verbal behaviour in humans (Ekman, 2004; Poyatos, 2010; Kappas, Krumhuber and Küster, 2013). They convey emotions and intentions that shape interpersonal dynamics and significantly influence how others perceive and react in social interactions (Frith, 2009; Jack and Schyns, 2015). The ability to interpret and respond to facial expressions is important in human interactions. Misinterpretations, however, are sometimes associated with mental health difficulties such as psychosis (Seiferth et al., 2008; Souto et al., 2013; Gupta et al., 2023), anxiety (Kang et al., 2019; Dyer et al., 2022; Fujihara, Guo and Liu, 2023), and autism (Christinaki, Vidakis and Triantafyllidis, 2013; Loth et al., 2018; Talaat, 2023). Researchers can gain deeper insights into the emotional and cognitive processes involved by examining the impact of facial expressions on social interactions. This understanding can lead to more accurate assessments and the creation of targeted interventions that enhance social engagement and emotional well-being.

The impact of the facial expressions of VHs during social interactions have been studied in both non-immersive (Bickmore and Picard, 2004; Oh Kruzic et al., 2020; Takemoto et al., 2023) and immersive modalities (Park et al., 2009; Nijman et al., 2020; Volonte et al., 2020). For instance, Oh Kruzic et al. (2020) tested interactions between pairs of the participants communicating with each other through VHs, with or without facial and body movement displayed on the screen. They found that the addition of facial movements made participants like their partner more and form more accurate impressions of them. Notably, facial motion has a stronger impact than body motion. Volonte et al. (2020) compared psychological and physiological responses when participants encountered groups of VHs with varying facial expressions (e.g., positive, negative, neutral, and mixed) in VR. Participants exposed to VHs with positive facial expressions exhibited less negative emotions and higher physiological arousal, as measured by electrodermal activity, compared to other groups. These studies suggest VH facial expressions can influence users' mood and their physiological responses. The effect sizes varied from small to large (0.3-0.9 for Cohen's d; 0.1-0.3 for η^2), highlighting the significant role and nuanced complexity of facial expressions. Despite this potential and complexity, there is still limited research explicitly studying facial expressions within VR environments in the context of mental health.

Advances in the technical pipeline and commercial software have made the creation and implementation of facial expressions on VHs more efficient and accessible than ever before. 3D software like Autodesk Maya[2] and Blender[3] and open-source character libraries such as Microsoft Rocketbox[4] offer a set of virtual characters with facial rigging, blend-shape animations, and are optimised for real-time rendering. Additionally, artificial intelligence (AI) powered 3D animation pipelines, such as iClone[5] and Daz3D[6], facilitate more efficient prototyping, making it easier to test facial animations in experimental studies. For example, iClone uses AI to generate synchronised lip movements from audio files, making the process more efficient. These tools enable the creation of detailed and dynamic facial expressions, which are crucial for producing realistic and emotionally plausible VHs. However, creating plausible and contextually appropriate facial expressions remains a complex and challenging task. It requires not only technical proficiency, but also a deep understanding of human facial anatomy and emotions to ensure that the expressions are suitable and believable within specific contexts (Maestri, 2006). In summary, technologies have made the process more cost-effective and accessible, but achieving high-quality facial animations still requires considerable expertise and understanding of the social context.

[2] https://www.autodesk.com/products/maya/overview
[3] https://www.blender.org/
[4] https://github.com/microsoft/Microsoft-Rocketbox
[5] https://www.reallusion.com/iclone/
[6] https://www.daz3d.com/

1.4 AIM FOR THE THESIS

The thesis aims to understand the potential of VHs and their facial expressions to optimise VR mental health applications. The objectives are as follows:

1. Provide a comprehensive literature review of the use of VHs in immersive VR for mental health applications.

2. Evaluate the effects of detailed programming of VHs (e.g. facial expressions) in VHs on people's perception and behaviour in mental health VR experience.

3. Investigate physiological responses and eye gaze behaviour in relation to the detailed features of VHs in VR mental health settings.

To address the first objective, a systematic review was conducted to examine current use and future opportunities for VHs in VR mental health research (Chapter 2). For the second objective, bespoke facial animations tailored to the social contexts were implemented on VHs and tested through two large-scale user studies. Specific target groups (i.e. individuals with a fear of heights and those with elevated paranoia) were recruited as exemplars of mental health difficulties, and their psychological and behavioural responses were assessed, taking into account their mental health states (Chapter 3, Chapter 5). For the third objective, Chapter 4 further examines participants' physiological responses (e.g. heart rate and skin conductance) to a virtual coach, a VH that guides users through VR therapy by providing instructions and feedback, displaying different emotional attributes. Chapter 5 explores how these VHs affected participants' behavioural responses (i.e. visual attention through their eye gaze data) during interactions. Finally, in Chapter 6, we present the overall conclusions, limitations, and future directions of research from the study results presented in this thesis.

REFERENCES

APA. (2020). *Publication manual of the American Psychological Association, 7th ed* (pp. xxii, 427). American Psychological Association. https://doi.org/10.1037/0000165-000

Atherton, S., Antley, A., Evans, N., Cernis, E., Lister, R., Dunn, G., Slater, M., & Freeman, D. (2016). Self-Confidence and Paranoia: An Experimental Study Using an Immersive Virtual Reality Social Situation. *Behavioural and Cognitive Psychotherapy*, *44*(1), 56–64. https://doi.org/10.1017/S1352465814000496

Bell, I. H., Nicholas, J., Alvarez-Jimenez, M., Thompson, A., & Valmaggia, L. (2020). Virtual reality as a clinical tool in mental health research and practice. *Dialogues in Clinical Neuroscience*, *22*(2), 169–177. https://doi.org/10.31887/DCNS.2020.22.2/lvalmaggia

Bickmore, T. W., & Picard, R. W. (2004). Towards caring machines. *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, 1489–1492. https://doi.org/10.1145/985921.986097

Bolinski, F., Etzelmüller, A., De Witte, N. A. J., van Beurden, C., Debard, G., Bonroy, B., Cuijpers, P., Riper, H., & Kleiboer, A. (2021). Physiological and self-reported arousal in virtual reality versus face-to-face emotional activation and cognitive restructuring in university students: A crossover experimental study using wearable monitoring. *Behaviour Research and Therapy*, *142*, 103877. https://doi.org/10.1016/j.brat.2021.103877

Bönsch, A., Radke, S., Ehret, J., Habel, U., & Kuhlen, T. W. (2020). The Impact of a Virtual Agent's Non-Verbal Emotional Expression on a User's Personal Space Preferences. *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*, 1–8. https://doi.org/10.1145/3383652.3423888

Brown, P., Waite, F., Lambe, S., Jones, J., Jenner, L., Diamond, R., & Freeman, D. (2022). Automated Virtual Reality Cognitive Therapy (gameChange) in Inpatient Psychiatric Wards: Qualitative Study of Staff and Patient Views Using an Implementation Framework. *JMIR Formative Research*, *6*(4), e34225. https://doi.org/10.2196/34225

Brown, P., Waite, F., Rovira, A., Nickless, A., & Freeman, D. (2020). Virtual reality clinical-experimental tests of compassion treatment techniques to reduce paranoia. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-64957-7

Cho, S., Ku, J., Park, J., Han, K., Lee, H., Choi, Y. K., Jung, Y.-C., Namkoong, K., Kim, J.-J., Kim, I. Y., Kim, S. I., & Shen, D. F. (2008). Development and Verification of an Alcohol Craving–Induction Tool Using Virtual Reality: Craving Characteristics in Social Pressure Situation. *CyberPsychology & Behavior*, *11*(3), Article 3. https://doi.org/10.1089/cpb.2007.0149

Christinaki, E., Vidakis, N., & Triantafyllidis, G. (2013). Facial expression recognition teaching to preschoolers with autism: A natural user interface approach. *Proceedings of* *the 6th Balkan Conference in Informatics*, 141–148. https://doi.org/10.1145/2490257.2490262

Dyer, M. L., Attwood, A. S., Penton-Voak, I. S., & Munafò, M. R. (2022). The role of state and trait anxiety in the processing of facial expressions of emotion. *Royal Society Open Science*, *9*(1), 210056. https://doi.org/10.1098/rsos.210056

Ekman, P. (2004). *Emotions Revealed: Understanding Faces and Feelings*. https://www.amazon.co.uk/Emotions-Revealed-Understanding-Faces-Feelings/dp/0753817659

Falconer, C. J., Rovira, A., King, J. A., Gilbert, P., Antley, A., Fearon, P., Ralph, N., Slater, M., & Brewin, C. R. (2016). Embodying self-compassion within virtual reality and its effects on patients with depression. *BJPsych Open*, *2*(1), 74–80. https://doi.org/10.1192/bjpo.bp.115.002147

Freeman, D., Bradley, J., Antley, A., Bourke, E., DeWeever, N., Evans, N., Černis, E., Sheaves, B., Waite, F., Dunn, G., Slater, M., & Clark, D. M. (2016). Virtual reality in the treatment of persecutory delusions: Randomised controlled experimental study testing how to reduce delusional conviction. *British Journal of Psychiatry*, *209*(1), 62–67. https://doi.org/10.1192/bjp.bp.115.176438

Freeman, D., Dunn, G., Murray, R. M., Evans, N., Lister, R., Antley, A., Slater, M., Godlewska, B., Cornish, R., Williams, J., Di Simplicio, M., Igoumenou, A., Brenneisen, R., Tunbridge, E. M., Harrison, P. J., Harmer, C. J., Cowen, P., & Morrison, P. D. (2015). How Cannabis Causes Paranoia: Using the Intravenous Administration of Δ 9 -Tetrahydrocannabinol (THC) to Identify Key Cognitive Mechanisms Leading to Paranoia. *Schizophrenia Bulletin*, *41*(2), 391–399. https://doi.org/10.1093/schbul/sbu098

Freeman, D., Freeman, J., Ahmed, M., Haynes, P., Beckwith, H., Rovira, A., Miguel, A. L., Ward, R., Bousfield, M., Riffiod, L., Kabir, T., Waite, F., & Rosebrock, L. (2024). Automated VR therapy for improving positive self-beliefs and psychological well-being in young patients with psychosis: A proof of concept evaluation of Phoenix VR self-confidence therapy. *Behavioural and Cognitive Psychotherapy*, *52*(3), 277–287. https://doi.org/10.1017/S1352465823000553

Freeman, D., Haselton, P., Freeman, J., Spanlang, B., Kishore, S., Albery, E., Denne, M., Brown, P., Slater, M., & Nickless, A. (2018). Automated psychological therapy using immersive virtual reality for treatment of fear of heights: A single-blind, parallel-group, randomised controlled trial. *The Lancet Psychiatry*, *5*(8), 625–632. https://doi.org/10.1016/S2215-0366(18)30226-8

Freeman, D., Lambe, S., Kabir, T., Petit, A., Rosebrock, L., Yu, L.-M., Dudley, R., Chapman, K., Morrison, A., O'Regan, E., Aynsworth, C., Jones, J., Murphy, E., Powling, R., Galal, U., Grabey, J., Rovira, A., Martin, J., Hollis, C., … West, J. (2022). Automated virtual reality therapy to treat agoraphobic avoidance and distress in patients with psychosis (gameChange): A multicentre, parallel-group, single-blind, randomised, controlled trial in England with mediation and moderation analyses. *The Lancet Psychiatry*, *9*(5), 375–388. https://doi.org/10.1016/S2215-0366(22)00060-8

Freeman, D., Pugh, K., Antley, A., Slater, M., Bebbington, P., Gittins, M., Dunn, G., Kuipers, E., Fowler, D., & Garety, P. (2008). Virtual reality study of paranoid thinking in the general population. *The British Journal of Psychiatry*, *192*(4), 258–263. https://doi.org/10.1192/bjp.bp.107.044677

Freeman, D., Rosebrock, L., Waite, F., Loe, B. S., Kabir, T., Petit, A., Dudley, R., Chapman, K., Morrison, A., O'Regan, E., Aynsworth, C., Jones, J., Murphy, E., Powling, R., Peel, H., Walker, H., Byrne, R., Freeman, J., Rovira, A., … Lambe, S. (2023). Virtual reality (VR) therapy for patients with psychosis: Satisfaction and side effects. *Psychological Medicine*, *53*(10), 4373–4384. https://doi.org/10.1017/S0033291722001167

Frith, C. (2009). Role of facial expressions in social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3453–3458. https://doi.org/10.1098/rstb.2009.0142

Fujihara, Y., Guo, K., & Liu, C. H. (2023). Relationship between types of anxiety and the ability to recognize facial expressions. *Acta Psychologica*, *241*, 104100. https://doi.org/10.1016/j.actpsy.2023.104100

Fusaro, M., Fanti, V., & Chakrabarti, B. (2023). Greater interpersonal distance in adults with autism. *Autism Research*, *16*(10), 2002–2007. https://doi.org/10.1002/aur.3013

Geraets, C. N. W., Klein Tuente, S., Lestestuiver, B. P., van Beilen, M., Nijman, S. A., Marsman, J. B. C., & Veling, W. (2021). Virtual reality facial emotion recognition in social environments: An eye-tracking study. *Internet Interventions*, *25*, 100432. https://doi.org/10.1016/j.invent.2021.100432

Gonzalez-Franco, M., Egan, Z., Peachey, M., Antley, A., Randhavane, T., Panda, P., Zhang, Y., Wang, C. Y., Reilly, D. F., Peck, T. C., Won, A. S., Steed, A., & Ofek, E. (2020). MoveBox: Democratizing MoCap for the Microsoft Rocketbox Avatar Library. *2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, 91–98. https://doi.org/10.1109/AIVR50618.2020.00026

Gonzalez-Franco, M., Ofek, E., Pan, Y., Antley, A., Steed, A., Spanlang, B., Maselli, A., Banakou, D., Pelechano, N., Orts-Escolano, S., Orvalho, V., Trutoiu, L., Wojcik, M., Sanchez-Vives, M. V., Bailenson, J., Slater, M., & Lanier, J. (2020). The Rocketbox Library and the Utility of Freely Available Rigged Avatars. *Frontiers in Virtual Reality*, *1*. https://doi.org/10.3389/frvir.2020.561558

Guillén, V., Baños, R. M., & Botella, C. (2018). Users' Opinion About a Virtual Reality System as an Adjunct to Psychological Treatment for Stress-Related Disorders: A Quantitative and Qualitative Mixed-Methods Study. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01038

Guldager, J. D., Kjær, S. L., Grittner, U., & Stock, C. (2022). Efficacy of the Virtual Reality Intervention VR FestLab on Alcohol Refusal Self-Efficacy: A Cluster-Randomized Controlled Trial. *International Journal of Environmental Research and Public Health*, *19*(6), 3293. https://doi.org/10.3390/ijerph19063293

Gupta, K., Zhang, Y., Gunasekaran, T. S., Krishna, N., Pai, Y. S., & Billinghurst, M. (2024). CAEVR: Biosignals-Driven Context-Aware Empathy in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, 1–11. https://doi.org/10.1109/TVCG.2024.3372130

Gupta, T., Osborne, K. J., Nadig, A., Haase, C. M., & Mittal, V. A. (2023). Alterations in facial expressions in individuals at risk for psychosis: A facial electromyography approach using emotionally evocative film clips. *Psychological Medicine*, *53*(12), 5829–5838. https://doi.org/10.1017/S0033291722003087

Gürerk, Ö., Bönsch, A., Kittsteiner, T., & Staffeldt, A. (2019). Virtual humans as co-workers: A novel methodology to study peer effects. *Journal of Behavioral and Experimental Economics*, *78*, 17–29. https://doi.org/10.1016/j.socec.2018.11.003

Günther, V., Kropidlowski, A., Schmidt, F. M., Koelkebeck, K., Kersting, A., & Suslow, T. (2021). Attentional processes during emotional face perception in social anxiety disorder: A systematic review and meta-analysis of eye-tracking findings. Progress in Neuro-Psychopharmacology and Biological Psychiatry, 111, 110353. https://doi.org/10.1016/j.pnpbp.2021.110353

Marloth, M., Chandler, J., & Vogeley, K. (2020). Psychiatric Interventions in Virtual Reality: Why We Need an Ethical Framework. *Cambridge Quarterly of Healthcare Ethics, 29(4),* 574–584. https://doi.org/10.1017/S0963180120000328

Halbig, A., & Latoschik, M. E. (2021). A Systematic Review of Physiological Measurements, Factors, Methods, and Applications in Virtual Reality. *Frontiers in Virtual Reality*, *2*. https://doi.org/10.3389/frvir.2021.694567

Han, K., Shin, J., Yoon, S. Y., Jang, D.-P., & Kim, J.-J. (2014). Deficient gaze pattern during virtual multiparty conversation in patients with schizophrenia. *Computers in Biology and Medicine*, *49*, 60–66. https://doi.org/10.1016/j.compbiomed.2014.03.012

Hidding, M., Veling, W., Pijnenborg, G. H. M., & van der Stouwe, E. C. D. (2024). A single-session VR intervention addressing self-compassion and self-criticism with and without perspective change: Results of a randomized controlled experiment. *Behaviour Research and Therapy*, *173*, 104466. https://doi.org/10.1016/j.brat.2023.104466

Jack, R. E., & Schyns, P. G. (2015). The Human Face as a Dynamic Tool for Social Communication. *Current Biology*, *25*(14), R621–R634. https://doi.org/10.1016/j.cub.2015.05.052

Kane, H. S., McCall, C., Collins, N. L., & Blascovich, J. (2012). Mere presence is not enough: Responsive support in a virtual world. *Journal of Experimental Social Psychology*, *48*(1), 37–44. https://doi.org/10.1016/j.jesp.2011.07.001

Kang, W., Kim, G., Kim, H., & Lee, S.-H. (2019). The Influence of Anxiety on the Recognition of Facial Emotion Depends on the Emotion Category and Race of the Target Faces. *Experimental Neurobiology*, *28*(2), 261–269. https://doi.org/10.5607/en.2019.28.2.261

Kappas, A., Krumhuber, E., & Küster, D. (2013). 6 Facial behavior. In J. A. Hall & M. L. Knapp (Eds.), *Nonverbal Communication* (pp. 131–166). De Gruyter Mouton. https://doi.org/10.1515/9783110238150.131

Kim, J., Jung, Y. H., Shin, Y.-B., Kim, M.-K., Eom, H., Kim, E., Kim, J., & Kim, J.-J. (2020). Development and Validation of a Virtual Reality-Based Training Program for Promoting Subjective Well-Being. *Psychiatry Investigation*, *17*(12), 1207–1215. https://doi.org/10.30773/pi.2020.0311

Lindner, P., Miloff, A., Fagernäs, S., Andersen, J., Sigeman, M., Andersson, G., Furmark, T., & Carlbring, P. (2019). Therapist-led and self-led one-session virtual reality exposure therapy for public speaking anxiety with consumer hardware and software: A randomized controlled trial. *Journal of Anxiety Disorders*, *61*, 45–54. https://doi.org/10.1016/j.janxdis.2018.07.003

Llanes-Jurado, J., Gómez-Zaragozá, L., Minissi, M. E., Alcañiz, M., & Marín-Morales, J. (2024). Developing conversational Virtual Humans for social emotion elicitation based on large language models. *Expert Systems with Applications*, *246*, 123261. https://doi.org/10.1016/j.eswa.2024.123261

Loth, E., Garrido, L., Ahmad, J., Watson, E., Duff, A., & Duchaine, B. (2018). Facial expression recognition as a candidate marker for autism spectrum disorder: How frequent and severe are deficits? *Molecular Autism*, *9*(1), 7. https://doi.org/10.1186/s13229-018-0187-7

Lundin, R. M., Yeap, Y., & Menkes, D. B. (2023). Adverse Effects of Virtual and Augmented Reality Interventions in Psychiatry: Systematic Review. *JMIR Mental Health*, *10*, e43240. https://doi.org/10.2196/43240

Luong, T., & Holz, C. (2022). Characterizing Physiological Responses to Fear, Frustration, and Insight in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, *28*(11), 3917–3927. IEEE Transactions on Visualization and Computer Graphics. https://doi.org/10.1109/TVCG.2022.3203113

Maestri, G. (2006). Facial and Dialogue Animation. In *Digital Character Animation 3* (pp. 196–219). New Riders.

Martens, M. A., Antley, A., Freeman, D., Slater, M., Harrison, P. J., & Tunbridge, E. M. (2019). It feels real: Physiological responses to a stressful virtual reality environment and its impact on working memory. *Journal of Psychopharmacology*, *33*(10), 1264–1273. https://doi.org/10.1177/0269881119860156

Nijman, S. A., Veling, W., Greaves-Lord, K., Vos, M., Zandee, C. E. R., Aan het Rot, M., Geraets, C. N. W., & Pijnenborg, G. H. M. (2020). Dynamic Interactive Social Cognition Training in Virtual Reality (DiSCoVR) for People With a Psychotic Disorder: Single-Group Feasibility and Acceptability Study. *JMIR Mental Health*, *7*(8). https://doi.org/10.2196/17808

Oh Kruzic, C., Kruzic, D., Herrera, F., & Bailenson, J. (2020). Facial expressions contribute more than body movements to conversational outcomes in avatar-mediated virtual environments. *Scientific Reports*, *10*(1), 20626. https://doi.org/10.1038/s41598-020-76672-4

Pan, X., & Hamilton, A. F. de C. (2018). Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology*, *109*(3), 395–417. https://doi.org/10.1111/bjop.12290

Park, I. H., Kim, J.-J., Ku, J., Jang, H. J., Park, S.-H., Kim, C.-H., Kim, I. Y., & Kim, S. I. (2009). Characteristics of social anxiety from virtual interpersonal interactions in patients with schizophrenia. *Psychiatry*, *72*(1), 79–93. https://doi.org/10.1521/psyc.2009.72.1.79

Ponton, J. L., Yun, H., Andujar, C., & Pelechano, N. (2022). Combining Motion Matching and Orientation Prediction to Animate Avatars for Consumer-Grade VR Devices. *Computer Graphics Forum*, *41*(8), 107–118. https://doi.org/10.1111/cgf.14628

Pot-Kolder, R. M. C. A., Geraets, C. N. W., Veling, W., Beilen, M. van, Staring, A. B. P., Gijsman, H. J., Delespaul, P. A. E. G., & Gaag, M. van der. (2018). Virtual-reality-based cognitive behavioural therapy versus waiting list control for paranoid ideation and social avoidance in patients with psychotic disorders: A single-blind randomised controlled trial. *The Lancet Psychiatry*, *5*(3), 217–226. https://doi.org/10.1016/S2215-0366(18)30053-1

Poyatos, F. (2010). Forms and Functions of Nonverbal Communication in the Novel: A New Perspective of the Author-Character-Reader Relationship. In A. Kendon (Ed.), *Nonverbal Communication, Interaction, and Gesture: Selections from SEMIOTICA* (pp. 107–150). De Gruyter Mouton. https://doi.org/10.1515/9783110880021.107

Reichenberger, J., Pfaller, M., Forster, D., Gerczuk, J., Youssef Shiban, & Mühlberger, A. (2019). Men Scare Me More: Gender Differences in Social Fear Conditioning in Virtual Reality. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.01617

Rizzo, A. S., Difede, J., Rothbaum, B. O., Reger, G., Spitalnick, J., Cukor, J., & McLay, R. (2010). Development and early evaluation of the Virtual Iraq/Afghanistan exposure therapy system for combat-related PTSD. *Annals of the New York Academy of Sciences*, *1208*, 114–125. https://doi.org/10.1111/j.1749-6632.2010.05755.x

Rothbaum, B. O., Price, M., Jovanovic, T., Norrholm, S. D., Gerardi, M., Dunlop, B., Davis, M., Bradley, B., Duncan, E. J., Rizzo, A., & Ressler, K. J. (2014). A Randomized, Double-Blind Evaluation of d-Cycloserine or Alprazolam Combined With Virtual Reality Exposure Therapy for Posttraumatic Stress Disorder in Iraq and Afghanistan War Veterans. *American Journal of Psychiatry*, *171*(6), 640–648. https://doi.org/10.1176/appi.ajp.2014.13121625

Rovira, A., & Slater, M. (2022). Encouraging bystander helping behaviour in a violent incident: A virtual reality study using reinforcement learning. *Scientific Reports*, *12*(1), 3843. https://doi.org/10.1038/s41598-022-07872-3

Rovira, A., Southern, R., Swapp, D., Campbell, C., Zhang, J. J., Levine, M., & Slater, M. (2021). Bystander Affiliation Influences Intervention Behavior: A Virtual Reality Study. *SAGE Open*, *11*(3), 21582440211040076. https://doi.org/10.1177/21582440211040076

Rovira, A., Swapp, D., Spanlang, B., & Slater, M. (2009). The use of virtual reality in the study of people's responses to violent incidents. *Frontiers in Behavioral Neuroscience*, *3*. https://doi.org/10.3389/neuro.08.059.2009

Rubin, M., Muller, K., Hayhoe, M. M., & Telch, M. J. (2022). Attention guidance augmentation of virtual reality exposure therapy for social anxiety disorder: A pilot randomized controlled trial. *Cognitive Behaviour Therapy*, *51*(5), 371–387. https://doi.org/10.1080/16506073.2022.2053882

Sanchez-Vives, M. V., & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience*, *6*(4), 332–339. https://doi.org/10.1038/nrn1651

Schmuckler, M. (2001). What Is Ecological Validity? A Dimensional Analysis. *Infancy*, *2*. https://doi.org/10.1207/S15327078IN0204_02

Schroeder, R. (2002). Social Interaction in Virtual Environments: Key Issues, Common Themes, and a Framework for Research. In R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments* (pp. 1–18). Springer. https://doi.org/10.1007/978-1-4471-0277-9_1

Seiferth, N. Y., Pauly, K., Habel, U., Kellermann, T., Jon Shah, N., Ruhrmann, S., Klosterkötter, J., Schneider, F., & Kircher, T. (2008). Increased neural response related to neutral faces in individuals at risk for psychosis. *NeuroImage*, *40*(1), 289–297. https://doi.org/10.1016/j.neuroimage.2007.11.020

Shiban, Y., Reichenberger, J., Neumann, I. D., & Mühlberger, A. (2015). Social conditioning and extinction paradigm: A translational study in virtual reality. *Frontiers in Psychology*, *6*, 400. https://doi.org/10.3389/fpsyg.2015.00400

Shin, Y.-B., Kim, J.-J., Kim, M.-K., Kyeong, S., Jung, Y. H., Eom, H., & Kim, E. (2018). Development of an effective virtual environment in eliciting craving in adolescents and young adults with internet gaming disorder. *PLOS ONE*, *13*(4), Article 4. https://doi.org/10.1371/journal.pone.0195677

Slater, M., Pertaub, D.-P., & Steed, A. (1999). Public Speaking in Virtual Reality: Facing an Audience of Avatars. *Computer Graphics and Applications, IEEE*, *19*, 6–9. https://doi.org/10.1109/38.749116

Souto, T., Baptista, A., Tavares, D., Queirós, C., & António, M. (2013). Facial emotional recognition in schizophrenia: Preliminary results of the virtual reality program for facial emotional recognition. *Archives of Clinical Psychiatry (São Paulo)*, *40*(4), 129–134. https://doi.org/10.1590/S0101-60832013000400001

Takemoto, A., Aispuriete, I., Niedra, L., & Dreimane, L. F. (2023). Differentiating depression using facial expressions in a virtual avatar communication system. *Frontiers in Digital Health*, *5*. https://doi.org/10.3389/fdgth.2023.1080023

Talaat, F. M. (2023).   facial emotion recognition system among children with autism based on deep learning and IoT. *Neural Computing and Applications*, *35*(17), 12717–12728. https://doi.org/10.1007/s00521-023-08372-9

Tsoli, A., Mahmood, N., & Black, M. J. (2014). Breathing life into shape: Capturing, modeling and animating 3D human breathing. *ACM Trans. Graph.*, *33*(4), 52:1-52:11. https://doi.org/10.1145/2601097.2601225

Volonte, M., Hsu, Y.-C., Liu, K.-Y., Mazer, J. P., Wong, S.-K., & Babu, S. V. (2020). Effects of Interacting with a Crowd of Emotional Virtual Humans on Users' Affective and Non-Verbal Behaviors. *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 293–302. https://doi.org/10.1109/VR46266.2020.00049

Volonte, M., Ofek, E., Jakubzak, K., Bruner, S., & Gonzalez-Franco, M. (2022). HeadBox: A Facial Blendshape Animation Toolkit for the Microsoft Rocketbox Library. *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 39–42. https://doi.org/10.1109/VRW55335.2022.00015

Wolf, A., Ueda, K., & Hirano, Y. (2021). Recent updates of eye movement abnormalities in patients with schizophrenia: A scoping review. *Psychiatry and Clinical Neurosciences*, 75(3), 82–100. https://doi.org/10.1111/pcn.13188

Wienrich, C., & Latoschik, M. E. (2021). eXtended Artificial Intelligence: New Prospects of Human-AI Interaction Research. *Frontiers in Virtual Reality*, *2*. https://doi.org/10.3389/frvir.2021.686783

Wieser, M., Pauli, P., Grosseibl, M., Molzow, I., & Mühlberger, A. (2010). Virtual Social Interactions in Social Anxiety-The Impact of Sex, Gaze, and Interpersonal Distance. *Cyberpsychology, Behavior and Social Networking*, *13*, 547–554. https://doi.org/10.1089/cyber.2009.0432

Witmer, B., & Singer, M. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, *7*(3), 225–240.

Wu, J., Sun, Y., Zhang, G., Zhou, Z., & Ren, Z. (2021). Virtual Reality-Assisted Cognitive Behavioral Therapy for Anxiety Disorders: A Systematic Review and Meta-Analysis. *Frontiers in Psychiatry*, *12*. https://doi.org/10.3389/fpsyt.2021.575094

Zibrek, K., Kokkinara, E., & Mcdonnell, R. (2018). The Effect of Realistic Appearance of Virtual Characters in Immersive Environments—Does the Character's Personality Play a Role? *IEEE Transactions on Visualization and Computer Graphics*, *24*(4), 1681–1690. IEEE Transactions on Visualization and Computer Graphics. https://doi.org/10.1109/TVCG.2018.2794638

2 Systematic Review of Virtual Humans in Virtual Reality Mental Health Research

ABSTRACT

This chapter examines the use of virtual humans (VHs) in immersive virtual reality (VR) in the context of mental health research by exploring their roles and characteristics. We conducted a systematic review that included 73 papers. VHs were used in research studies related to social anxiety (n = 18), eating disorders (n = 18), and psychosis (n = 15). They were primarily used as active social interaction partners (n = 37), part of a virtual crowds (n = 13), and virtual bodies (n = 23). Explicit interactions were associated with active interaction partners, and implicit and passive interactions were frequently implemented in virtual crowds and virtual body studies. Over half of the studies examined the effects of manipulating specific characteristics of VHs, with common modifications as body size, gender and age. Research also focused on behavioural aspects such as levels of attention and gaze direction, while few studies addressed VH emotional characteristics or personality. This review underscores the importance of VHs in mental health VR and highlights the need for the systematic examinations to better understand the impact of VH design on mental health outcomes.

2.1 INTRODUCTION

Virtual humans (VHs) have been part of VR mental health applications since the late 1990s. The pioneering work by Riva & Melis (1997) used images of a female silhouette in VR to represent the participant's body, marking one of the earliest uses of VH in mental health research. Other early studies showed that virtual audiences could trigger people's anxiety effectively (Pertaub et al., 2002; Slater et al., 1999). The improvement of VR technologies and cost reduction since approximately 2013 made VR more accessible, allowing VH research to expand into areas such as understanding paranoia (Fornells-Ambrojo et al., 2016; Valmaggia et al., 2016), treating autism (Amaral et al., 2018; Fusaro et al., 2023) or improving PTSD therapy (A. 'Skip' Rizzo & Shilling, 2017). More recently, some applications incorporated virtual coaches to guide patients and automate treatments (Freeman et al., 2022, 2024), as well as body-swapping experiences that allow participants to switch between different entities to experience perspective changes (Falconer et al., 2016; Halim et al., 2023; Slater et al., 2019).

Implementing VHs for mental health research is a complex task that requires collaboration between psychologists and computer scientists but has not received the attention its complexity requires. Chapter 1 provided an overview of the main concepts surrounding VHs in the context of mental health. In this Chapter, we present a systematic review of research studies in mental health that used VR scenarios featuring VHs, from the early development of VR until May 2024. The review includes a classification of how VHs were used in these studies. This review addresses the following questions:

- What roles and functions have VHs played in VR mental health research?
- What have been the chosen agencies for VHs in VR mental health research?
- What have been the examined characteristics of VHs in VR mental health research?

2.2 CATEGORISING VHS

We categorise VHs and discuss their features from three main aspects – role, agency, and the interaction type.

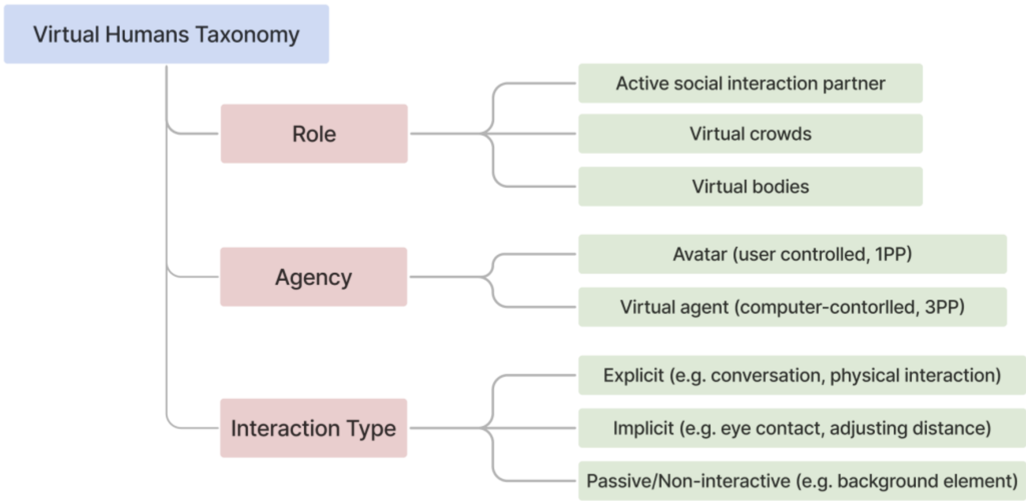


Figure2.1. Virtual human taxonomy based on existing mental health VR study.

2.2.1 Role

The role of VHs can be divided into three categories: active social interaction partner, virtual crowds and virtual bodies. These categories were chosen to reflect the distinct ways VHs are utilised in VR mental health research, offering a simple, clear and functional framework to classify their roles and contributions across diverse applications. Active social interaction partners engage with participants to simulate interactions through verbal and non-verbal communication. A virtual coach is an example of active interaction partner that focuses on providing guidance or feedback to facilitate psychological treatment (Freeman et al., 2022). Unlike active social interaction partners, virtual crowds have limited contact with participants or normally interact passively with implicit interaction such as making eye contact only (Jorjafki et al., 2018). It is common for virtual crowds to act as part of the environmental stimuli to populate the virtual world (A. S. Rizzo et al., 2010). Lastly, virtual bodies focus on body representation, and are used either to represent a range of different body sizes or to embody the user (e.g. user's avatar) for an embodiment experience (Ferrer-Garcia M. et al., 2017).

2.2.2 Agency - Avatars and agents

Agency refers to the state or entity of taking action or exerting control (Nowak & Biocca, 2003), and can be categorised as avatars and virtual agents (VAs). An avatar is a digital representation controlled by a user, serving as their embodiment in virtual environments (Nowak & Biocca, 2003). A virtual agent (VA) is a computer-generated character designed to simulate human interaction autonomously, and normally controlled by a computer rather than a person in real time (Cassell, 2001).

Avatars represent the users' body in the virtual environment and are often viewed from a first-person perspective (1PP). In a typical setup, when users look down in VR, they see the virtual body instead of their own (co-location). With the use of a real-time full-body tracking, the user's body movements can be mapped onto the avatar, allowing it to follow their movements in real time. This co-location and the synchronised movements are the two key factors that create the illusion of body ownership, a foundational element of embodiment (Kilteni et al., 2012). However, the term 'avatar' is sometimes misused in mental health VR studies to describe all virtual humans (VHs), even when they do not represent participants.

VAs, on the other hand, are typically controlled by computers and viewed from a third-person perspective (3PP). They can exhibit varying degrees of autonomy. A notable example is the semi-autonomous agent, which combines computer automation with operator intervention. These are often used in psychological experiments, where the researcher can modify the VA's behaviours in real time based on participants' response (Pan & Hamilton, 2018). This approach is rooted in the Wizard of Oz technique, where a hidden operator simulates autonomous system behaviour (Salber & Coutaz, 1993). This paradigm has also been used in VR, such as triggering different pre-recorded responses of a virtual interviewer in a job interview practise (Hartanto et al., 2014).

2.2.3 Interaction types of VAs

VAs exhibit different degrees of interaction with the human participants. We have divided them into three main blocks – explicit, implicit, and passive or non-interactive.

Explicit Interactions: We include any direct engagement between a VA and a human participant in this category. A good example is having a conversation or physical interactions such as shaking hands (Joundi et al., 2021), or even performing collaborative work, such as playing a sport together (Kothgassner et al., 2021). This level of interaction aims to create social experiences similar to a real-life situation, although it can be very expensive to control all the possible cases of where the interaction may lead to.

Implicit Interactions: They include subtler responses, where VAs acknowledge the participant's presence with simple actions such as looking at the person, or stepping back to increase the interpersonal distance, without fully engage in an explicit interaction. Such interactions are cheaper to implement but still increase the sense of presence (Slater & Wilbur, 1997; Spyridonis et al., 2024), making participants feel as if they are part of the scene rather than being spectators.

Passive or No Interactions: These VAs act as non-interactive background elements in VR environments, unresponsive to participant actions. They're easier to implement and useful in scenarios require crowd simulation, such as populating a virtual restaurant (Brinkman et al., 2011) or constructing an office scene (Tarnanas et al., 2003).

2.3 METHODS

2.3.1 Inclusion and exclusion criteria

Inclusion criteria: Papers were required to: (1) Use VHs; (2) Use immersive VR; (3) Used for assessment or intervention of mental health conditions through empirical tests with human participants.

Exclusion criteria: Papers were excluded if they: (1) Had no quantitative data or case studies (number of participants = 1); (2) Were focused on psychology concepts (e.g. the rubber hand illusion paradigm); (3) Were not available as a full text.

2.3.1 Searching strategy

The literature search was performed on MEDLINE, Embase, PsycINFO, and Ovid, with title and abstract searches for the keyword combinations around (1) virtual reality, (2) VH, and (3) mental health with the search query: [(virtual reality OR immersive virtual reality OR VR) AND (virtual human OR virtual character OR virtual agent OR avatar OR humanoid) AND ((assessment OR treatment OR therapy OR mental health) OR (mood disorders OR depress* OR bipolar OR mania OR paranoia OR psychosis OR psychotic OR schizophren* OR schizotyp* OR delus* OR hallucinat* OR phobias OR obsessive compulsive disorder OR OCD OR anxiety OR post traumatic stress disorder OR PTSD OR trauma OR anorexia nervosa OR bulimia nervosa OR eating disorders OR binge eating OR insomnia OR sleep OR nightmares OR circadian OR panic OR substance OR abuse OR cannabis OR tobacco OR alcohol OR amphetamine OR hallucinogens OR heroin))]. We included the empirical research that reported quantitative data from recruited participants.

2.3.2 Article screening

The search returned 1194 articles published between 1997 and 2024 (final search performed in June 2024). From these, 199 articles passed the abstract check and 67 fitted the inclusion criteria. Six additional studies were also identified through citation chains, which resulted in 73 studies being included. Figure 2.2 presents a PRISMA diagram (Page et al., 2021) summary of the search process.

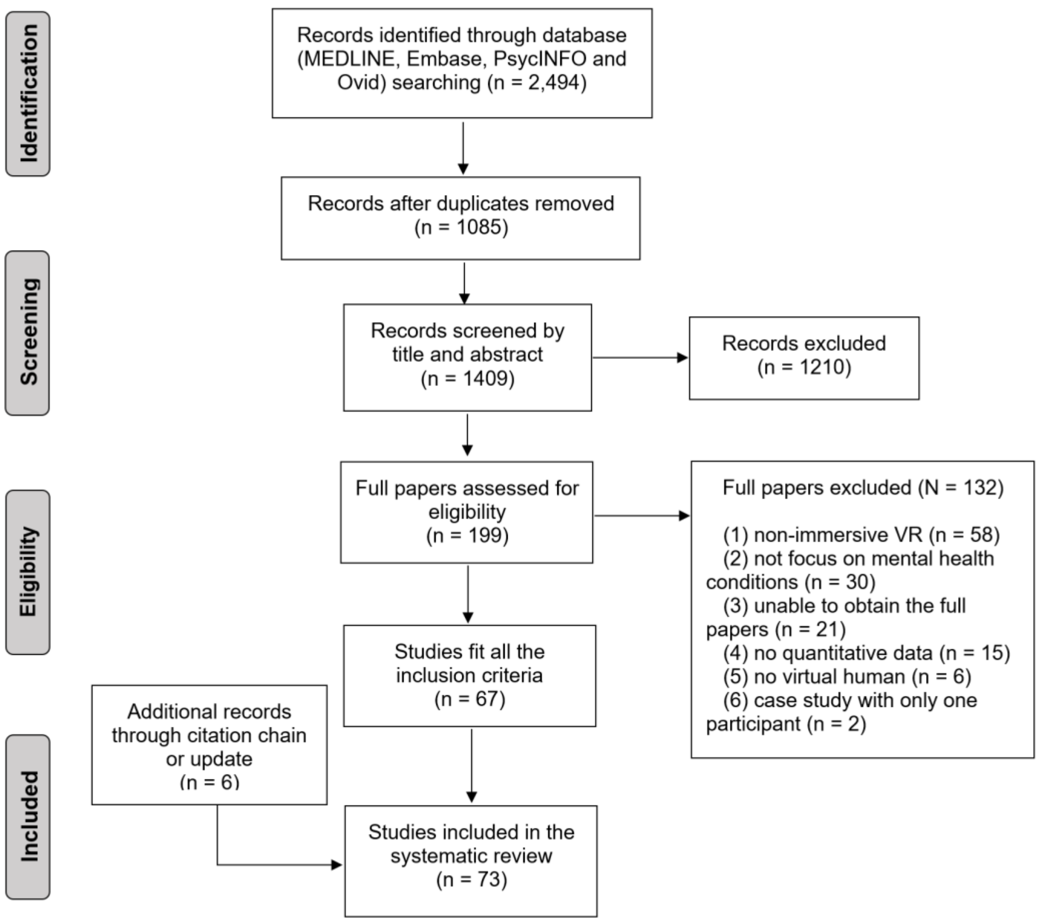


Figure 2.2. PRISMA flow diagram. Study screening flowchart of the systematic review.

2.4 RESULTS

A total of 73 studies meeting the inclusion criteria were found. Summaries of the studies can be seen in in Table 2.1 - Table 2.3.

VHs were used in the following research areas in mental health: social anxiety (n = 18), eating disorders (n = 18), psychosis (n = 15), substance abuse (n = 4), phobias (n = 3), depression (n = 3), autism (n = 3), post-traumatic stress disorder (PTSD) (n = 2), dissociative disorders (n = 1) and others without specific mental health conditions (n = 6).

2.4.1 Roles of VHs

The roles of virtual humans in the included studies could be categorised into active social interaction partners (n = 37), virtual crowds (n = 13), and virtual bodies (n = 23). We organise the studies based on our previous taxonomy as it offers a clear and functional framework to classify the use of VHs in mental health research. For studies that overlapped across multiple roles, we categorised them according to their primary role or the role with the highest level of interactivity.

Active social interaction partners

In thirty-seven studies VHs served primarily as active social interaction partners. These studies tested these VHs in three different ways: (1) manipulating specific VH characteristics to assess their influence, (2) comparing the responses of clinical and non-clinical individuals, and (3) validating whether VHs in VR could trigger participants' responses or replicate the effects of real-world experiences.

Ten studies explicitly examined VH characteristics. Wieser et al. (2010) investigated the anxiety-inducing effects of VH gender and gaze behaviours.

Compared to female agents, male agents with direct gaze elicited stronger anxiety and heart rate elevation among socially anxious women. A similar gender-related effect was observed in a VR fear conditioning program by Reichenberger et al. (2019), where female patients reported higher anxiety levels in response to aversive behaviours (e.g. unpleasant air blasts) from male VAs than from female VAs. Park et al. (2009) and Nijman et al. (2020) focused on non-verbal behaviours of VHs, where they implemented different facial expressions on VHs to help individuals with psychotic disorders practise emotion recognition.

Researchers also manipulated the dialogue responses (positive or negative) (Hartanto et al., 2014) and responsive frequency (high or low) (Fornells-Ambrojo et al., 2016) in conversation agents. Both the attitude and reply frequency of the VHs in a conversation were shown to affect participants' behavioural (e.g. interpersonal distance) and physiological (e.g. heart rate) responses. In another study by Slater et al. (2019), Participants engaged in a self-conversation by switching between two virtual bodies. They first embodied a VH resembling themselves to explain their problem to a virtual psychologist modelled after Sigmund Freud. Then, embodying Freud's VH, they provided advice to their virtual self. This approach resulted in a better perception of help compared to scripted dialogue. Furthermore, Pan, Gillies and Slater (2015) explored the impact of VH personality traits (confident vs. shy) on participants' perceptions, showing that the shy VH was perceived as more positive and friendly. Last the perceived agency of VHs was studied by Kothgassner et al. (2019, 2021), who compared participants' reactions when they believed they were interacting with a human versus a computer. Both experiments indicated that interactions with VHs can be as effective as real-life counterparts, provided participants believe the VH is controlled by a human in real time.

In another use case, five studies used active interaction partners to assess the responses differences between the clinical and non-clinical populations. In a conversation involving three people (two VHs and the participant), Han et al. (2014)

observed that patients with schizophrenia avoided eye contact more frequently during interactions with the VHs. Lee et al. (2021) used a ball game to study self-disturbances (e.g., a distorted perception of one's own body boundary) from schizophrenia. When a virtual man threw a ball to the participants, patients with schizophrenia exhibited longer reaction times in picking up the ball and had a weaker sense of body boundary compared to the non-clinical group. Robitaille et al. (2017) designed a scenario to detect executive functioning difficulties in military personnel with traumatic brain injury. In a military patrol scene, a virtual male dressed as a civilian was used to simulate a divided attention task by approaching the participant. Both groups tolerated the experience, but those with post-traumatic brain injury displayed navigation difficulties following the addition of the divided attention task. Lastly, Fusaro et al. (2023) and Artiran et al. (2024) studied the differences in responses to VHs between participants with autism and non-clinical controls. People with autism maintained a greater interpersonal distance from the VHs and displayed fewer responses to VHs' social cues (e.g. head turning and direct eye gaze) in virtual interview scenarios.

Twenty-two studies used active interaction partners to understand whether VHs in VR could trigger participants' responses or replicate the effects of real-world experience. Ten studies involved participants with mental health conditions or specific phobias. Three of them were related to psychosis. Du Sert et al. (2018) tested a therapy for auditory verbal hallucinations (AVH) in psychosis. Patients interacted with a VH representing the harmful voice they heard, and this exposure reduced their AVH-related depressive symptoms. Freeman et al. (2022) developed gameChange, an automated VR therapy designed to address agoraphobia in psychosis. Patients were guided by a virtual coach through everyday scenarios like visiting a GP or shopping, resulting in significant anxiety reduction, particularly in those with severe symptoms. More recently, Freeman et al. (2024) described an ongoing study Phoenix, in which a virtual coach named Farah helps young psychosis patients build self-confidence by guiding them through a series of tasks, offering instructions and encouragement. Two other studies focused on eating disorders. Bektas et al. (2023) used a virtual kitchen scene with anorexia nervosa patients, finding that a VH encouraging patients to challenge their disorder elicited stronger disgust reactions towards food. Natali et al. (2024) used the same scenario to induce positive mood, showing that supportive VHs can reduce anxiety and improve mood in anorexia patients. Regarding phobias, Freeman et al. (2018) conducted a clinical trial for individuals with a fear of heights. A VR coach called Nic guided the participant through the different stages of the experience, leading to a significant reduction in their fear. Similarly, Miloff et al. (2020) used a holographic-style virtual coach to guide patients confront their fear of spiders, with positive relationships toward the coach predicting better outcomes. Falconer et al. (2016) explored self-compassion in depression. Participants first delivered compassionate lines to a virtual child and then embodied the child, watching their compassionate actions from a third-person perspective, which fostered self-compassion. Amaral et al. (2018) conducted a clinical trial using VHs for social and cognition training for autistic individuals. A further study (Giguère et al., 2023) applied VR to treat addiction to cannabis. Participants interacted with a VH representing a key figure in their cannabis consumption, such as a friend, a drug dealer or those triggering their cravings, showing potential to reduce usage.

For the remaining twelve validation studies with participants from the non-clinical populations, five focused on social anxiety. Kwon, Powell and Chalmers (2013) compared the effects of displaying VHs on a desktop monitor versus an immersive VR system, finding that immersive VR elicited stronger anxiety potentially due to the higher sense of presence. This was further supported by a VR self-compassion induction (Ryan & Griffin, 2016) and a joint attention study by Seo et al. (2019). Both studies concluded that VR had a greater capacity to induce anxiety compared to non-immersive setups, with better training outcomes in reducing anxiety levels. Powers et al. (2013) examined participants' fear levels during VR and real-life interactions, noting that VR interactions triggered fear and anxiety similar to real-life situations. Quintana et al. (2019) investigated the effects of smells on perceptions of VHs. They discovered that exposure to fear-related odorants increased anxiety and reduced the trust towards the virtual character.

Halim et al. (2023) tested a VR system designed to improve depression by allowing users to practice giving and receiving compassion with a VH. It demonstrated early benefits in enhancing self-compassion and reducing depressive symptoms. Brander et al. (2021) examined the usability of a therapeutic VR tool for treating auditory hallucinations (AH), presenting it to psychiatric hospital staff. The tool allows operators to customise the appearance and behaviour of VHs, to whom patients attribute their AH. Additionally, Guldager et al. (2022) introduced VR FestLab, a party simulation game where adolescents practice refusing alcohol in peer-pressure scenarios with VHs. A similar approach was also employed by Shin et al. (2018) to address gaming addiction and by Lee et al. (2004) to help with nicotine addiction, where participants practiced regulating their cravings to refuse invitations to play or smoke from other characters. Lastly, VR was used to promote subjective well-being and self-confidence among young people by introducing positive perspectives that helped them acknowledge their strengths (Kim et al., 2020) and reducing self-criticism (Hidding et al., 2024) through structured conversation with virtual characters.

Virtual crowds

There were thirteen studies that used a virtual crowd. Six of them were related to anxiety. In one of the earliest study carried out, Slater et al. (1999) used a virtual audience and participants gave a speech in front of them. The study compared the performance of the speaker giving a speech in two different versions of the scenario, one where the audience was listening and the other showed a disruptive behaviour. The results provided early evidence that a virtual crowd can trigger levels of anxiety comparable to a real-life situation, and participants performed better when the audience was more listening. Similar findings were shown in the study by Takac et al. (2019), they suggested that the distress created in virtual public speaking scenario could have lingering effects after the VR session. Virtual crowds were also used to populate a virtual office, finding that watching other VHs working was effective in triggering work-related stress (Tarnanas et al., 2003), or creating peer effects to make people work more efficiently (Gürerk et al., 2019). Brinkman et al. (2011) explored the effect of the size and ethnicity of virtual crowds in a bar scene, showing that the VR exposure to the increased density and proportion of VHs with other ethnicity could induce stronger distress. Additionally, Shiban et al. (2015) incorporated air blasts and the sound of a female character screaming when participants approached a virtual crowd to study how individuals develop social fears and how these fears can be diminished or overcome through repeated exposure.

Four studies used a virtual crowd for research in psychosis. Early tests examined people's appraisals of neutral characters in a library scene. They found that virtual crowds can effectively elicit persecutory thoughts and these results can help to understand psychosis (Freeman et al., 2003, 2005). Freeman et al. (2008) later used an underground train scene to further explore people's paranoid interpretations and the cognitive factors behind paranoid thinking, showing paranoia could be predicted by anxiety, worry, perceptual anomalies and cognitive inflexibility. Atherton et al. (2016) also used a similar scenario of a train to understand the factors that contribute to paranoia, finding that participants' own self-confidence levels have an impact on their paranoid thoughts towards the virtual characters.

Mountford, Tchanturia and Valmaggia (2016) used a virtual bus ride to study eating disorders. Participants were tested for body image satisfaction, and the results showed that exposure to the virtual scenario with virtual humans (VHs) did not necessarily change participants' feelings or perceptions about their body image. Cho et al. (2008) designed a VR bar to study alcohol craving and found that the introduction of drinking characters led to a stronger drinking desire compared to simply displaying the alcohol in the scene. Rizzo and colleagues (2010) used virtual soldiers and Middle Eastern civilians to simulate Iraqi military experiences for training. Service members suffering from trauma were exposed to the VR environment, which resulted in a reduction of PTSD severity for 80% of the participants who completed the study.

Virtual bodies

VHs were used as own body representation in twenty studies – fifteen of them were related to eating disorders. Riva and Melis (1997) conducted first study found in the literature of body image representation using VR. Participants were asked to choose the virtual figures that best represented their current and the ideal body size. Perpiñá et al. (1999) introduced a feature allowing patients with eating disorders to present their body image by scaling the virtual body. Both VR programs demonstrated the feasibility of using virtual figures to understand body image disturbances. It was also suggested that the use of a virtual body as reference can provide comparable evaluation results to traditional paper-based measurements but with a higher level of engagement (Fisher et al., 2020). While there was no intention to elicit the illusion of body ownership in the above studies, many eating disorders studies aimed to make participants regard the virtual figures as their own.

The studies that involved body ownership illusion paradigm showed that when embodied in a larger figure, people tended to check their virtual body more often (Porras-Garcia et al., 2018; Porras-Garcia et al., 2019a) and experienced higher degree of self-dissatisfaction and anxiety (Ferrer-Garcia M. et al., 2017; Mendoza-Medialdea et al., 2023; Porras-Garcia et al., 2019b, 2020; Provenzano et al., 2020). In contrast, Schroeder et al. (2023) found that body dissatisfaction was not necessarily affected by the weight of the avatar, although people with higher body dissatisfaction exhibited more body-checking behaviours on weight-related parts.

Ferrer-Garcia et al. (2018) created a paradigm to reduce anxiety related to the body image in female students by making them return to the normal-size virtual figure after experiencing a larger body. A similar method was also implemented to ease the body image distortion from both healthy-weight participants and participants with obesity (Scarpina et al., 2019). Wolf and colleagues (2021) compared the bodyweight perceptions of two groups of females - those who embodied a VH and those who observed the VH as other people's avatar. The self-embodiment group underestimated the bodyweight of the VH compared to those who rated the VH as another's avatar. Furthermore, Ascione et al. (2023) conducted a body-related attention bias modification task to reduce excessive visual focus on weight-related areas of a self-represented VR body, which showed significant effects in reducing body dissatisfaction.

Three other studies used virtual bodies for psychosis research. Spanlang et al. (2018) examined whether self-fragmentation in VR, where participants embodied a VH while retaining physical presence in the real world, impacted the biological responses in patients with schizophrenia. Yamamoto and Nakao (2022) explored different association relationships between the participants and virtual characters finding that perceiving a fake body as one's own reduced the sense of body ownership in individuals with depersonalization. Gorisse et al. (2021) used a virtual body-double resembling each participant, created by scanning them and generating a digital figure for a social VR task. They concluded that observing a body-double engaging in VR social interactions could reduce persecutory thoughts during the task.

Additionally, Van Gelder et al. (2022) demonstrated that alternating convicted offenders between their current and aged future selves reduced self-defeating behaviours. Similarly, Vahle and Tomasik (2022) found that the embodiment of an older increased social motivation among young adults. Besides, Burin et al. (2022) presented participants with a moving character from a first-person and third-person perspective. They concluded that body illusion from the first-person perspective created stronger physiological activation to help people practise stress coping. Furthermore, Aymerich-Franch et al. (2014) assessed the impact of the facial similarity of a VH on participants, finding that embodying a figure with a dissimilar face reduced anxiety during VR presentations.

2.4.2 VA and avatar

Out of 73 studies, 68 involved the use of VAs. Nine of them adopted a semi-autonomous agent design, where facilitators manually triggered an audio effect (du Sert et al., 2018; Slater et al., 1999), adjusted VH voices (Giguère et al., 2023) or played pre-recorded animations (Fornells-Ambrojo et al., 2016; Powers et al., 2013) to customise characters' behaviours in real time.

Regarding the interaction types of VAs, around half of them (35/68) engaged in explicit interactions with the participants, primarily serving as active interaction partners. In eight studies, VAs had implicit interactions (e.g. walking towards or looking at participants without responses to participants' behaviours). Passive interactions were observed in 26 studies where VHs were used as virtual crowds and virtual bodies.

In the reviewed studies, the term 'avatar' was often used to describe any type of VH, even when it was not controlled by a human in real time or did not represent the participant's own body. Only five studies specifically used VHs exclusively as self-avatars for the participants. For example, Aymerich-Franch et al (2014) used an avatar with varying levels of visual similarity to the participant's face to understand the self-embodiment experience. Additionally, twelve studies combined the use of agents and avatars to create perspective changes through body-swapping experiences in social interactions (Falconer et al., 2016; Slater et al., 2019; Van Gelder et al., 2022) or to explore the effects of vicarious agency (the perception of control over the VH) (Gorisse et al., 2021). Kothgassner et al. (2019, 2021) also examined the effect of perceived agency by comparing participants' VR experiences with VAs and avatars. The results indicated that avatars could elicit stronger emotional responses compared to VAs.

2.4.3 Human characteristics

Forty-one out of the 73 studies in our review examined the effects of manipulating specific characteristics of VHs. The manipulated features included their visual appearance (e.g. gender, ethnicity, body size), interaction behaviour (e.g. interpersonal distance, eye gaze behaviour patterns and responsiveness towards participants), agency (VA or avatar), emotions, and personality. The most researched traits were their body size (n = 13), level of attention towards the participants (n = 5), and age and gender (n = 5). The exploration of VHs' emotional attributes, such as facial expressions (Nijman et al., 2020; Park et al., 2009) or personality (Pan et al., 2015) is limited. These manipulations aimed to understand how different physical and behavioural characteristics of VHs influence participant reactions, engagement, and emotional responses.

In addition to physical and behavioural traits, four studies explored the effects of verbal behaviour through various aspects such as dialogue content (Slater et al., 2019), attitude (Han et al., 2014; Hartanto et al., 2014) and the voice of the VHs (Giguère et al., 2023).

Besides individual character traits, two studies also tested the effects of crowd size and density as independent factors in virtual crowd settings (Brinkman et al., 2011; Takac et al., 2019). These studies aimed to explore how variations in the number and arrangement of virtual characters influence participants' sense of presence and anxiety in crowded environments.

Citation	Role	Agency	Virtual Agent Interaction Type	Parameters for manipulation	Condition	N	Participant characteristics	Comment on findings
Lee et al. (2004)	Active interaction partner	Agent	Implicit Interaction	-	Substance abuse	16	Males who smoked at least 10 cigarettes a day	VHs in cue exposure therapy exerted stronger impact compared to seductive objects.
Park et al. (2009)	Active interaction partner	Agent	Explicit Interaction	Gender; Emotional facial expression	Psychosis	54	Patients diagnosed with schizophrenia and non-clinical general population	VHs with emotional facial expression successfully triggered emotional and biological response.
Wieser et al. (2010)	Active interaction partner	Agent	Implicit Interaction	Gender; Gaze pattern; Interpersonal distance	Social anxiety	39	Females with social anxiety	Women paid more attention to the weight gained areas on VH than men.
Kwon, Powell and Chalmers (2013)	Active interaction partner	Agent	Explicit Interaction	-	Social anxiety	20	Students with social anxiety	VHs provoked both stronger anxiety and presence than those in non-immersive display.
Powers et al. (2013)	Active interaction partner	Semi-autonomous agent	Explicit Interaction	-	Social anxiety	26	Female undergraduate	Encountering VHs could trigger strong anxiety and fear.

Citation	Role	Agency	Virtual Agent Interaction Type	Parameters for manipulation	Condition	N	Participant characteristics	Comment on findings
Han et al. (2014)	Active interaction partner	Agent	Explicit Interaction	Attitude (positive, negative dialog)	Psychosis	45	Patients with schizophrenia and non-clinical general population	Patients with schizophrenia had active avoidance of eye contact with VHs compared to the healthy controls.
Hartanto et al. (2014)	Active interaction partner	Semi-autonomous agent	Explicit Interaction	Attitude (positive, negative dialog)	Social anxiety	24	General population	Positive dialogue feedback from VH resulted in less anxiety, lower heart rate, and longer answers.
Pan, Gillies and Slater (2015)	Active interaction partner	Semi-autonomous agent	Explicit Interaction	Personality (confident, shy)	Social anxiety	24	Male participants	VH with a shy personality can be perceived as more friendly.
Fornells-Ambrojo et al. (2016)	Active interaction partner	Semi-autonomous agent	Explicit Interaction	Responsiveness in social interaction	Psychosis	61	Male participants	Highly contingent VH were perceived more trustworthy for extremely paranoid individuals.
Ryan and Griffin (2016)	Active interaction partner	Agent	Implicit Interaction	-	Social anxiety	27	Students	VR exposure to a VH could successfully triggered social anxiety.

Citation	Role	Agency	Virtual Agent Interaction Type	Parameters for manipulation	Condition	N	Participant characteristics	Comment on findings
Falconer et al. (2016)	Active interaction partner	Agent and/or avatar	Explicit Interaction	-	Depression	15	Patients with depression	Changing the viewer perspective in a virtual social interaction could decrease self-criticism and increase self-compassion.
Robitaille et al. (2017)	Active interaction partner	Agent and/or avatar	Explicit Interaction	Responsiveness in social interaction	PTSD	12	Non-clinical military members and military members with mild traumatic brain injury	Participants with PTSD displayed a lack of navigational behaviour during a VH interaction.
Amaral et al. (2018)	Active interaction partner	Agent	Implicit Interaction	-	Autism	15	Patients with high-functioning ASD	VH affected social attentions among patients with autism.
Du Sert et al. (2018)	Active interaction partner	Semi-autonomous agent	Explicit Interaction	-	Psychosis	15	Patients diagnosed with schizophrenia	VR therapy with VH could improve auditory verbal hallucinations and depressive symptoms.
Freeman et al. (2018)	Active interaction partner (virtual coach)	Agent	Explicit Interaction	-	Phobias	100	Individuals with fear of height	A virtual therapist could help automate the VR therapy.

Citation	Role	Agency	Virtual Agent Interaction Type	Parameters for manipulation	Condition	N	Participant characteristics	Comment on findings
Shin et al. (2018)	Active interaction partner	Agent		Different levels of social pressure	Others	64	General population and patients with internet gaming disorder	VHs elicited cravings in gaming in gaming addiction exposure.
Kothgassner et al. (2019)	Active interaction partner	Agent and/or avatar	Explicit Interaction	Perceived agency (avatar VS agent)	Social anxiety	56	Students	Virtual social support from a VH human could be effective when the recipient thought it was from a human.
Quintana et al. (2019)	Active interaction partner	Agent	Explicit Interaction	-	Social anxiety	53	Female adults	Perception towards the VH could be affected by odorants manipulation.
Reichenberger et al. (2019)	Active interaction partner	Agent	Explicit Interaction	Gender; Interaction pattern (with or without aversive behaviour)	Social anxiety	60	General population	VH could be used to learn about social fear, and women reported higher fear compared to men.
Slater et al. (2019)	Active interaction partner (virtual coach)	Agent and/or avatar	Explicit Interaction	Dialogue contents (self-dialogue VS scripted dialogue)	Depression	58	General population	Self-dialogue VR coaching had better results than the scripted VR counselling.

Citation	Role	Agency	Virtual Agent Interaction Type	Parameters for manipulation	Condition	N	Participant characteristics	Comment on findings
Seo et al. (2019)	Active interaction partner	Agent	Explicit Interaction	Interaction pattern (combination of gaze and pointing)	Social anxiety	33	General population	VHs with non-verbal behaviour could be used to measure and promote joint attention.
Miloff et al. (2020)	Active interaction partner (virtual coach)	Agent	Explicit Interaction	-	Phobias	70	Patients with fear of spider	Alliance toward a virtual therapist is a significant predictor of treatment outcome in VR.
Nijman et al. (2020)	Active interaction partner	Semi-autonomous agent	Explicit Interaction	Facial expression	Psychosis	22	Patients with psychotic disorder	VHs' facial expressions can be utilised to train participants' social cognition.
Kim et al. (2020)	Active interaction partner	Agent	Explicit Interaction	-	Others	36	Non-clinical male volunteers	VHs could be used for assessment and intervention to promote subjective well-being.

Citation	Role	Agency	Virtual Agent Interaction Type	Parameters for manipulation	Condition	N	Participant characteristics	Comment on findings
Lee et al. (2021)	Active interaction partner	Agent and/or avatar	Explicit Interaction	-	Psychosis	48	Patients diagnosed with schizophrenia and non-clinical general population	The social context created by VHs impacted the difference in peripersonal space recognition.
Kothgassner et al. (2021)	Active interaction partner	Agent and/or avatar	Explicit Interaction	Perceived agency (avatar VS agent)	Social anxiety	84	Females	Social interaction experience with VHs was comparable to real-life experience of cyberbullying.
Brander et al. (2021)	Active interaction partner	Semi-autonomous agent	Explicit Interaction	Visual and behaviour traits of the virtual character through customization	Psychosis	109	Psychiatric hospital staff (psychotherapists, nursing staff, and administrators)	The function to customise VH for VR treatment was highly valued by clinical staff.
Guldager et al. (2022)	Active interaction partner	Agent	Explicit Interaction	-	Substance abuse	372	Students aged 15-18	VR party simulation could improve adolescents' drinking refusal skills.

Citation	Role	Agency	Virtual Agent Interaction Type	Parameters for manipulation	Condition	N	Participant characteristics	Comment on findings
Freeman et al. (2022)	Active interaction partner (virtual coach)	Agent	Explicit Interaction	-	Psychosis	346	Patients with psychosis	Automated VR therapy with VH could reduce anxious avoidance and distress in everyday situations.
Fusaro et al. (2023)	Active interaction partner	Agent and/or avatar	Explicit Interaction	-	Autism	53	General population and individuals with autism spectrum disorder	Autistic adults exhibited greater inter-personal distance with VHs.
Bektas et al. (2023)	Active interaction partner	Agent	Explicit Interaction	-	Eating Disorder	70	Patients with anorexia nervosa	The addition of a VH to the virtual kitchen scene could increase the association between food-specific state disgust and symptoms of eating disorders.
Giguère et al. (2023)	Active interaction partner	Semi-autonomous agent	Explicit Interaction	Voice and facial expressions	Substance abuse	19	Patients with cannabis use disorder and severe mental disorders	The VH intervention could reduce the cannabis use among people with cannabis use disorder and severe mental disorders.

Citation	Role	Agency	Virtual Agent Interaction Type	Parameters for manipulation	Condition	N	Participant characteristics	Comment on findings
Halim et al. (2023)	Active interaction partner	Agent and/or avatar	Explicit Interaction	-	Depression	36	Young adults	Changing the viewer perspective through a virtual body swapping experience can help enhance self-compassion and reduce depressive symptoms.
Artiran et al. (2024)	Active interaction partner	Agent	Explicit Interaction	Levels of attention (degrees of backchannels)	Autism	30	Patients with autism and non-clinical population	The behaviour of virtual interviewers influenced the gazes and head movements in individuals with autism.
Natali et al. (2024)	Active interaction partner	Agent	Explicit Interaction	-	Eating Disorder	145	Patients with anorexia nervosa	VHs could provide positive social support to patients with anorexia nervosa in a VR food exposure scene.

Citation	Role	Agency	Virtual Agent Interaction Type	Parameters for manipulation	Condition	N	Participant characteristics	Comment on findings
Hidding et al. (2024)	Active interaction partner	Agent	Explicit Interaction	-	Others	68	undergraduate students with high levels of self-criticism	Expressing compassion to a VH with similar self-criticism could reduce self-criticism and increase self-compassion.
Freeman et al. (2024)	Active interaction partner (virtual coach)	Agent	Explicit Interaction	-	Psychosis	11	Patients with non-affective psychosis and low positive self-beliefs, aged 16–26 years old	Automated VR therapy with a virtual coach could improve self-beliefs and well-being in young psychosis patients.

Table 2.1. Active social interaction partners.

Citation	Role	Agency	Virtual Agent Interaction Type	Parameters for manipulation	Condition	N	Participant characteristics	Comment on findings
Slater et al. (1999)	Virtual crowds	Semi-autonomous agent	Implicit Interaction	Attentional focus level	Social anxiety	10	General population	The use of virtual audience could treat public speaking fears; Higher perceived audience interest reduced public speaking anxiety.
Tarnanas, Tsoukalas and Stogiannidou (2003)	Virtual crowds	Agent	Passive Interaction	-	Social anxiety	120	General population and patients with work-related stress disorders	Virtual crowds triggered psychological and biological reactions in social anxiety.
Freeman et al. (2003)	Virtual crowds	Agent	Passive Interaction	-	Psychosis	24	General population	People tend to attribute their mental states to VHs in VR.
Freeman et al. (2005)	Virtual crowds	Agent	Passive Interaction	-	Psychosis	30	People cover the whole range of the Paranoia.	Neutral characters could elicit people's persecutory thoughts and be used to understand psychosis.
Freeman et al. (2008)	Virtual crowds	Agent	Passive Interaction	-	Psychosis	200	General population	Neutral characters could elicit paranoid ideation among general population.

Citation	Role	Agency	Virtual Agent Interaction Type	Parameters for manipulation	Condition	N	Participant characteristics	Comment on findings
Cho et al. (2008)	Virtual crowds	Agent	Implicit Interaction	-	Substance abuse	10	Population with a low level of alcohol dependence	The presence of a VH in cue exposure induced stronger cravings.
Rizzo et al. (2010)	Virtual crowds	Agent	Implicit Interaction	-	PTSD	20	Active duty service members	The VR environment with VHs could simulate thematic scene (i.e. war) for clinical treatment on PTSD and depression.
Brinkman et al. (2011)	Virtual crowds	Agent	Passive Interaction	Density; Ethnic appearance (white-European or North-African)	Social anxiety	26	General population and patients diagnosed with schizophrenia	VR exposure to the increased density and proportion of VHs with other ethnicity induced strong physiological arousal and distress.
Shiban et al. (2015)	Virtual crowds	Agent	Passive Interaction	-	Social anxiety	40	General population	VHs can be used to understand the learning and unlearning of social fears in VR.

Citation	Role	Agency	Virtual Agent Interaction Type	Parameters for manipulation	Condition	N	Participant characteristics	Comment on findings
Mountford, Tchanturia and Valmaggia (2016)	Virtual crowds	Agent	Passive Interaction	-	Eating disorder	18	General population	The exposure to the virtual scenario with VHs did not necessarily change people's in-state body image.
Atherton et al. (2016)	Virtual crowds	Agent	Passive Interaction	-	Psychosis	26	Males reporting paranoid ideation within the past month	People' own self-confidence could affect their paranoid thoughts towards the VHs.
Gürerk et al. (2019)	Virtual crowds	Agent	Passive Interaction	Different working behaviours	Social anxiety	108	General population	VHs with better working performance could motivate people to perform better.
Takac et al. (2019)	Virtual crowds	Agent	Implicit Interaction	Crowd size; Attentional focus level	Social anxiety	19	General population	Virtual audiences could elicit public speaking distress, which also lasted beyond VR session.

Table 2.2. Virtual crowds.

Citation	Role	Agency	Virtual Agent Interaction Type	Parameters for manipulation	Condition	N	Participant characteristics	Comment on findings
Riva and Melis (1997)	Virtual body	Agent	Passive Interaction	Body size	Eating disorder	119	General population	Virtual figures could be used to understand people's body images.
Perpiñá et al. (1999)	Virtual body	Agent	Passive Interaction	Body size	Eating disorder	13	Patients with eating disorders	VR body image treatment could have stronger impact than standard treatment.
Aymerich-Franch, Kizilcec and Bailenson (2014)	Virtual body	Avatar	Passive Interaction	Different levels of visual similarity to the participant	Social anxiety	187	General population	Having an avatar with an in-similar face could reduce anxiety.
Ferrer-Garcia et al. (2017)	Virtual body	Agent	Passive Interaction	Body size	Eating disorder	23	College students	Owning a fatter virtual body triggered higher levels of body anxiety and fear of weight gain.
Ferrer-Garcia et al. (2018)	Virtual body	Agent	Passive Interaction	Body size	Eating disorder	40	Female students	Owning a virtual body with different body sizes produced changes in body dissatisfaction.
Porras-Garcia et al. (2018)	Virtual body	Agent	Passive Interaction	Body size	Eating disorder	35	College students	Owning the larger-size virtual body increased people's body checking behaviours.
Porras-Garcia et al. (2019a)	Virtual body	Agent	Passive Interaction	Body size	Eating disorder	85	College students	Woman paid more attention to the weight gained areas on the virtual figure than man.

Citation	Role	Agency	Virtual Agent Interaction Type	Parameters for manipulation	Condition	N	Participant characteristics	Comment on findings
Porras-Garcia et al. (2019b)	Virtual body	Agent	Passive Interaction	Body size	Eating disorder	50	Undergraduates	Participants embodied through synchronous visuo-tactile technique had higher anxiety associated body dissatisfaction.
Mertens, Wagensveld and Engelhard (2019)	Virtual body	Agent	Passive Interaction	-	Phobias	48	General population and spider fearful individuals	Embodying participant into a virtual body from 1PP could improve the validity of fear conditioning.
Provenzano et al. (2019)	Virtual body	Agent	Passive Interaction	Body size	Eating disorder	20	Patients with eating disorders	Different perspectives (1PP vs 3PP) of viewing the virtual body could change people's body image perception.
Scarpina et al. (2019)	Virtual body	Agent	Passive Interaction	Body size	Eating disorder	30	General population and participants with obesity	Body ownership illusion could be created for both people with obesity and general population.
Spanlang et al. (2018)	Virtual body	Agent	Passive Interaction	Interaction pattern	Psychosis	27	General population	The body ownership illusion could trigger people's biological reactions.

Citation	Role	Agency	Virtual Agent Interaction Type	Parameters for manipulation	Condition	N	Participant characteristics	Comment on findings
Fisher et al. (2020)	Virtual body	Agent	Passive Interaction	Body size	Eating disorder	31	Female adolescents	Body size evaluation result from VR was comparable to the paper-based measurement.
Porras-Garcia et al. (2020)	Virtual body	Agent	Passive Interaction	Body size	Eating disorder	41	Female college students	People could have attentional bias towards the embodied virtual figure because of their own body dissatisfaction.
Gorisse et al. (2021)	Virtual body	Agent and/or avatar	Explicit Interaction	Interaction pattern (engagement level with the virtual crowds)	Psychosis	30	General population	Observing a body-double successfully engage in a VR social interaction could reduce paranoia.
Wolf et al. (2021)	Virtual body	Agent and/or avatar	Passive Interaction	-	Eating disorder	56	Females of normal weight	Embodiment through a VH could affect people's self-evaluation of their body weights.
Van Gelder et al. (2022)	Virtual body	Agent and/or avatar	Explicit Interaction	Age (current self and future self)	Others	24	Convicted male offenders	Offenders' interaction with a VH representing their future self could reduce their self-defeating behaviours.
Yamamoto and Nakao (2022)	Virtual body	Agent	Passive Interaction	Colour	Depersonalization	31	Male with various depersonalization tendencies	People with depersonalization had weaker body ownership of the self-representative figure.

Citation	Role	Agency	Virtual Agent Interaction Type	Parameters for manipulation	Condition	N	Participant characteristics	Comment on findings
Burin et al. (2022)	Virtual body	Agent and/or avatar	Passive Interaction	Embodiment perspectives (first person perspective or third person perspective)	Others	52	General population	Experiencing a moving virtual body from the 1PP could enhance stress coping ability.
Vahle and Tomasik (2022)	Virtual body	Avatar	-	Age (age-congruent and age-incongruent with the participants)	Others	74	General population aged between 18 and 30	The embodiment of an older avatar increased young adults' social motivation.
Mendoza-Medialdea et al. (2023)	Virtual body	Avatar	-	Body size	Eating disorder	57	College women	Manipulating the VH body size could produce changes in body dissatisfaction.
Ascione et al. (2023)	Virtual body	Avatar	-	-	Eating disorder	23	Adolescent patients with anorexia nervosa	The VR body attentional bias modification task could reduce body-related attentional bias among people with anorexia nervosa.
Schroeder et al. (2023)	Virtual body	Avatar	-	Body size	Eating disorder	43	Women with high or low body dissatisfaction	People with low body dissatisfaction show stronger disembodiment during self-avatar mirror exposure.

Table 2.3. Virtual bodies.

2.5 DISCUSSION

This chapter reviewed VR studies that included VHs in the scenarios to understand and treat mental health difficulties. We examined 73 studies and categorised them based on the primary role of the VHs - active social interaction partners, virtual crowds, and virtual bodies. We focused on the use of agency, interaction types and characteristics of the VH. VHs have served diverse purposes and demonstrated the ability to create psychological, behavioural, and physiological influences. However, very few studies provide detailed descriptions of the VHs, such as their visual appearance or behavioural capabilities. Without a clear understanding of the VHs' characteristics, it becomes difficult to assess their true impact or recreate experimental conditions. Therefore, the potential of each VH role requires further exploration, particularly by addressing these gaps and the associated challenges.

Active social interaction partners interact explicitly with participants. A long-standing challenge in this context has been enhancing the realism and responsiveness of these interactions, particularly during close-up engagements (Magnenat-Thalmann & Thalmann, 2005; Pan & Hamilton, 2018; Zell et al., 2019). Improving both visual fidelity and behavioural plausibility is essential for achieving lifelike interactions (Bombari et al., 2015; Gillies, 2018; Vinayagamoorthy et al., 2006). Consistent design of both visual and behavioural elements is needed to avoid the Uncanny Valley effect—where characters with realistic visual appearance evoke discomfort due to subtle imperfections (Mori et al., 2012), as also seen in VR interaction (Felnhofer et al., 2024; Stein & Ohler, 2017). Additionally, it is important to consider VH behavioural plausibility and consistency across multiple behavioural channels, including both verbal and non-verbal communication (e.g. body movement, facial expressions, voice, and dialogue content) (Chattopadhyay et al., 2020). For example, Choudhary et al. (2023) demonstrated that mismatches between facial expressions and vocal tone in VHs (e.g. a happy facial expression combined with an unhappy vocal tone) led to negative effects on participants' immersion and engagement, underscoring the importance of aligning these communication modes.

Regarding responsiveness and interactivity, current feedback in mental health VR often relies on pre-recorded actions triggered by participant inputs, such as their UI selection, position, or response time (Quintana et al., 2019; Robitaille et al., 2017; Tarnanas et al., 2003), or VH eye contacts that track the participant's location (Freeman et al., 2018, 2022). Future applications are likely to integrate machine learning algorithms, such as reinforcement learning, to create autonomous VHs that learn from experience (Dobre et al., 2022; Rovira & Slater, 2022). These VHs could adapt elements like ideal interpersonal distance and optimal attention levels to keep participants engaged. Algorithms could detect interaction dynamics (e.g., participant attitudes, responsiveness) and adjust VH behaviours in real time to meet specific goal, which would be particularly helpful in applications encouraging positive behaviours. Another area to explore is the integration of physiological monitoring to create adaptive interaction based on participants' physiological and affective states. Although a small number of reviewed studies used biometric data (e.g. salivary alpha-amylase, heart rate, galvanic skin) to evaluate physiological responses (Burin et al., 2022; Park et al., 2009; Spanlang et al., 2018), none has used it for enhancing interaction fidelity like others studies in computer science (Gupta et al., 2024; Jing et al., 2024). As this data can provide real-time estimates of the user's emotional arousal and valence, future implementations could focus on identifying and responding meaningfully to participants' emotional changes (Gupta et al., 2024).

A virtual coach represents a specific type of active social interaction partner, designed to guide and motivate users towards positive behavioural changes while facilitating the automation of therapy delivery. Focus group feedback indicated that users valued the convenience, accessibility, and anonymity provided by a virtual coach, but they also expressed concerns about the coach feeling unrelatable and difficult to engage with

(Venning et al., 2021). The development of a VR coach should incorporate knowledge and feedback from domain experts and targeted user groups to ensure it is tailored to the specific age range and mental health conditions it aims to address (Knight et al., 2021; Lambe et al., 2020). Generally, increasing the engagement and friendliness of communication from virtual coaches is likely to improve their effectiveness and the outcomes of VR therapy (Miloff et al., 2020).

In virtual crowds studies, it is standard practice to use identical designs and behaviours for characters that share common objectives to achieve a balance between realism and computational efficiency (Dickinson et al., 2019; Ulicny & Thalmann, 2001). However, visual and behavioural repetition within a group of VHs are very noticeable and decrease fidelity. It is therefore crucial to introduce variations in the appearance and movements of the characters to diversify the crowds and avoid repetitive patterns (Dong & Peng, 2019; Thalmann, 2000). The visual appearance can be easily diversified by randomising the characters' height, body shape, hair colour and style, and clothing. It is also important that their animations are not repetitive. Techniques to avoid undesired patterns in their motions include blending animations, using motion-capture data from multiple individuals, and implementing procedural animation systems that generate varied movements in real time based on the environment and the characters' interactions.

There are additional considerations regarding virtual crowds when they have a certain level of autonomy to move around. Each individual must move naturally through the environment, avoiding collisions with objects and other people, both static and dynamic obstacles. For example, Trivedi and Mousas (2023) implemented a crowd behaviour system for a VR street scene, where all the VAs in the virtual crowd have a navigation script with steering and pathfinding capabilities to avoid collisions. At the same time, each agent has its own customisable target location, walking speed and animations (e.g. head direction and gaze patterns). Other advanced virtual crowd control systems generate real-time behaviours driven by user input and the dynamics of other virtual entities within the crowd (Jain et al., 2024). These designs could also be applied in mental health applications to improve crowd interactions and better understand people's responses to busy environments with large numbers of people.

Studies involving virtual bodies predominantly helped individuals reassess their body image biases or gain new perspectives by owning another virtual body. These use cases require improved methods to establish and manipulate the sense of embodiment. Existing methods like visual-motor synchrony (Mertens et al., 2019) and visual-tactile synchrony (Porras Garcia et al., 2019; Scarpina et al., 2019) are designed to align visual feedback with participants’ movements or tactile sensations to enhance the illusion of ownership over the virtual body. While these methods have been effective in individually initiating a sense of virtual body ownership, sustaining or manipulating this sense throughout the VR experience presents a challenge. Issues such as mismatched sensory feedback or slight discrepancies in motion can lead to a breakdown in the feeling of embodiment (Cerritelli et al., 2021; Spanlang et al., 2014). Integrating other sensory modalities such as auditory (e.g. footsteps or voice modulation that correspond to the participant’s actions) or olfactory feedback (e.g. simulating scent experience based on the participant’s location) could perhaps enhance the body ownership illusion. It is also suggested that combined multisensory feedback strengthens and prolongs the influence on the sense of embodiment, affect, and behaviours (Flavián et al., 2021; Leonardis et al., 2014).

Customising the virtual body to closely resemble each participant can enhance the sense of embodiment. Aymerich-Franch et al. (2014) found that when participants embodied in a virtual body with similar face as themselves they experienced stronger self-association (feeling that the body belonged to them or relate to the VH in a stronger way) and higher anxiety in the public speaking task compared to that were dissimilar. This underlines the importance of avatar personalisation and its potentials to modify the extent of elicited emotions to achieve certain goals. This customisation should consider not only physical appearance but also the features such as participant's ethnic background (Banakou et al., 2020) and voice (Kao et al., 2021), as these factors can significantly impact the experience of body ownership.

2.5.1 Limitations

There are several limitations to this systematic review. First, our search strategy could not capture all VR mental health studies that used VHs in their scenarios. We examined four databases using title and abstract searches with specific keyword combinations, so any study that did not simultaneously address VR, VHs, and mental health conditions was likely to be overlooked. We identified six additional studies through citation chaining that were not included in the initial query results. Second, there may be limitations in our categorisation of the roles and characteristics of VHs. Existing taxonomies used to classify virtual characters and their behaviours (Saberi et al., 2014; Vinayagamoorthy et al., 2006) are at least ten years old, and none specifically focus on VR studies in mental health. We created our own classification according on the usage of the VHs in the reviewed studies. This classification would likely need to be adapted to cover other research fields outside the context of mental health. Third, regarding result analysis, we opted for a narrative review approach instead of a quantitative evaluation to measure the quality and effectiveness of VHs across studies. This decision was made given the variability in study objectives (ranging from validation and assessment to clinical treatments) and the diversity in sample sizes. The lack of detailed presentations and information on virtual characters in most mental health VR studies also complicated this evaluation. The design of a rigorous comparative matrix study would be nearly infeasible, and it is beyond the scope of this thesis.

2.6 CONCLUSIONS

VHs play a central role in several fields of VR mental health research, yet many studies lack comprehensive descriptions of their characteristics. In summary, VHs can be classified as active interaction partner, part of a virtual crowd, and virtual body. They are predominantly controlled by computer rather than by a person in real time, with the perception of agency (agent vs. avatar) influencing participants' actions and emotions. While the impact of VH characteristics varies, most studies have focused primarily on body size and gender, leaving an insufficient understanding of their emotional expressions and personality. Future research should systematically examine the full range of VH features, and researchers are encouraged to provide more detailed depictions of VHs, including multimedia (e.g. images and videos) to enhance replicability and comprehension.

REFERENCES

Amaral, C., Mouga, S., Simões, M., Pereira, H. C., Bernardino, I., Quental, H., Playle, R., McNamara, R., Oliveira, G., & Castelo-Branco, M. (2018). A Feasibility Clinical Trial to Improve Social Attention in Autistic Spectrum Disorder (ASD) Using a Brain Computer Interface. *Frontiers in Neuroscience*, *12*, 477. https://doi.org/10.3389/fnins.2018.00477

Artiran, S., Bedmutha, P. S., & Cosman, P. (2024). Analysis of Gaze, Head Orientation, and Joint Attention in Autism With Triadic VR Interviews. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *32*, 759–769. https://doi.org/10.1109/TNSRE.2024.3363728

Ascione, M., Carulla-Roig, M., Miquel-Nabau, H., Porras-Garcia, B., Meschberger-Annweiler, F.-A., Serrano-Troncoso, E., Ferrer-Garcia, M., Moreno-Sánchez, M., & Gutierrez-Maldonado, J. (2023). Attentional Bias Modification Training Based on Virtual Reality and Eye Tracking in Anorexia Nervosa Patients. *Journal of Clinical Medicine*, *12*(18), 5932. https://doi.org/10.3390/jcm12185932

Atherton, S., Antley, A., Evans, N., Cernis, E., Lister, R., Dunn, G., Slater, M., & Freeman, D. (2016). Self-Confidence and Paranoia: An Experimental Study Using an Immersive Virtual Reality Social Situation. *Behavioural and Cognitive Psychotherapy*, *44*(1), 56–64. https://doi.org/10.1017/S1352465814000496

Aymerich-Franch, L., Kizilcec, R. F., & Bailenson, J. N. (2014). The Relationship between Virtual Self Similarity and Social Anxiety. *Frontiers in Human Neuroscience*, *8*, 944. https://doi.org/10.3389/fnhum.2014.00944

Banakou, D., Beacco, A., Neyret, S., Blasco-Oliver, M., Seinfeld, S., & Slater, M. (2020). Virtual body ownership and its consequences for implicit racial bias are dependent on social context. *Royal Society Open Science*, *7*(12), 201848. https://doi.org/10.1098/rsos.201848

Bektas, S., Natali, L., Rowlands, K., Valmaggia, L., Di Pietro, J., Mutwalli, H., Himmerich, H., Treasure, J., & Cardi, V. (2023). Exploring Correlations of Food-Specific Disgust with Eating Disorder Psychopathology and Food Interaction: A Preliminary Study Using Virtual Reality. *Nutrients*, *15*(20), Article 20. https://doi.org/10.3390/nu15204443

Bombari, D., Schmid Mast, M., Canadas, E., & Bachmann, M. (2015). Studying social interactions through immersive virtual environment technology: Virtues, pitfalls, and future challenges. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00869

Brander, M., Egger, S. T., Hürlimann, N., Seifritz, E., Sumner, R. W., Vetter, S., & Magnenat, S. (2021). Virtual Reality Human–Human Interface to Deliver Psychotherapy to People Experiencing Auditory Verbal Hallucinations: Development and Usability Study. *JMIR Serious Games*, *9*(2). https://doi.org/10.2196/26820

Brinkman, W.-P., Veling, W., Dorrestijn, E., Sandino, G., Vakili, V., & Gaag, M. (2011). Virtual Reality to study responses to social environmental stressors in individuals with and without psychosis. *Annual Review of CyberTherapy and Telemedicine*, 86–91. https://doi.org/10.3233/978-1-6-750-766-6-86

Burin, D., Cavanna, G., Rabellino, D., Kotozaki, Y., & Kawashima, R. (2022). Neuroendocrine Response and State Anxiety Due to Psychosocial Stress Decrease after a Training with Subject's Own (but Not Another) Virtual Body: An RCT Study. *International Journal of Environmental Research and Public Health*, *19*(10), 6340. https://doi.org/10.3390/ijerph19106340

Cassell, J. (2001). Embodied Conversational Agents: Representation and Intelligence in User Interfaces. *AI Magazine*, *22*(4), Article 4. https://doi.org/10.1609/aimag.v22i4.1593

Cerritelli, F., Chiera, M., Abbro, M., Megale, V., Esteves, J., Gallace, A., & Manzotti, A. (2021). The Challenges and Perspectives of the Integration Between Virtual and Augmented Reality and Manual Therapies. *Frontiers in Neurology*, *12*, 700211. https://doi.org/10.3389/fneur.2021.700211

Chattopadhyay, D., Ma, T., Sharifi, H., & Martyn-Nemeth, P. (2020). Computer-Controlled Virtual Humans in Patient-Facing Systems: Systematic Review and Meta-Analysis. *Journal of Medical Internet Research*, *22*(7), e18839. https://doi.org/10.2196/18839

Cho, S., Ku, J., Park, J., Han, K., Lee, H., Choi, Y. K., Jung, Y.-C., Namkoong, K., Kim, J.-J., Kim, I. Y., Kim, S. I., & Shen, D. F. (2008). Development and Verification of an Alcohol Craving–Induction Tool Using Virtual Reality: Craving Characteristics in Social Pressure Situation. *CyberPsychology & Behavior*, *11*(3), 302–309. https://doi.org/10.1089/cpb.2007.0149

Choudhary, Z., Norouzi, N., Erickson, A., Schubert, R., Bruder, G., & Welch, G. F. (2023). Exploring the Social Influence of Virtual Humans Unintentionally Conveying Conflicting Emotions. *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, 571–580. https://doi.org/10.1109/VR55154.2023.00072

Dickinson, P., Gerling, K., Hicks, K., Murray, J., Shearer, J., & Greenwood, J. (2019). Virtual reality crowd simulation: Effects of agent density on user experience and behaviour. *Virtual Reality*, *23*(1), 19–32. https://doi.org/10.1007/s10055-018-0365-0

Dobre, G. C., Gillies, M., & Pan, X. (2022). Immersive machine learning for social attitude detection in virtual reality narrative games. *Virtual Reality*, *26*(4), 1519–1538. https://doi.org/10.1007/s10055-022-00644-4

Dong, Y., & Peng, C. (2019). Real-Time Large Crowd Rendering with Efficient Character and Instance Management on GPU. *International Journal of Computer Games Technology*, *2019*, e1792304. https://doi.org/10.1155/2019/1792304

du Sert, O. P., Potvin, S., Lipp, O., Dellazizzo, L., Laurelli, M., Breton, R., Lalonde, P., Phraxayavong, K., O'Connor, K., Pelletier, J.-F., Boukhalfi, T., Renaud, P., &

Dumais, A. (2018). Virtual reality therapy for refractory auditory verbal hallucinations in schizophrenia: A pilot clinical trial. *Schizophrenia Research*, *197*, 176–181. https://doi.org/10.1016/j.schres.2018.02.031

Falconer, C. J., Rovira, A., King, J. A., Gilbert, P., Antley, A., Fearon, P., Ralph, N., Slater, M., & Brewin, C. R. (2016). Embodying self-compassion within virtual reality and its effects on patients with depression. *BJPsych Open*, *2*(1), 74–80. https://doi.org/10.1192/bjpo.bp.115.002147

Felnhofer, A., Knaust, T., Weiss, L., Goinska, K., Mayer, A., & Kothgassner, O. D. (2024). A Virtual Character's Agency Affects Social Responses in Immersive Virtual Reality: A Systematic Review and Meta-Analysis. *International Journal of Human–Computer Interaction*, *40*(16), 4167–4182. https://doi.org/10.1080/10447318.2023.2209979

Ferrer-Garcia M., Porras-Garcia B., Gonzalez-Ibanez C., Gracia-Blanes M., Vilalta-Abella F., Pla-Sanjuanelo J., Riva G., Dakanalis A., Achotegui-Loizate J., Talarn-Caparros A., Ribas-Sabate J., Andreu-Gracia A., Diaz-Marsa M., Monras-Arnau M., Serrano-Troncoso E., Treasure J., & Gutierrez-Maldonado J. (2017). Does owning a 'fatter' virtual body increase body anxiety in college students? *Annual Review of CyberTherapy and Telemedicine*, *15*, 147–153.

Ferrer-Garcia, M., Porras-Garcia, B., Moreno, M., Bertomeu, P., & Maldonado, J. G. (2018). Embodiment in different size virtual bodies produces changes in women's body image distortion and dissatisfaction. *Annual Review of CyberTherapy and Telemedicine*, *16*(Human Factors Engineering [4010]), 111–117.

Fisher, S., Abdullah, A., Charvin, I., Da Fonseca, D., & Bat-Pitault, F. (2020). Comparison of body image evaluation by virtual reality and paper-based figure rating scales in adolescents with anorexia nervosa: Retrospective study. *Eating and Weight Disorders: EWD*, *25*(3), 735–743. https://doi.org/10.1007/s40519-019-00680-1

Flavián, C., Ibáñez-Sánchez, S., & Orús, C. (2021). The influence of scent on virtual reality experiences: The role of aroma-content congruence. *Journal of Business Research*, *123*, 289–301. https://doi.org/10.1016/j.jbusres.2020.09.036

Fornells-Ambrojo, M., Elenbaas, M., Barker, C., Swapp, D., Navarro, X., Rovira, A., Sanahuja, J. M. T., & Slater, M. (2016). Hypersensitivity to Contingent Behavior in Paranoia: A New Virtual Reality Paradigm. *The Journal of Nervous and Mental Disease*, *204*(2), 148–152. https://doi.org/10.1097/NMD.0000000000000414

Freeman, D., Freeman, J., Ahmed, M., Haynes, P., Beckwith, H., Rovira, A., Miguel, A. L., Ward, R., Bousfield, M., Riffiod, L., Kabir, T., Waite, F., & Rosebrock, L. (2024). Automated VR therapy for improving positive self-beliefs and psychological well-being in young patients with psychosis: A proof of concept evaluation of Phoenix VR self-confidence therapy. *Behavioural and Cognitive Psychotherapy*, *52*(3), 277–287. https://doi.org/10.1017/S1352465823000553

Freeman, D., Garety, P. A., Bebbington, P., Slater, M., Kuipers, E., Fowler, D., Green, C., Jordan, J., Ray, K., & Dunn, G. (2005). The psychology of persecutory ideation II: A virtual reality experimental study. *The Journal of Nervous and Mental Disease*, *193*(5), 309–315. https://doi.org/10.1097/01.nmd.0000161686.53245.70

Freeman, D., Haselton, P., Freeman, J., Spanlang, B., Kishore, S., Albery, E., Denne, M., Brown, P., Slater, M., & Nickless, A. (2018). Automated psychological therapy using immersive virtual reality for treatment of fear of heights: A single-blind, parallel-group, randomised controlled trial. *The Lancet. Psychiatry*, *5*(8), 625–632. https://doi.org/10.1016/S2215-0366(18)30226-8

Freeman, D., Lambe, S., Kabir, T., Petit, A., Rosebrock, L., Yu, L.-M., Dudley, R., Chapman, K., Morrison, A., O'Regan, E., Aynsworth, C., Jones, J., Murphy, E., Powling, R., Galal, U., Grabey, J., Rovira, A., Martin, J., Hollis, C., … West, J. (2022). Automated virtual reality therapy to treat agoraphobic avoidance and distress in patients with psychosis (gameChange): A multicentre, parallel-group, single-blind, randomised, controlled trial in England with mediation and moderation analyses. *The Lancet Psychiatry*, *9*(5), 375–388. https://doi.org/10.1016/S2215-0366(22)00060-8

Freeman, D., Pugh, K., Antley, A., Slater, M., Bebbington, P., Gittins, M., Dunn, G., Kuipers, E., Fowler, D., & Garety, P. (2008). Virtual reality study of paranoid thinking in the general population. *The British Journal of Psychiatry*, *192*(4), 258–263. https://doi.org/10.1192/bjp.bp.107.044677

Freeman, D., Slater, M., Bebbington, P. E., Garety, P. A., Kuipers, E., Fowler, D., Met, A., Read, C. M., Jordan, J., & Vinayagamoorthy, V. (2003). Can Virtual Reality be Used to Investigate Persecutory Ideation? *The Journal of Nervous and Mental Disease*, *191*(8), 509–514. https://doi.org/10.1097/01.nmd.0000082212.83842.fe

Fusaro, M., Fanti, V., & Chakrabarti, B. (2023). Greater interpersonal distance in adults with autism. *Autism Research*, *16*(10), 2002–2007. https://doi.org/10.1002/aur.3013

Giguère, S., Potvin, S., Beaudoin, M., Dellazizzo, L., Giguère, C.-É., Furtos, A., Gilbert, K., Phraxayavong, K., & Dumais, A. (2023). Avatar Intervention for Cannabis Use Disorder in Individuals with Severe Mental Disorders: A Pilot Study. *Journal of Personalized Medicine*, *13*(5), Article 5. https://doi.org/10.3390/jpm13050766

Gillies, M. (2018). Creating Virtual Characters. *Proceedings of the 5th International Conference on Movement and Computing*, 1–8. https://doi.org/10.1145/3212721.3212835

Gorisse, G., Senel, G., Banakou, D., Beacco, A., Oliva, R., Freeman, D., & Slater, M. (2021). Self-observation of a virtual body-double engaged in social interaction reduces persecutory thoughts. *Scientific Reports*, *11*(1), 23923. https://doi.org/10.1038/s41598-021-03373-x

Guldager, J. D., Kjær, S. L., Grittner, U., & Stock, C. (2022). Efficacy of the Virtual Reality Intervention VR FestLab on Alcohol Refusal Self-Efficacy: A Cluster-Randomized Controlled Trial. *International Journal of Environmental Research and Public Health*, *19*(6), 3293. https://doi.org/10.3390/ijerph19063293

Gupta, K., Zhang, Y., Gunasekaran, T. S., Krishna, N., Pai, Y. S., & Billinghurst, M. (2024). CAEVR: Biosignals-Driven Context-Aware Empathy in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, 1–11. https://doi.org/10.1109/TVCG.2024.3372130

Gürerk, Ö., Bönsch, A., Kittsteiner, T., & Staffeldt, A. (2019). Virtual humans as co-workers: A novel methodology to study peer effects. *Journal of Behavioral and Experimental Economics*, *78*, 17–29. https://doi.org/10.1016/j.socec.2018.11.003

Halim, I., Stemmet, L., Hach, S., Porter, R., Liang, H.-N., Vaezipour, A., Henry, J. D., & Baghaei, N. (2023). Individualized Virtual Reality for Increasing Self-Compassion: Evaluation Study. *JMIR Mental Health*, *10*(1), e47617. https://doi.org/10.2196/47617

Han, K., Shin, J., Yoon, S. Y., Jang, D.-P., & Kim, J.-J. (2014). Deficient gaze pattern during virtual multiparty conversation in patients with schizophrenia. *Computers in Biology and Medicine*, *49*, 60–66. https://doi.org/10.1016/j.compbiomed.2014.03.012

Hartanto, D., Kampmann, I. L., Morina, N., Emmelkamp, P. G. M., Neerincx, M. A., & Brinkman, W.-P. (2014). Controlling Social Stress in Virtual Reality Environments. *PLOS ONE*, *9*(3), e92804. https://doi.org/10.1371/journal.pone.0092804

Hidding, M., Veling, W., Pijnenborg, G. H. M., & van der Stouwe, E. C. D. (2024). A single-session VR intervention addressing self-compassion and self-criticism with and without perspective change: Results of a randomized controlled experiment. *Behaviour Research and Therapy*, *173*, 104466. https://doi.org/10.1016/j.brat.2023.104466

Jain, D., Jain, U., Rao, K., Dhingra, D., Nagrath, P., Varshney, A. R., & Nagrath, P. (2024). A Real-Time Crowd Tracking and Control System Using Deep Learning. In A. E. Hassanien, S. Anand, A. Jaiswal, & P. Kumar (Eds.), *Innovative Computing and Communications* (pp. 587–601). Springer Nature. https://doi.org/10.1007/978-981-97-3817-5_42

Jorjafki, E. M., Sagarin, B. J., & Butail, S. (2018). Drawing power of virtual crowds. *Journal of The Royal Society Interface*, *15*(145), 20180335. https://doi.org/10.1098/rsif.2018.0335

Jing, A., Teo, T., McDade, J., Zhang, C., Wang, Y., Mitrofan, A., Thareja, R., Shin, H., Lee, Y., Gil, Y.-H., Billinghurst, M., & Lee, G. A. (2024). Superpowering Emotion Through Multimodal Cues in Collaborative VR. 2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 160–169. https://doi.org/10.1109/ISMAR62088.2024.00030

Kao, D., Ratan, R., Mousas, C., & Magana, A. J. (2021). The Effects of a Self-Similar Avatar Voice in Educational Games. *Proc. ACM Hum.-Comput. Interact.*, *5*(CHI PLAY), 238:1-238:28. https://doi.org/10.1145/3474665

Kilteni, K., Groten, R., & Slater, M. (2012). The Sense of Embodiment in Virtual Reality. *Presence: Teleoperators and Virtual Environments*, *21*(4), 373–387. https://doi.org/10.1162/PRES_a_00124

Kim, J., Jung, Y. H., Shin, Y.-B., Kim, M.-K., Eom, H., Kim, E., Kim, J., & Kim, J.-J. (2020). Development and Validation of a Virtual Reality-Based Training Program for Promoting Subjective Well-Being. *Psychiatry Investigation*, *17*(12), 1207–1215. https://doi.org/10.30773/pi.2020.0311

Knight, I., West, J., Matthews, E., Kabir, T., Lambe, S., Waite, F., & Freeman, D. (2021). Participatory design to create a VR therapy for psychosis. *Design for Health*, *5*(1), 98–119. https://doi.org/10.1080/24735132.2021.1885889

Kothgassner, O. D., Goreis, A., Glenk, L. M., Kafka, J. X., Beutl, L., Kryspin-Exner, I., Hlavacs, H., Palme, R., & Felnhofer, A. (2021). Virtual and real-life ostracism and its impact on a subsequent acute stressor. *Physiology & Behavior*, *228*, 113205. https://doi.org/10.1016/j.physbeh.2020.113205

Kothgassner, O. D., Goreis, A., Kafka, J. X., Kaufmann, M., Atteneder, K., Beutl, L., Hennig-Fast, K., Hlavacs, H., & Felnhofer, A. (2019). Virtual social support buffers stress response: An experimental comparison of real-life and virtual support prior to a social stressor. *Journal of Behavior Therapy and Experimental Psychiatry*, *63*, 57–65. https://doi.org/10.1016/j.jbtep.2018.11.003

Kwon, J. H., Powell, J., & Chalmers, A. (2013). How level of realism influences anxiety in virtual reality environments for a job interview. *International Journal of Human-Computer Studies*, *71*(10), 978–987. https://doi.org/10.1016/j.ijhcs.2013.07.003

Lambe, S., Knight, I., Kabir, T., West, J., Patel, R., Lister, R., Rosebrock, L., Rovira, A., Garnish, B., Freeman, J., M. Clark, D., Waite, F., & Freeman, D. (2020). Developing an automated VR cognitive treatment for psychosis: gameChange VR therapy. *Journal of Behavioral and Cognitive Therapy*, *30*(1), 33–40. https://doi.org/10.1016/j.jbct.2019.12.001

Lee, H.-S., Hong, S.-J. J., Baxter, T., Scott, J., Shenoy, S., Buck, L., Bodenheimer, B., & Park, S. (2021). Altered Peripersonal Space and the Bodily Self in Schizophrenia: A Virtual Reality Study. *Schizophrenia Bulletin*, *47*(4), 927–937. https://doi.org/10.1093/schbul/sbab024

Lee, J., Lim, Y., Graham, S. J., Kim, G., Wiederhold, B. K., Wiederhold, M. D., Kim, I. Y., & Kim, S. I. (2004). Nicotine craving and cue exposure therapy by using virtual environments. *Cyberpsychology & Behavior : The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society*, *7*(6), 705–713. https://doi.org/10.1089/cpb.2004.7.705

Leonardis, D., Frisoli, A., Barsotti, M., Carrozzino, M., & Bergamasco, M. (2014). Multisensory Feedback Can Enhance Embodiment Within an Enriched Virtual Walking Scenario. *Presence: Teleoperators and Virtual Environments*, *23*(3), 253–266. https://doi.org/10.1162/PRES_a_00190

Magnenat-Thalmann, N., & Thalmann, D. (2005). Virtual humans: Thirty years of research, what next? *The Visual Computer*, *21*(12), 997–1015. https://doi.org/10.1007/s00371-005-0363-6

Mendoza-Medialdea, M. T., Meschberger-Annweiler, F.-A., Ascione, M., Rueda-Pina, A., Rabarbari, E., Porras-Garcia, B., Ferrer-Garcia, M., & Gutiérrez-Maldonado, J. (2023). Body Dissatisfaction and Body-Related Attentional Bias: Is There a Causal Relationship? *Journal of Clinical Medicine*, *12*(17), Article 17. https://doi.org/10.3390/jcm12175659

Mertens, G., Wagensveld, P., & Engelhard, I. M. (2019). Cue conditioning using a virtual spider discriminates between high and low spider fearful individuals. *Computers in Human Behavior*, *91*, 192–200. https://doi.org/10.1016/j.chb.2018.10.006

Miloff, A., Carlbring, P., Hamilton, W., Andersson, G., Reuterskiöld, L., & Lindner, P. (2020). Measuring Alliance Toward Embodied Virtual Therapists in the Era of Automated Treatments With the Virtual Therapist Alliance Scale (VTAS): Development and Psychometric Evaluation. *Journal of Medical Internet Research*, *22*(3), e16660. https://doi.org/10.2196/16660

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine*, *19*(2), 98–100. IEEE Robotics & Automation Magazine. https://doi.org/10.1109/MRA.2012.2192811

Mountford, V. A., Tchanturia, K., & Valmaggia, L. (2016). 'What Are You Thinking When You Look at Me?' A Pilot Study of the Use of Virtual Reality in Body Image. *Cyberpsychology, Behavior and Social Networking*, *19*(2), 93–99. https://doi.org/10.1089/cyber.2015.0169

Natali, L., Meregalli, V., Rowlands, K., Di Pietro, J., Treasure, J., Collantoni, E., Meneguzzo, P., Tenconi, E., Favaro, A., Fontana, F., Ceccato, E., Sala, A., Valmaggia, L., & Cardi, V. (2024). Virtual food exposure with positive mood induction or social support to reduce food anxiety in anorexia nervosa: A feasibility study. *International Journal of Eating Disorders*, *57*(3), 703–715. https://doi.org/10.1002/eat.24155

Nijman, S. A., Veling, W., Greaves-Lord, K., Vos, M., Zandee, C. E. R., Aan het Rot, M., Geraets, C. N. W., & Pijnenborg, G. H. M. (2020). Dynamic Interactive Social Cognition Training in Virtual Reality (DiSCoVR) for People With a Psychotic Disorder: Single-Group Feasibility and Acceptability Study. *JMIR Mental Health*, *7*(8). https://doi.org/10.2196/17808

Nowak, K., & Biocca, F. (2003). The Effect of the Agency and Anthropomorphism on Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments. *Presence Teleoperators & Virtual Environments*, *12*, 481–494. https://doi.org/10.1162/105474603322761289

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ*, *372*, n160. https://doi.org/10.1136/bmj.n160

Pan, X., Gillies, M., & Slater, M. (2015). Virtual Character Personality Influences Participant Attitudes and Behavior – An Interview with a Virtual Human Character about Her Social Anxiety. *Frontiers in Robotics and AI*, *2*. https://doi.org/10.3389/frobt.2015.00001

Pan, X., & Hamilton, A. F. de C. (2018). Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology*, *109*(3), Article 3. https://doi.org/10.1111/bjop.12290

Park, I. H., Kim, J.-J., Ku, J., Jang, H. J., Park, S.-H., Kim, C.-H., Kim, I. Y., & Kim, S. I. (2009). Characteristics of social anxiety from virtual interpersonal interactions in patients with schizophrenia. *Psychiatry*, *72*(1), 79–93. https://doi.org/10.1521/psyc.2009.72.1.79

Perpiñá, C., Botella, C., Baños, R., Marco, H., Alcañiz, M., & Quero, S. (1999). Body Image and Virtual Reality in Eating Disorders: Is Exposure to Virtual Reality More Effective than the Classical Body Image Treatment? *CyberPsychology & Behavior*, *2*(2), 149–155. https://doi.org/10.1089/cpb.1999.2.149

Pertaub, D.-P., Slater, M., & Barker, C. (2002). An Experiment on Public Speaking Anxiety in Response to Three Different Types of Virtual Audience. *Presence: Teleoperators and Virtual Environments*, *11*(1), 68–78. https://doi.org/10.1162/105474602317343668

Porras Garcia, B., Ferrer Garcia, M., Olszewska, A., Yilmaz, L., González Ibañez, C., Gracia Blanes, M., Gültekin, G., Serrano Troncoso, E., & Gutiérrez Maldonado, J. (2019). Is This My Own Body? Changing the Perceptual and Affective Body Image Experience among College Students Using a New Virtual Reality Embodiment-Based Technique. *Journal of Clinical Medicine*, *8*(7). https://doi.org/10.3390/jcm8070925

Porras-Garcia, B., Ferrer-Garcia, M., Ghita, A., Moreno, M., López-Jiménez, L., Vallvé-Romeu, A., Serrano-Troncoso, E., & Gutiérrez-Maldonado, J. (2019). The influence of gender and body dissatisfaction on body-related attentional bias: An eye-tracking and virtual reality study. *International Journal of Eating Disorders*, *52*(10), 1181–1190. https://doi.org/10.1002/eat.23136

Porras-Garcia, B., Ferrer-Garcia, M., Yilmaz, L., Sen, Y. O., Olszewska, A., Ghita, A., Serrano-Troncoso, E., Treasure, J., & Gutiérrez-Maldonado, J. (2020). Body-related attentional bias as mediator of the relationship between body mass index and body dissatisfaction. *European Eating Disorders Review : The Journal of the Eating Disorders Association*, *28*(4), 454–464. https://doi.org/10.1002/erv.2730

Porras-Garcia, B., Ghita, A., Moreno, M., Ferrer-Garcia, M., Bertomeu-Panisello, P., Serrano-Troncoso, E., Riva, G., Dakanalis, A., Achotegui-Loizate, J., Talarn-Caparros, A., Andreu-Gracia, A., Treasure, J., Martinez-Mallen, E., Moreno-Perez, E., & Gutierrez-Maldonado, J. (2018). Gender differences in attentional bias after owning a virtual avatar with increased weight. *Annual Review of CyberTherapy and Telemedicine*, *16*(Human Factors Engineering [4010]), 73–79.

Powers, M. B., Briceno, N. F., Gresham, R., Jouriles, E. N., Emmelkamp, P. M. G., & Smits, J. A. J. (2013). Do conversations with virtual avatars increase feelings of social anxiety? *Journal of Anxiety Disorders*, *27*(4), 398–403. https://doi.org/10.1016/j.janxdis.2013.03.003

Provenzano, L., Porciello, G., Ciccarone, S., Lenggenhager, B., Tieri, G., Marucci, M., Dazzi, F., Loriedo, C., & Bufalari, I. (2020). Characterizing Body Image Distortion and Bodily Self-Plasticity in Anorexia Nervosa via Visuo-Tactile Stimulation in Virtual Reality. *Journal of Clinical Medicine*, *9*(1), Article 1. https://doi.org/10.3390/jcm9010098

Quintana, P., Nolet, K., Baus, O., & Bouchard, S. (2019). The Effect of Exposure to Fear-Related Body Odorants on Anxiety and Interpersonal Trust Toward a Virtual Character. *Chemical Senses*, *44*(9), 683–692. https://doi.org/10.1093/chemse/bjz063

Reichenberger, J., Pfaller, M., Forster, D., Gerczuk, J., Youssef Shiban, & Mühlberger, A. (2019). Men Scare Me More: Gender Differences in Social Fear Conditioning in Virtual Reality. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.01617

Riva, G., & Melis, L. (1997). Virtual Reality for the Treatment of Body Image Disturbances. *Studies in Health Technology and Informatics*, 95–112.

Rizzo, A. S., Difede, J., Rothbaum, B. O., Reger, G., Spitalnick, J., Cukor, J., & McLay, R. (2010). Development and early evaluation of the Virtual Iraq/Afghanistan exposure therapy system for combat-related PTSD. *Annals of the New York Academy of Sciences*, *1208*, 114–125. https://doi.org/10.1111/j.1749-6632.2010.05755.x

Rizzo, A. 'Skip', & Shilling, R. (2017). Clinical Virtual Reality tools to advance the prevention, assessment, and treatment of PTSD. *European Journal of Psychotraumatology*, *8*(sup5). https://doi.org/10.1080/20008198.2017.1414560

Robitaille, N., Jackson, P. L., Hébert, L. J., Mercier, C., Bouyer, L. J., Fecteau, S., Richards, C. L., & McFadyen, B. J. (2017). A Virtual Reality avatar interaction (VRai) platform to assess residual executive dysfunction in active military personnel with previous mild traumatic brain injury: Proof of concept. *Disability and Rehabilitation. Assistive Technology*, *12*(7), 758–764. https://doi.org/10.1080/17483107.2016.1229048

Rovira, A., & Slater, M. (2022). Encouraging bystander helping behaviour in a violent incident: A virtual reality study using reinforcement learning. *Scientific Reports*, *12*(1), Article 1. https://doi.org/10.1038/s41598-022-07872-3

Ryan, V., & Griffin, R. (2016). An Investigation Into Anxiety In Virtual Reality Following A Self-Compassion Induction. *Annual Review of CyberTherapy and Telemedicine*, *14*, pp 109-114.

Saberi, M., Bernardet, U., & DiPaola, S. (2014). *An Architecture for Personality-based, Nonverbal Behavior in Affective Virtual Humanoid Character | Elsevier Enhanced Reader*. https://doi.org/10.1016/j.procs.2014.11.104

Salber, D., & Coutaz, J. (1993). Applying the Wizard of Oz technique to the study of multimodal systems. In L. J. Bass, J. Gornostaev, & C. Unger (Eds.), *Human-Computer Interaction* (pp. 219–230). Springer. https://doi.org/10.1007/3-540-57433-6_51

Scarpina, F., Serino, S., Keizer, A., Chirico, A., Scacchi, M., Castelnuovo, G., Mauro, A., & Riva, G. (2019). The Effect of a Virtual-Reality Full-Body Illusion on Body Representation in Obesity. *Journal of Clinical Medicine*, *8*(9). https://doi.org/10.3390/jcm8091330

Schroeder, P. A., Gehrer, N. A., Reents, M., Reimer, N., Vagedes, J., & Svaldi, J. (2023). Body Dissatisfaction Directs Avatar Perception: Embodiment and Selective Visual Attention in Body Mass-Modified Self-Avatars. *Cyberpsychology, Behavior, and Social Networking*, *26*(11), 850–860. https://doi.org/10.1089/cyber.2022.0385

Seo, S., Kim, E., Mundy, P., Heo, J., & Kim, K. (Kenny). (2019). Joint Attention Virtual Classroom: A Preliminary Study. *Psychiatry Investigation*, *16*(4), 292–299. https://doi.org/10.30773/pi.2019.02.08

Shiban, Y., Reichenberger, J., Neumann, I. D., & Mühlberger, A. (2015). Social conditioning and extinction paradigm: A translational study in virtual reality. *Frontiers in Psychology*, *6*, 400. https://doi.org/10.3389/fpsyg.2015.00400

Shin, Y.-B., Kim, J.-J., Kim, M.-K., Kyeong, S., Jung, Y. H., Eom, H., & Kim, E. (2018). Development of an effective virtual environment in eliciting craving in adolescents and young adults with internet gaming disorder. *PLOS ONE*, *13*(4), e0195677. https://doi.org/10.1371/journal.pone.0195677

Slater, M., Neyret, S., Johnston, T., Iruretagoyena, G., Crespo, M. Á. de la C., Alabèrnia-Segura, M., Spanlang, B., & Feixas, G. (2019). An experimental study of a virtual reality counselling paradigm using embodied self-dialogue. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-46877-3

Slater, M., Pertaub, D.-P., & Steed, A. (1999). Public Speaking in Virtual Reality: Facing an Audience of Avatars. *Computer Graphics and Applications, IEEE*, *19*, 6–9. https://doi.org/10.1109/38.749116

Slater, M., & Wilbur, S. (1997). A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, *6*(6), 603–616. https://doi.org/10.1162/pres.1997.6.6.603

Spanlang, B., Nierula, B., Haffar, M., & Debruille, J. B. (2018). Mimicking Schizophrenia: Reducing P300b by Minimally Fragmenting Healthy Participants' Selves Using Immersive Virtual Reality Embodiment. *Frontiers in Human Neuroscience*, *12*, 504. https://doi.org/10.3389/fnhum.2018.00504

Spanlang, B., Normand, J.-M., Borland, D., Kilteni, K., Giannopoulos, E., Pomés, A., González-Franco, M., Perez-Marcos, D., Arroyo-Palacios, J., Muncunill, X. N., & Slater, M. (2014). How to Build an Embodiment Lab: Achieving Body

Representation Illusions in Virtual Reality. *Frontiers in Robotics and AI*, *1*, 9. https://doi.org/10.3389/frobt.2014.00009

Spyridonis, F., Daylamani-Zad, D., & Nightingale, J. (2024). PublicVR: A virtual reality exposure therapy intervention for adults with speech anxiety. *Virtual Reality*, *28*. https://doi.org/10.1007/s10055-024-00998-x

Stein, J.-P., & Ohler, P. (2017). Venturing into the uncanny valley of mind—The influence of mind attribution on the acceptance of human-like characters in a virtual reality setting. *Cognition*, *160*, 43–50. https://doi.org/10.1016/j.cognition.2016.12.010

Takac, M., Collett, J., Blom, K. J., Conduit, R., Rehm, I., & De Foe, A. (2019). Public speaking anxiety decreases within repeated virtual reality training sessions. *PloS One*, *14*(5), e0216288. https://doi.org/10.1371/journal.pone.0216288

Tarnanas, loannis, Tsoukalas, loannis, & Stogiannidou, A. (2003). Virtual reality as a psychosocial coping environment. *Annual Review of CyberTherapy and Telemedicine*, *1*(Bowers, D., Meichenbaum, D. (1994). The unconscious reconsidered. New York: John Willey.), 77–81.

Thalmann, D. (2000). *Challenges for the Research in Virtual Humans*. 9.

Trivedi, H., & Mousas, C. (2023). Human-virtual crowd interaction: Towards understanding the effects of crowd avoidance proximity in an immersive virtual environment. *Computer Animation and Virtual Worlds*, *34*(3–4), e2169. https://doi.org/10.1002/cav.2169

Ulicny, B., & Thalmann, D. (2001). Crowd simulation for interactive virtual environments and VR training systems. In N. Magnenat-Thalmann & D. Thalmann (Eds.), *Computer Animation and Simulation 2001* (pp. 163–170). Springer Vienna. https://doi.org/10.1007/978-3-7091-6240-8_15

Vahle, N., & Tomasik, M. J. (2022). The Embodiment of an Older Avatar in a Virtual Reality Setting Impacts the Social Motivation of Young Adults. *Experimental Aging Research*, *48*(2), 164–176. https://doi.org/10.1080/0361073X.2021.1943793

Valmaggia, L. R., Day, F., & Rus-Calafell, M. (2016). Using virtual reality to investigate psychological processes and mechanisms associated with the onset and maintenance of psychosis: A systematic review. *Social Psychiatry and Psychiatric Epidemiology*, *51*(7), 921–936. https://doi.org/10.1007/s00127-016-1245-0

Van Gelder, J.-L., Cornet, L. J. M., Zwalua, N. P., Mertens, E. C. A., & van der Schalk, J. (2022). Interaction with the future self in virtual reality reduces self-defeating behavior in a sample of convicted offenders. *Scientific Reports*, *12*(1), Article 1. https://doi.org/10.1038/s41598-022-06305-5

Venning, A., Herd, M. C., Oswald, T. K., Razmi, S., Glover, F., Hawke, T., Quartermain, V., & Redpath, P. (2021). Exploring the acceptability of a digital mental health platform incorporating a virtual coach: The good, the bad, and the opportunities. *Health Informatics Journal*, *27*(1), 1460458221994873. https://doi.org/10.1177/1460458221994873

Vinayagamoorthy, V., Gillies, M., Steed, A., Tanguy, E., Pan, X., Loscos, C., & Slater, M. (2006). *Building Expression into Virtual Characters*. 41.

Wang, Z., Giannopoulos, E., Slater, M., Peer, A., & Buss, M. (2011). Handshake: Realistic Human-Robot Interaction in Haptic Enhanced Virtual Reality. *Presence: Teleoperators and Virtual Environments*, *20*(4), 371–392. https://doi.org/10.1162/PRES_a_00061

Wieser, M., Pauli, P., Grosseibl, M., Molzow, I., & Mühlberger, A. (2010). Virtual Social Interactions in Social Anxiety-The Impact of Sex, Gaze, and Interpersonal Distance. *Cyberpsychology, Behavior and Social Networking*, *13*, 547–554. https://doi.org/10.1089/cyber.2009.0432

Wolf, E., Merdan, N., Dölinger, N., Mal, D., Wienrich, C., Botsch, M., & Latoschik, M. E. (2021). The Embodiment of Photorealistic Avatars Influences Female Body Weight Perception in Virtual Reality. *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, 65–74. https://doi.org/10.1109/VR50410.2021.00027

Yamamoto, K., & Nakao, T. (2022). Top-Down Influence Leads to a Reduced Sense of Body Ownership in Individuals With Depersonalization Tendencies: A Focus on Full Body Illusion. *Frontiers in Psychology*, *13*. https://www.frontiersin.org/articles/10.3389/fpsyg.2022.790960

Zell, E., Zibrek, K., & McDonnell, R. (2019). Perception of virtual characters. *ACM SIGGRAPH 2019 Courses*, 1–17. https://doi.org/10.1145/3305366.3328101

3 A Randomised Controlled Test of Emotional Attributes of A Virtual Coach Within A Virtual Reality Mental Health Treatment

ABSTRACT

Background: Virtual coach plays an important role in some automated mental health therapies in VR, yet the impact of their characteristics and emotional attributes has not been studied in depth.

Aim: This chapter aimed to test whether positive non-verbal behaviours of a virtual coach can enhance people's engagement and confidence in an automated VR therapy.

Method: We developed the initial consultation with a VR coach for an automated VR therapy for fear of heights. 120 individuals scoring highly for fear of heights participated. In a two-by-two factor, between-groups, randomised design, participants met one of the four versions of the virtual coach, with warm facial expression (with/without) and affirmative nods (with/without). Participants rated the therapeutic alliance, treatment credibility, and treatment expectancy after the VR consultation.

Results: Both warm facial expressions (difference = 7.44 [3.25, 11.62], p = 0.001, η_p^2 = 0.10) and affirmative nods (difference = 4.36 [0.21, 8.58], p = 0.040, η_p^2 = 0.04) by the virtual coach independently increased therapeutic alliance. Affirmative nods increased the treatment credibility (difference = 1.76 [0.34, 3.11], p = 0.015, η_p^2 = 0.05) and expectancy (difference = 2.28 [0.45, 4.12], p=0.015, η_p^2 = 0.05) but warm facial expressions did not increase treatment credibility (difference = 0.64 [-0.75, 2.02], p = 0.363, η_p^2 = 0.01) or expectancy (difference = 0.36 [-1.48, 2.20], p = 0.700, η_p^2=0.001). There were no significant interactions between head nods and facial expressions in the occurrence of therapeutic alliance (p = 0.403, η_p^2 = 0.01), credibility (p = 0.072, η_p^2 = 0.03), or expectancy (p = 0.275, η_p^2 = 0.01).

Conclusions: There is likely to be therapeutic value in detailed consideration of the animations of virtual coaches for automated VR therapies.

This chapter is adapted from the paper:

Wei, S., Freeman, D. & Rovira, A. A randomised controlled test of emotional attributes of a virtual coach within a virtual reality (VR) mental health treatment. Sci Rep 13, 11517 (2023). https://doi.org/10.1038/s41598-023-38499-7

3.1 INTRODUCTION

Automated VR therapies are likely to be a valid approach to scale up the delivery of efficacious psychological treatment for mental health difficulties (Freeman et al., 2018, 2022). Without reliance on the relatively scarce resource of trained therapists, but with the opportunity for patients to access help in their own homes via the latest standalone consumer VR headsets, automated VR therapies offer a route to greater mental health treatment provision. Virtual coaches – who provide instruction, education, encouragement, and feedback to patients – an important element in current VR therapies. Our systematic review in Chapter 2 indicated that certain characteristics (e.g. voice, facial emotion) of the VHs influence user perception and behaviour (Giguère et al., 2023; Nijman et al., 2020). However, the virtual coach is a relatively novel element and only few studies have included one, there is limited evidence in the literature regarding how the detailed design of the coach affects the therapeutic experience in VR.

In this chapter, we aim to understand the potential influence of the VR coach's emotional attributes by testing two specific characteristics of the virtual coach's non-verbal behaviour that could enhance the VR treatment experience. If characteristics of the VR therapist do affect the patient experience – including markers of better treatment outcomes – then there could be a programme of work testing a range of potentially important factors in their realisation.

Therapeutic alliance, a positive relationship between patient and therapist, is a reliable predictor of better mental health treatment outcomes (Ardito & Rabellino, 2011; Horvath & Symonds, 1991), and even affects the efficacy of psychological treatments delivered in digital forms (Miloff et al., 2020; Mitruț et al., 2021; Sagui-Henson et al., 2022). Similarly, patient belief in the credibility of a therapy offered, and expectations of successful outcomes, predict better treatment outcomes (Schulte, 2008; Thompson-Hollands et al., 2014). Therefore, creating VR coaches that enhance therapeutic alliance and treatment credibility and expectancy could help maximise outcomes from automated VR therapies. Conducting randomized controlled clinical trials to compare treatment outcomes for every change in the characteristics of a virtual coach is infeasible, since clinical trials are typically labour and resource-intensive studies. Instead, the use of proxy measures for good outcomes, such as therapeutic alliance and treatment credibility and expectancy, provide a practical solution for examining potential treatment effects of variation in a virtual coach.

A growing body of research has focused on the experience of VHs in coaching and therapies within non-immersive modalities. For example, an early test from Bickmore and Picard (2004) compared empathic and neutral versions of a virtual exercise advisor presented on a desktop computer. The empathic advisor displayed caring behaviours, such as direct gaze towards the participant and a concerned facial expression when participants felt unwell. Participants perceived more care from the empathic advisor and were more willing to continue the consultation. Likewise, Lawson and Mayer (2022) found that people reported a favourable social connection with a virtual instructor that had a positive voice and body gestures in video coaching. Furthermore, Ter Stal et al. (2021) tested the effects of positive facial expressions and response texts of an online virtual coach, who provided tips on physical activity and healthy nutrition. Results showed that positive text responses from the coach, programmed as responses with a greater number of positive words and longer word count, significantly increased participants' perceived rapport with the coach. However, positive facial expressions did not have a significant effect.

Other studies have looked at VHs in mental health digital interventions. DeVault and colleagues (2014) created a virtual interview program on a desktop computer, where a virtual interviewer assessed people's distress indicators. They compared two versions of the interviewer – an automated version and a Wizard-of-Oz version in which the human operators triggered the virtual interviewer's spoken and gestural responses. The results showed that people who experienced the Wizard-of-Oz version reported greater rapport, high system usability, and a strong sense that the VH was a good listener. Lisetti et al. (2013) evaluated an intervention for alcohol dependence delivered with an empathic or non-empathic virtual counsellor presented on a computer screen. Adding empathic qualities (e.g. nodding, smiling, head posture mimicry, and eyebrow movement) led to a higher level of trust in the counsellor and a more significant social influence. On the other hand, Ranjbartabar and colleagues (2021) reported in a study of virtual therapists presented on a computer screen that empathic virtual therapists might not necessarily deliver better emotional outcomes than neutral therapists. Overall, reviews of the use of VHs have highlighted the potential benefits of realisation of emotional behaviours in facilitating participant engagement (Elshan et al., 2022; Mitruț et al., 2021).

In studies of VHs in VR, research has suggested that characters' behavioural realism and positive non-verbal communication can enhance their social impact (Aburumman et al., 2022; Volonte et al., 2020; Wu et al., 2021). Wu et al. (2021) reported that people perceived stronger social presence and interpersonal attractions when collaborating with a highly expressive VH, featuring detailed facial movements and body tracking, compared to a low expressive version. More specifically, non-verbal behaviours such as positive facial expressions with smiles (Volonte et al., 2020) and responsive nodding (Aburumman et al., 2022) by characters increased perceived friendliness, trust, and bonding in VR social situations. However, relationships with virtual coaches in automated VR therapies for mental health difficulties have not been experimentally examined. Furthermore, the potential influence of participant factors on the experience of a VR coach is unknown. For instance, individuals who are especially mistrustful in everyday life situations may find it harder to form a therapeutic alliance with a virtual coach (Trotta et al., 2021), but this has not been tested.

The current study tested the impact of a VR coach's positive non-verbal behaviours on therapeutic alliance and treatment credibility and expectancy for an acrophobia treatment. Warm facial expressions and affirmative head nods were selected as key emotional attributes, as previous research suggests these features enhance perceived friendliness and trust in VHs (Volonte et al., 2020; Aburumman et al., 2022). Additionally, we tested whether a participant's level of mistrust may moderate the relationship with a virtual coach. Acrophobia was selected as the therapy focus because it is a prevalent mental difficulty among the non-clinical population, making it easier to recruit participants. This ensured that eligible individuals were genuinely motivated to confront and reduce their fear of heights. Our primary hypotheses were that the addition of warm facial expressions and affirmative nods would independently enhance the therapeutic alliance and treatment credibility and expectancy. Furthermore, we hypothesised that the combined use of warm facial expressions and affirmative nods would have the strongest positive effect (i.e. there would be a significant interaction).

3.2 VR SCENARIO DEVELOPMENT

3.2.1 Apparatus and VR scenario

We used a Windows 10 computer (Intel i7-8700K, Nvidia GeForce GTX 1080Ti, 32 GB RAM) to run the VR scenario and render it on a Meta Quest 2 (Meta, formerly Facebook, 2022) through a wireless connection (Air Link). The VR headset resolution is 1832*1920 pixels per eye and was set up at a 90 Hz refresh rate. Physiological data were recorded with an Empatica E4[7] wristband.

We developed the VR experience in Unity game engine, version 2020.3.22. The experience consisted of an indoor scene where participants met the virtual coach for the first time (Figure 3.1-a) and then they were taken to an outdoor area for a walking

[7] https://www.empatica.com/e4

task (Figure 3.1-b). A video of the VR playthrough experience can be found via the provided link[8].

Indoor scene: Participants met the virtual coach for an introductory consultation, positioned approximately 1.2 meters away to simulate a typical social conversation distance. This distance was the same in all four experimental conditions. The consultation script was adapted from a previously published VR fear of heights clinical trial (Freeman et al., 2018). The virtual coach first introduced herself and explained the cognitive approach to understanding fear of heights (e.g. "The reason we're afraid of heights is because we think something bad is going to happen. And that makes us feel anxious. Then we end up avoiding heights because they feel so scary"). The coach then asked participants questions related to their own fears about heights. Participants answered the questions through a 3D user interface. They went through this interactive conversation at their own pace, which typically took around 4 minutes.

Outdoor scene: Participants were instructed to walk along an elevated walkway. They started in the middle of a virtual terrace to receive instructions from the virtual coach. From there, they had to step out on the walkway (Figure 3.1-b), walk until they reached the circular platform at the other end, and return to the starting point. The scene concluded once they were back to the starting point or at any point when participant decided to end it before completion.

[8] https://youtu.be/j4-2hCrVjo8

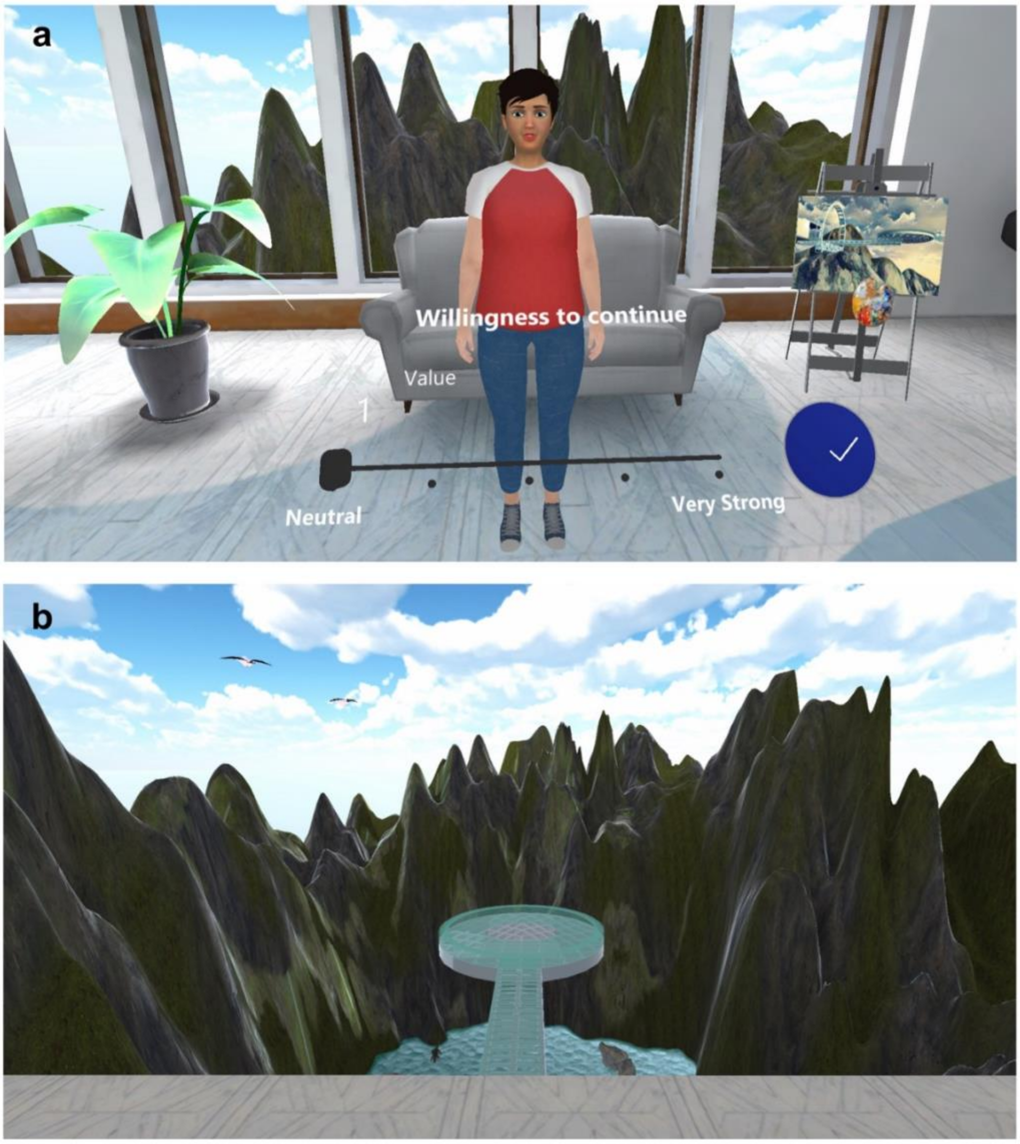


Figure 3.1. Screenshots of the VR experience. (a) Indoor scene: the virtual coach provided an introductory consultation about fear of heights and its treatment. The scene ended with a question about willingness to continue the VR therapy. (b) Outdoor scene: participants were instructed to step out on a glass-floor walkway.

3.2.2 Animation Creation

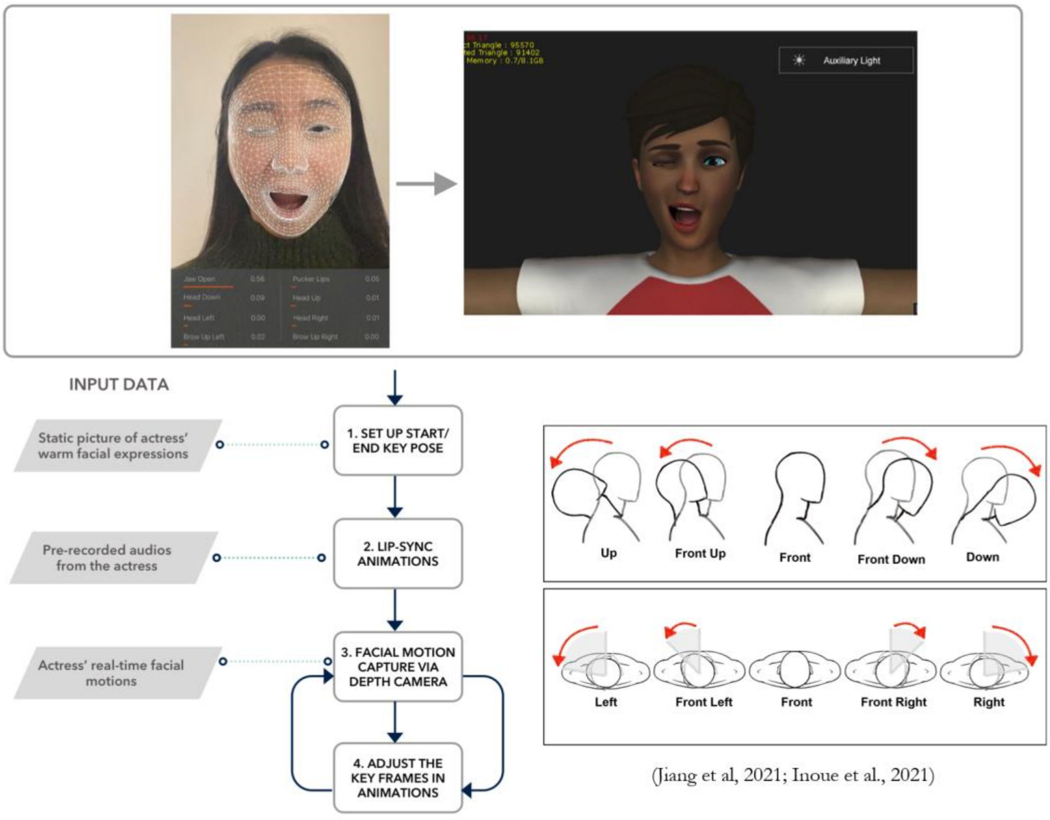


Figure 3.2. Animation creation pipeline. Customised steps to capture and refine facial and head animations for the VR coach.

We combined the use of motion capture, blend-shape and bone animation to create realistic facial expressions and affirmative nods (Maestri, 2006). A female psychologist was invited as the voice and motion capture actress for the virtual coach. Further adjustments for head movements were made according to head movement patterns in real-life psychotherapy communication (Inoue et al., 2021). We created animation in Iclone7[9] with the LiveFace[10] plugin and followed a customised 4-step pipeline (Figure 3.2):

[9] https://www.reallusion.com/iclone/
[10] https://mocap.reallusion.com/iclone-motion-live-mocap/iphone-live-face.html

1) **Creating Key Poses:** We started by creating the initial and final key poses to model the actress's warm facial expressions.

2) **Generating Lip-Sync Animations:** We then generated lip-sync animations from pre-recorded audio using the AccuLIPS system, which generates visemes, aligning lip movements with their corresponding voice sections.

3) **Motion Capture**: We used an iPhone 11 depth camera to capture the actress's facial movements, which were then mapped onto the virtual coach. We applied an adjusted weight profile, fine-tuning the overall motion to approximately 80% and muting the mouth movements for better control.

4) **Refining Animations**: Finally, we reviewed and adjusted the distortion keys to further refine the facial animations. We then added head movements as a separate layer using bone animation, targeting the rigged coach.

The animation pipeline generated the base animation for facial motion and head nods. This process, particularly steps 3 and 4, was typically repeated multiple times until a satisfactory result was achieved. For the virtual coach's body animation, we combined a pre-set standing idle animation from the iClone7 animation library with keyframe bone adjustments. Arm movements were created through keyframe bone adjustments to match the talking script.

We ran a pilot test with 12 individuals to verify our character animations of warm facial expressions and affirmative nods. During the pilot, volunteers were informed that they were shown a virtual coach varying in warm facial expressions (with/without) and affirmative head nods (with/without). They were then asked to identify the exact condition after the trial. All of the pilot volunteers successfully identified the correct animations they observed, confirming the validity of the animations. The result comparison of four types of animation can be found in the playthrough video[11].

3.3 METHODS

3.3.1 Experimental Design

We used a balanced two-by-two factorial between-groups experimental design. The two factors were warm facial expression (with/without, i.e. neutral face) and affirmative head nods (with/without). Participants were randomised and allocated to one of four virtual coach versions: (1) neutral face (2) neutral face and affirmative nods (3) warm facial expressions and (4) warm facial expressions and affirmative nods. In all experimental conditions the virtual coach included basic behaviours such as eye blinking and lip syncing. The study was single-blind. Participants were unaware of the study hypotheses or that they were being randomised to interact with one of the different versions of the virtual coach.

We calculated a target sample size for a between-factors ANOVA using G*power 3.1 (Faul et al., 2009). We specified a medium effect size of partial eta-squared = 0.06 and conventional values of power = 0.80 and $\alpha = 0.05$. A total of 120 participants (30 per condition) would be needed. A randomization list was created using Research Randomizer (Urbaniak & Plous, 2013).

Ethical approval was received from the University of Oxford Medical Sciences Interdivisional Research Ethics Committee (R77367/RE001). The study was performed in accordance with relevant guidelines and regulations and written informed consent was obtained from all participants.

[11] https://youtu.be/j4-2hCrVjo8

3.3.2 Measures

Therapeutic alliance. Alliance with the virtual coach was measured by the Virtual Therapist Alliance Scale (VTAS) (Miloff et al., 2020). It is a 17-item self-report questionnaire describing the perception and relationship with the therapist, such as "The way that the virtual coach communicated was captivating" and "The virtual coach gave me new perspectives on my troubles". All items are scored from 0 (Do not agree at all) to 4 (Agree completely) using the same response format with total scores ranging from 0 to 68. Higher scores reflect a stronger alliance with the virtual coach. The measure had very high internal reliability in this study (Cronbach's $\alpha = 0.94$, $N = 120$).

Treatment credibility/expectancy. Treatment credibility and improvement expectancy of the VR fear of heights treatment was measured by the Credibility/expectancy questionnaire (CEQ) (Devilly & Borkovec, 2000). It is a six-item questionnaire assessing two factors credibility (three items) and expectancy (three items) separately. Each item is rated in a Likert scale and computed to a score from 1 to 9 (responses to the fourth and the sixth item were linear interpolated from 0 to 100% to 1 to 9), with total scores ranging from 3 to 27 for each factor. Both factors had good internal reliability in this study (credibility: Cronbach's $\alpha = 0.81$; expectancy: Cronbach's $\alpha = 0.89$).

Mistrust. Level of mistrust was measured by The Revised Green et al., Paranoid Thoughts Scale (R-GPTS) (Freeman et al., 2021). It is an 18-item scale assessing ideas of persecution, such as "I have been thinking a lot about people avoiding me" and "I was certain people did things in order to annoy me". All items are scored from 0 (Do not agree at all) to 4 (Totally) with total scores ranging from 0 to 72. Higher scores reflect higher levels of mistrust. The measure had very high internal reliability in this study (Cronbach's $\alpha = 0.92$).

Fear of heights. Fear of heights was measured by the Heights Interpretation Questionnaire (HIQ) (Steinman & Teachman, 2011). It is a 16-item self-report questionnaire predicting subjective distress and avoidance of heights. The items assess people's anxious fears such as the fear of falling or getting hurt, when imagining two height situations (i.e. being on a ladder against a two-story house and on the balcony of a 15th-floor building). Each item was rated from 1 to 5 with the total score ranging from 16 to 80. The measure had good internal reliability in this study (Cronbach's α = 0.88).

Presence. We used a single item from the Igroup Presence Questionnaire (Schubert et al., 2001) to measure sense of presence ("In the computer-generated world I had a sense of 'being there'"). This measure was used to verify that both scenarios elicited an acceptable sense of presence. The item was scored on a 5-point Likert scale, from 1 (Not at all) to 5 (Very much).

Warmness of voice. We used a single item to measure the perceived warmness of voice ("The voice of the virtual coach was warm and friendly"). The item was scored from 0 (Do not agree at all) to 4 (Agree completely).

VR behavioural data. We recorded participants' head position and orientation. We extracted these data from the VR headset tracking system, with a sample rate of 120 Hz. In the outdoor scene, we also wrote a log with the timestamp and duration related to the key events - first step on the walkway, reached the circular platform, and the moment they were back to the terrace.

Physiological data. We record participants' cardiac activity and electrodermal activity (EDA). The cardiac data was captured via a photoplethysmography (PPG) sensor, which measured blood volume changes at 64 Hz for heartbeat detection. EDA data was captured in the form of skin conductance at 4 Hz to measure skin autonomic response. Physiological data were collected during the baseline, indoor VR, and outdoor VR sessions. Baseline data were recorded while participants were in a seated resting position, to capture their resting physiological states. We collected the data using the E4 streaming server mode with a script integrated in the VR system in Unity. All signals were logged with a system timestamp in milliseconds synchronised with the VR experience. Details on physiological data processing and modelling are discussed in the next chapter (Chapter 4).

3.3.1 Participants

Participants were primarily recruited via social media advertisements in Oxfordshire. We screened for fear of heights using the *Heights Interpretation Questionnaire (HIQ)* (Steinman & Teachman, 2011) HIQ score > 29, as used in a pervious VR trial for fear of heights (Freeman et al., 2018)) among the general population. Exclusion criteria were individuals who were (a) under 18 years of age, or who reported (b) having photosensitive epilepsy or a significant visual, hearing or mobility impairment that meant that they would not be able to use VR or (c) taking medication which can cause motion sicknesses.

120 participants (female = 66, male = 50, non-binary = 4) with a mean age of 44.4 (SD = 16.4) took part in the VR study. Participants had a mean fear of heights score of 43.8 (SD = 10.8). Table 3.1 presents a summary of participant characteristics.

	Neutral Face (n=30)	**Neutral Face with Nod (n=30)**	**Warm Face (n=30)**	**Warm Face with Nod (n=30)**
Age in years, Mean (SD), range	40.7 (16.6), 18-70 (range)	48.2 (16.4), 18-72 (range)	41.1 (17), 19-77 (range)	47.8 (14.9), 24-74 (range)
Gender female (F), male (M), non-binary (NB)	15 F/13 M/2 NB	15 F/15 M	18 F/12 M	18 F/10 M/2 NB
Fear of Heights Scores (HIQ scores)	43.8 (10.5)	43.7 (10.8)	43.9 (11.2)	43.8 (11.0)

VR Experience	1.97 (0.85)	1.63 (0.72)	1.80 (0.96)	1.80 (0.85)

Table 3.1. Participant characteristics by randomisation group.

3.3.2 Experimental Procedures

Participants were invited for a single session at our VR lab. They were informed that they would try the introductory part of a VR therapy for fear of heights, during which their heart rate and skin conductance data would be recorded. After obtaining written consent to participate in the study, the researcher first helped participants fit the Empatica E4 wristband and recorded five minutes of data as baseline. Following that, they put the VR headset on and they then experienced the indoor scene in the selected parameters according to the experimental group they were allocated in. Once the indoor scene ended, participants took the VR headset off and filled out the measures of therapeutic alliance, warmness of voice, treatment credibility/expectancy, and presence. Once they completed them, they went back into VR to experience the outdoor scene. The researcher ensured that participants knew beforehand they could stop the VR scene at any time without having to give any reason for that. After the outdoor experience, participants took the headset off and filled out the presence and mistrust questionnaires. After all data were collected, they were debriefed about the purpose of the study. The entire session lasted approximately 45 minutes, and participants were reimbursed for their time. Figure 3.3 shows a summary of the study procedure.

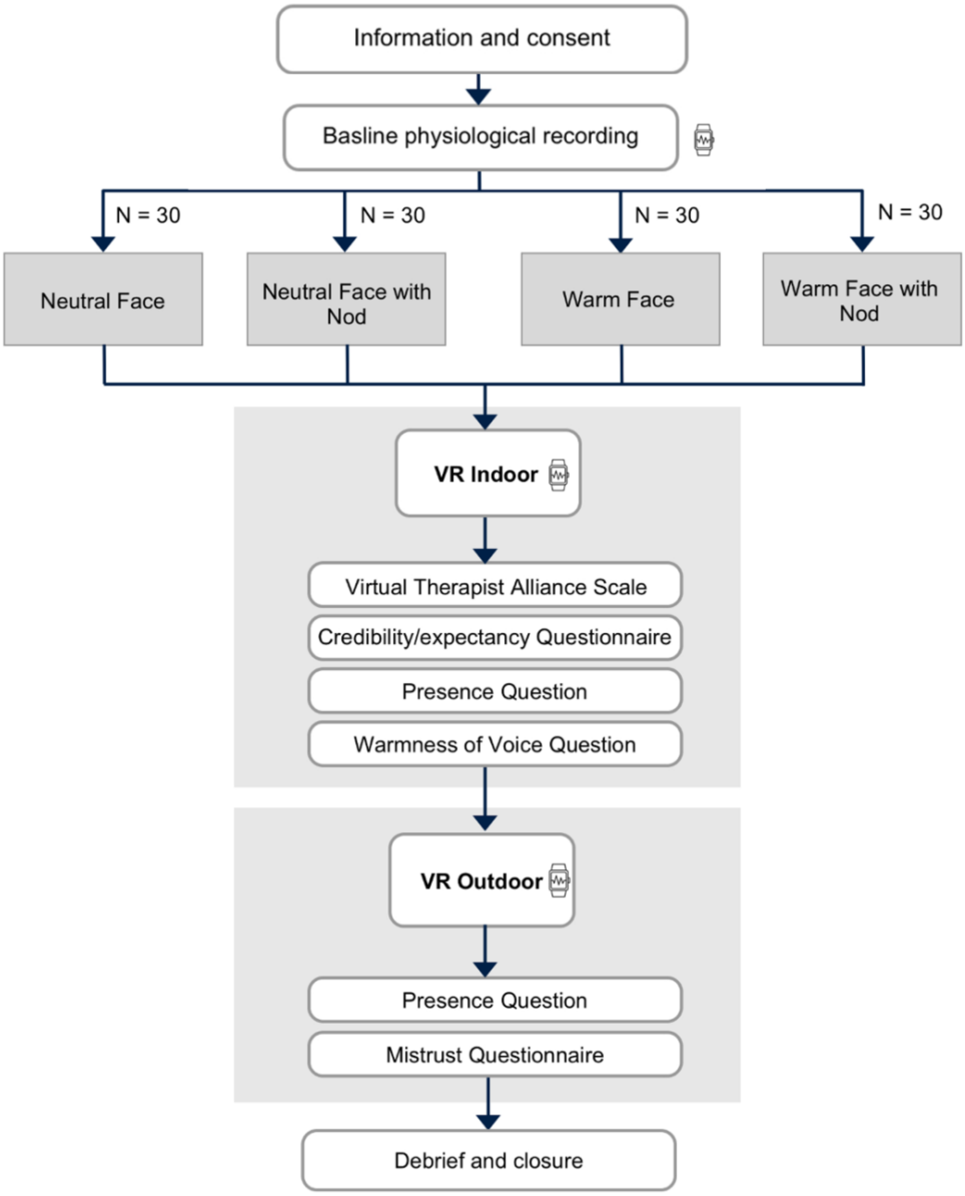


Figure 3.3. Experimental procedures.

3.3.3 Statistical Analysis

We checked that the data were suitable for two-way analysis of variance (ANOVA), using Levene's test for homogeneity of variance and Shapiro-Wilk test of normality (see Appendices). The homogeneity of variance was satisfied for all the variables, while the normality assumption was not met for therapeutic alliance, treatment expectancy, presence and warmness of voice. We maintained the original data without transformations due to the robustness of ANOVA to deviations from normality and the sufficient sample size (Sawyer, 2009).

To assess the effects of warm facial expressions and affirmative head nods on the therapeutic alliance, treatment credibility and expectancy, and other subjective measures we used a two-way ANOVA test with interaction terms. The effect sizes were computed as the partial eta-squared (η_p^2). Tukey's honest significant difference test (Tukey's HSD) was used for multiple pairwise comparisons. All tests for significance were made at the $\alpha = 0.05$ level. We report the results as mean differences and 95% confidence interval (95% CI) of the difference between conditions.

To assess whether mistrust would moderate the effect of warm facial expressions and affirmative head nods on therapeutic alliance, we used a multiple linear regression model with the interaction:

$$VirtualCoachAlliance = WarmFace + AffirmativeNod + AffirmativeNod * Mistrust + Mistrust * WarmFace$$

We evaluated the moderating effect based on the significance of the regression coefficient for the interaction term.

For the behavioural data on participants' performance in the walking task in outdoor VR, we calculated a standardised walking distance along the single direction of the virtual walkway. The standardised distance was calculated from a uniform starting point on the edge of the balcony to the furthest point each participant reached along the walkway. The distance ranged from 0 to 6.4 m, with a minimum value of 0 for those who did not step out onto the walkway, and a maximum value of twice the walkway length (6.4 m), for those who completed the task that reaching to the end of the plank and then walk back. This measure provides a precise assessment of participants' progress in the task. We then applied a two-way ANOVA with interaction terms to evaluate the effects of virtual human (VH) design on task performance.

Data cleaning and processing (i.e. removing data with empty entries, grouping by the virtual coach conditions, and aggregating to calculate scores for each measure) was performed using Python's Pandas and NumPy libraries (Harris et al., 2020; McKinney, 2010). Analyses were conducted using R with R Studio 1.4 (RStudio Team, 2021).

3.4 RESULTS

Figure 3.4 shows the raw data box plots for the primary measures of therapeutic alliance, treatment credibility, and expectancy. Descriptive statistics for the measures are shown in Table 3.2.

There were two sets of incomplete responses for treatment expectancy items as missing data. The full details of the statistical results can be found in Appendices-Chapter3.

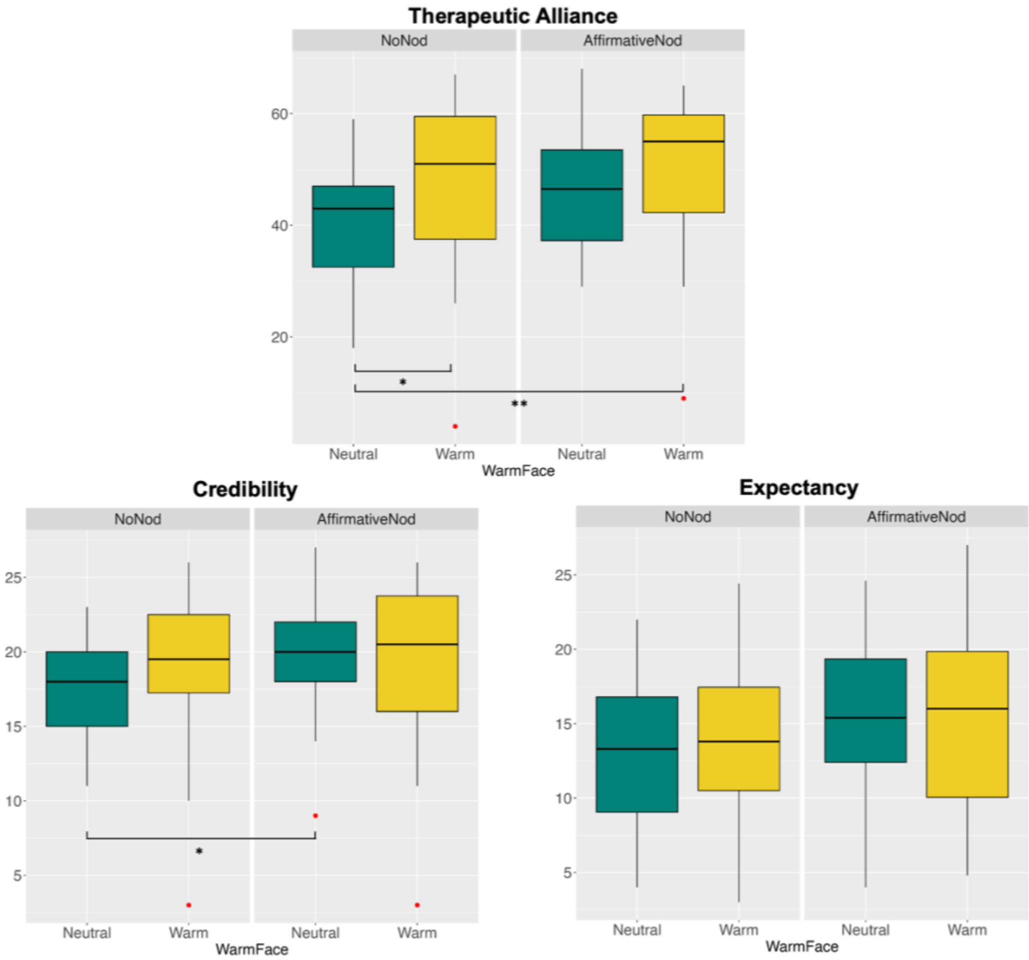


Figure 3.4. Boxplots of the scores of therapeutic alliance, credibility and expectancy. Red dots represent outliers, detected at 1.5 times the interquartile range, star indicates significant difference between groups.

Measures	**Neutral Face**	**Neutral Face with Nod**	**Warm Face**	**Warm Face with Nod**
	Mean (SD)	*Mean (SD)*	*Mean (SD)*	*Mean (SD)*
Therapeutic alliance	40.1 (11.5)	46.2 (10.5)	47.8 (14.9)	50.5 (13.5)
Credibility	17.5 (3.4)	20.1 (4.0)	18.8 (4.9)	19.3 (5.1)
Expectancy	12.8 (5.0)	15.7 (4.9)	14.1 (4.8)	15.1 (6.0)
Presence	7.7 (1.7)	8.2 (1.4)	8.3 (1.2)	8.6 (1.4)
Warmness of the voice	2.8 (1.1)	3.1 (1.1)	3.2 (1.0)	3.5 (0.8)
Mistrust	10.5 (10.3)	9.4 (9.2)	13.8 (11.9)	8.5 (11.5)

Table 3.2. Descriptive data of measures by experimental group.

3.4.1 Therapeutic Alliance

We removed two extreme outliers (< Q1 - 3 * IQR) before the two-way ANOVA statistical test. Simple main effects analysis showed that warm facial expressions (group difference = 7.44, 95% CI = [3.25, 11.62], $F(1, 114) = 12.389$, $p < 0.001$, $\eta_p^2 = 0.10$) and affirmative nods (group difference = 4.36, 95% CI = [0.21, 8.58], $F(1, 114) = 4.318$, $p = 0.040$, $\eta_p^2 = 0.04$) led to a significant increase in therapeutic alliance. The interaction term between warm facial expressions and affirmative nods was not statistically different ($F(1, 114) = 0.705$, $p = 0.403$, $\eta_p^2 = 0.01$). Tukey's HSD Test for multiple comparisons found that therapeutic alliance was significantly greater in the warm face compared to the neutral face condition (p-adj = 0.014) and in the warm face with nod compared to the neutral face condition (p-adj < 0.001).

3.4.2 Treatment Credibility and Expectancy

Simple main effects analysis showed that affirmative nods (group difference = 1.76, 95% CI = [0.34, 3.11], $F(1, 113) = 6.11$, $p = 0.015$, $\eta_p^2 = 0.05$) led to a significant increase in treatment credibility but that warm facial expressions did not (group difference = 0.64, 95% CI = [-0.75, 2.02], $F(1, 113) = 0.833$, $p = 0.363$, $\eta_p^2 = 0.01$). There was no statistically significant interaction between warm facial expressions and affirmative nods ($F(1, 113) = 3.293$, $p = 0.072$, $\eta_p^2 = 0.03$), although there was a trend that the interaction factor leading to greater credibility ratings. Tukey's HSD Test for multiple comparisons found that credibility was significantly greater in the neutral face with head nod compared to the neutral face condition without the head nod (p-adj = 0.016).

Two participants had incomplete data completion for the expectancy items and were removed from the statistical analysis. Simple main effects analysis showed that affirmative nods (group difference = 2.28, 95% CI = [0.45, 4.12], $F(1, 114) = 6.055$, p

$= 0.015, \eta_p^2 = 0.05$) led to a significant increase in expectancy but that warm facial expressions did not (group difference = 0.36, 95% CI = [-1.48, 2.20], $F(1, 114) = 0.833$, $p = 0.700, \eta_p^2 = 0.001$). The interaction term between warm facial expressions and affirmative nods was not statistically significant ($F(1, 114) = 1.202$, $p = 0.275, \eta_p^2 = 0.01$). Tukey's HSD Test for multiple comparisons found that mean expectancy was not significantly different between the experimental groups.

3.4.3 Moderator Effect of Mistrust

A multiple regression was used to predict therapeutic alliance by the variables of warm facial expression, affirmative nods, and their interaction with mistrust ($F(5, 112) = 4.21$, $p = 0.002$, $R^2 = 15.83$). Both interaction terms WarmFace*Mistrust ($p = 0.961$) and AffirmativeNods*Mistrust ($p = 0.971$) were not statistically significant, suggesting mistrust did not moderate the effects.

3.4.4 Presence

A two-way ANOVA showed that warm facial expressions (group difference = 0.70, 95% CI = [0.21, 1.19], $F(1, 113) = 8.119$, $p = 0.005, \eta_p^2 = 0.07$) led to significantly higher levels of presence but that affirmative nods did not (group difference = 0.40, 95% CI = [-0.09, 0.89], $F(1, 113) = 2.649$, $p = 0.106, \eta_p^2 = 0.02$). The interaction between warm facial expressions and affirmative nods was not statistically significant ($F(1, 113) = 0.178$, $p = 0.674$, $\eta_p^2 = 0.001$). Tukey's HSD Test for multiple comparisons found that presence was significantly higher in the warm face with nod compared to the neutral face condition (p-adj = 0.011).

3.4.5 Warmness of Voice

A two-way ANOVA showed that warm facial expressions (group difference = 0.45, 95% CI = [0.16, 0.75], F(1, 110) = 9.44, p = 0.003, η_p^2 = 0.08) and affirmative nods (group difference = 0.39, 95% CI = [0.09, 0.67], F(1, 110) = 6.54, p = 0.01, η_p^2 = 0.06) led to significantly higher ratings of voice warmness. The interaction between warm facial expressions and affirmative nods was not statistically significant (F(1, 110) = 1.579, p = 0.212, η_p^2 = 0.01). Tukey's HSD Test for multiple comparisons found that the warmness of the voice was significantly greater in the warm face with nod compared to the neutral face condition (p-adj < 0.001), the warm face condition compared to the neutral face condition (p-adj = 0.017) and the neutral face with nod condition compared to the neutral face condition (p-adj = 0.040).

3.4.6 Behavioural Data

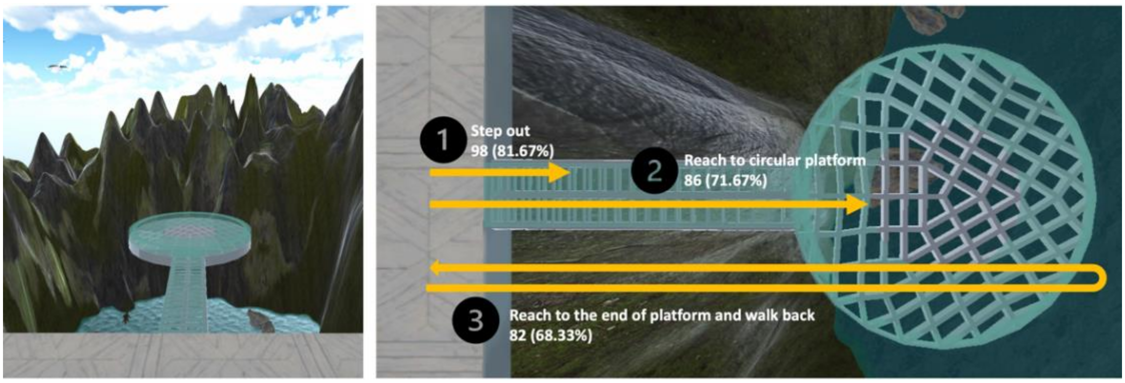


Figure 3.5. Three stages of virtual heights walking task.

We conducted an exploratory analysis of participants' walking task performance across the virtual height based on three stages of the task (see Figure 3.5). Table 3.3 shows the summary statistics. 82 out of 120 participants (68.33%) completed the task. The average time to move forward and step on to the walkway was 38.0 seconds (SD = 47.1), and the average duration spent in outdoor VR after the task brief was 113.1 seconds (SD = 82.0).

We also calculated the normalised walking distance based on the horizontal distance of the virtual walkway. Two sets of data were excluded; one participant experienced a VR connection loss and another opted out of the walking task in the outdoor scene. A two-way ANOVA suggested that warm facial expressions ($p = 0.187$) and affirmative nods ($p = 0.374$) did not have statistically significant effects on walking distance. Similarly, warm facial expressions and affirmative nods did not have statistically significant effects on the time to step on to the virtual walkway (warm facial expressions: $p = 0.356$, affirmative nods: $p = 0.978$) and the time spent in the outdoor scene (warm facial expressions: $p = 0.732$, affirmative nods: $p = 0.511$).

	Task Completion	**Normalised Distance (m)**	**Duration - StepOnWalkway (s)**	**Duration - Outdoor (s)**
	Number (%)	Mean (SD)	Mean (SD)	Mean (SD)
All groups	82 (68.33%)	4.65 (2.73)	38.0 (47.1)	113.1 (82.0)
Neutral Face	21 (70.00%)	4.78 (2.68)	36.1 (60.9)	111.5 (93.8)
Neutral Face with Nod	17 (56.67%)	3.86 (3.00)	50.1 (53.1)	109.5 (71.1)
Warm Face	22 (73.33%)	4.96 (2.62)	40.4 (40.1)	124.6 (87.0)
Warm Face with Nod	22 (73.33%)	4.99 (2.55)	27.5 (28.6)	106.7 (77.2)

Table 3.3. Summary statistics of the VR walking task.

3.5 DISCUSSION

The virtual coach is an important element for automating therapies for mental health in VR. This chapter provides evidence towards the potential therapeutic value of two non-verbal behaviours related to the virtual coach's animation. The evidence was gathered through the analysis of scores for VHs in therapeutic alliance, treatment credibility, and expectancy across different experimental groups.

The results support our initial hypotheses partially. We hypothesised that warm facial expressions and affirmative head nods would enhance the therapeutic alliance, treatment credibility, and expectancy, and their combination would have the strongest effect. The results showed that warm facial expressions and affirmative head nods individually affected therapeutic alliance, and the effect of warm facial expressions was greater than the head nods. Additionally, affirmative head nods increased people's beliefs in both the credibility of the treatment and the expectancy of good outcomes. The interaction between warm facial expressions and affirmative head nods was not statistically significant but there was a trend in the direction that the combination led to greater treatment credibility. In summary, how a virtual coach is designed affects the treatment experience and potentially the therapeutic outcomes. In this study we showed that there is likely to be value in implementing facial expressions and positive non-verbal behaviours for the virtual coach.

The primary finding that warm facial expressions and affirmative head nods increase alliance provides further evidence from previous studies of VHs outside of the context of VR mental health therapies (Aburumman et al., 2022; Lawson & Mayer, 2022; Oh et al., 2016; Osugi & Kawahara, 2018; Volonte et al., 2020). Similar to the conclusion by Oh et al. (2020) that virtual agents' facial expressions contribute more than body movements (such as raising of hands and head tilts), the effect size of warm facial expressions of the virtual coach in the current study on the therapeutic alliance was larger than affirmative head nods. Unexpectedly, we did not detect a main effect of warm facial expressions on treatment credibility or expectancy. However, when warm facial expressions were combined with affirmative head nods, there was a trend towards higher credibility ratings. This result might be due to the head nods giving the impression that the therapist was attentively listening and acknowledging participant responses (Inoue et al., 2021). Such an impression could have then enhanced the potential positive effects of warm facial expressions on treatment credibility when they were displayed simultaneously. Interestingly, positive non-verbal behaviours also led to positive voice perception, which highlights an interplay between perceptions of different sensory traits of VHs. Future research could examine other attributes (e.g. visual, auditory, and other non-verbal behaviours such as eye gaze and hand gestures) and their interactive effects.

Our main focus was the effect of characteristics of a coach on established proxies for good therapeutic outcomes. But we also took an exploratory look at potential effects on participants' behaviours in relation to virtual heights. Approximately one-third of participants did not complete the task in the outdoor scenario. There was no significant difference in task completion rate or distance covered between experimental groups. Since the study exposed participants to the feared situation only once, as opposed to multiple therapy sessions, we did not make any specific predictions. It is plausible that the relationship with the virtual coach would have little to no noticeable effect during participants' first exposure to the therapy. Indeed, no group differences were observed in whether participants stepped onto the platform or in the distance they covered.

The study has several limitations. First, we do not know whether the effects of the non-verbal behaviours do translate to better outcomes in a mental health therapy. This would require a clinical trial to provide evidence. Our view is that using proxies of good outcomes such as therapeutic alliance and treatment credibility is a more sensible testing strategy than conducting multiple clinical trials on small changes to a programme. When such treatments get used at scale then it may be possible to look at outcome effects by programming modifications. Second, we only focused on the virtual coach's facial expressions and head nods and did not account for factors such as gender, ethnicity, and age of the participants. Previous research indicates that people tend to have stronger bonds with VHs with similar characteristics as the person (Gamberini et al., 2015). In the future, it is likely that people will be able to customise the coach's appearance, style, and even their animations, which could be tested with the measurement of therapeutic alliance. Third, we used single-blind testing, with the experimenter being aware of a participant's allocated condition since there was only one experimenter running the study. This design choice may have introduced potential bias during the conduct of the experiment, including the experimenter's greeting style, which could have subsequently influenced participants' subjective ratings. Fourth, mistrust was measured at the end of testing, and this may have affected ratings, and therefore was not actually a true moderator variable. However, there was no clear evidence that mistrust was linked to perceptions of the therapeutic alliance or treatment credibility or expectancy. Finally, the violation of normality in the two-way ANOVA can result in overestimating test significance and increase the chances of Type I error. For example, the p-value of 0.04 for the relationship between nodding and alliance is close to the significance threshold, indicating that a larger sample size will be needed for more robust conclusions.

In this chapter, we investigated the positive non-verbal behaviours of a virtual coach, a special example of an active interaction partner, during an automated VR consultation for the treatment of fear of heights. The inclusion of warm facial expressions and affirmative head nods independently increased therapeutic alliance. Furthermore, affirmative head nods by the virtual coach improved perceptions of treatment credibility and positive outcome expectancy. The findings highlight the potential to enhance the experience and effectiveness of VR therapies through tailored VR character design. In the next chapter, we will examine the impacts of virtual coach design and also reflect on participants' physiological responses. While our study focused on the cognitive treatment of fear of heights, further study is needed to examine the degree to which there is generalization to other mental health difficulties and different treatment techniques. The development of VR therapies would benefit from a systematic programme of research of the best attributes of virtual coaches, which may vary depending on the conditions and treatment techniques, and require strong collaborations between clinical staff, people with lived experiences, and software developer.

REFERENCES

Aburumman, N., Gillies, M., Ward, J. A., & Hamilton, A. F. de C. (2022). Nonverbal communication in virtual reality: Nodding as a social signal in virtual interactions. *International Journal of Human-Computer Studies*, *164*, 102819. https://doi.org/10.1016/j.ijhcs.2022.102819

Ardito, R. B., & Rabellino, D. (2011). Therapeutic Alliance and Outcome of Psychotherapy: Historical Excursus, Measurements, and Prospects for Research. *Frontiers in Psychology*, *2*, 270. https://doi.org/10.3389/fpsyg.2011.00270

Bickmore, T. W., & Picard, R. W. (2004). Towards caring machines. *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, 1489–1492. https://doi.org/10.1145/985921.986097

DeVault, D., Artstein, R., Benn, G., Dey, T., Fast, E., Gainer, A., Georgila, K., Gratch, J., Hartholt, A., Lhommet, M., Lucas, G., Marsella, S., Morbini, F., Nazarian, A., Scherer, S., Stratou, G., Suri, A., Traum, D., Wood, R., … Morency, L.-P. (2014). SimSensei kiosk: A virtual human interviewer for healthcare decision support. *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-Agent Systems*, 1061–1068.

Devilly, G. J., & Borkovec, T. D. (2000). Psychometric properties of the credibility/expectancy questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*, *31*(2), 73–86. https://doi.org/10.1016/S0005-7916(00)00012-4

Elshan, E., Zierau, N., Engel, C., Janson, A., & Leimeister, J. M. (2022). Understanding the Design Elements Affecting User Acceptance of Intelligent Agents: Past, Present and Future. *Information Systems Frontiers*, *24*(3), 699–730. https://doi.org/10.1007/s10796-021-10230-9

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Freeman, D., Haselton, P., Freeman, J., Spanlang, B., Kishore, S., Albery, E., Denne, M., Brown, P., Slater, M., & Nickless, A. (2018). Automated psychological therapy using immersive virtual reality for treatment of fear of heights: A single-blind, parallel-group, randomised controlled trial. *The Lancet Psychiatry*, *5*(8), 625–632. https://doi.org/10.1016/S2215-0366(18)30226-8

Freeman, D., Lambe, S., Kabir, T., Petit, A., Rosebrock, L., Yu, L.-M., Dudley, R., Chapman, K., Morrison, A., O'Regan, E., Aynsworth, C., Jones, J., Murphy, E., Powling, R., Galal, U., Grabey, J., Rovira, A., Martin, J., Hollis, C., … West, J. (2022). Automated virtual reality therapy to treat agoraphobic avoidance and distress in patients with psychosis (gameChange): A multicentre, parallel-group, single-blind, randomised, controlled trial in England with mediation and moderation analyses. *The Lancet Psychiatry*, *9*(5), 375–388. https://doi.org/10.1016/S2215-0366(22)00060-8

Freeman, D., Loe, B. S., Kingdon, D., Startup, H., Molodynski, A., Rosebrock, L., Brown, P., Sheaves, B., Waite, F., & Bird, J. C. (2021). The revised Green *et al.* , Paranoid Thoughts Scale (R-GPTS): Psychometric properties, severity ranges, and clinical cut-offs. *Psychological Medicine*, *51*(2), 244–253. https://doi.org/10.1017/S0033291719003155

Gamberini, L., Chittaro, L., Spagnolli, A., & Carlesso, C. (2015). Psychological response to an emergency in virtual reality: Effects of victim ethnicity and emergency type on helping behavior and navigation. *Computers in Human Behavior*, *48*, 104–113. https://doi.org/10.1016/j.chb.2015.01.040

Giguère, S., Potvin, S., Beaudoin, M., Dellazizzo, L., Giguère, C.-É., Furtos, A., Gilbert, K., Phraxayavong, K., & Dumais, A. (2023). Avatar Intervention for Cannabis Use Disorder in Individuals with Severe Mental Disorders: A Pilot Study. *Journal of Personalized Medicine*, *13*(5), Article 5. https://doi.org/10.3390/jpm13050766

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), Article 7825. https://doi.org/10.1038/s41586-020-2649-2

Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology*, *38*(2), 139–149. https://doi.org/10.1037/0022-0167.38.2.139

Inoue, M., Irino, T., Furuyama, N., & Hanada, R. (2021). Observational and Accelerometer Analysis of Head Movement Patterns in Psychotherapeutic Dialogue. *Sensors*, *21*, 3162. https://doi.org/10.3390/s21093162

Lawson, A. P., & Mayer, R. E. (2022). Does the emotional stance of human and virtual instructors in instructional videos affect learning processes and outcomes? *Contemporary Educational Psychology*, *70*, 102080. https://doi.org/10.1016/j.cedpsych.2022.102080

Lisetti, C., Amini, R., Yasavur, U., & Rishe, N. (2013). I Can Help You Change! An Empathic Virtual Agent Delivers Behavior Change Health Interventions. *ACM Transactions on Management Information Systems*, *4*(4), 1–28. https://doi.org/10.1145/2544103

Maestri, G. (2006). Facial and Dialogue Animation. In *Digital Character Animation 3* (pp. 196–219). New Riders.

McKinney, W. (2010). *Data Structures for Statistical Computing in Python*. 56–61. https://doi.org/10.25080/Majora-92bf1922-00a

Miloff, A., Carlbring, P., Hamilton, W., Andersson, G., Reuterskiöld, L., & Lindner, P. (2020). Measuring Alliance Toward Embodied Virtual Therapists in the Era of Automated Treatments With the Virtual Therapist Alliance Scale (VTAS):

Development and Psychometric Evaluation. *Journal of Medical Internet Research*, *22*(3), e16660. https://doi.org/10.2196/16660

Mitruţ, O., Moldoveanu, A., Petrescu, L., Petrescu, C., & Moldoveanu, F. (2021). A Review of Virtual Therapists in Anxiety and Phobias Alleviating Applications. In J. Y. C. Chen & G. Fragomeni (Eds.), *Virtual, Augmented and Mixed Reality* (pp. 71–79). Springer International Publishing. https://doi.org/10.1007/978-3-030-77599-5_6

Nijman, S. A., Veling, W., Greaves-Lord, K., Vos, M., Zandee, C. E. R., Aan het Rot, M., Geraets, C. N. W., & Pijnenborg, G. H. M. (2020). Dynamic Interactive Social Cognition Training in Virtual Reality (DiSCoVR) for People With a Psychotic Disorder: Single-Group Feasibility and Acceptability Study. *JMIR Mental Health*, *7*(8). https://doi.org/10.2196/17808

Oh Kruzic, C., Kruzic, D., Herrera, F., & Bailenson, J. (2020). Facial expressions contribute more than body movements to conversational outcomes in avatar-mediated virtual environments. *Scientific Reports*, *10*(1), 20626. https://doi.org/10.1038/s41598-020-76672-4

Oh, S. Y., Bailenson, J., Krämer, N., & Li, B. (2016). Let the Avatar Brighten Your Smile: Effects of Enhancing Facial Expressions in Virtual Environments. *PLOS ONE*, *11*(9), e0161794. https://doi.org/10.1371/journal.pone.0161794

Osugi, T., & Kawahara, J. I. (2018). Effects of Head Nodding and Shaking Motions on Perceptions of Likeability and Approachability. *Perception*, *47*(1), 16–29. https://doi.org/10.1177/0301006617733209

Ranjbartabar, H., Richards, D., Bilgin, A. A., & Kutay, C. (2021). First Impressions Count! The Role of the Human's Emotional State on Rapport Established with an Empathic versus Neutral Virtual Therapist. *IEEE Transactions on Affective Computing*, *12*(3), 788–800. IEEE Transactions on Affective Computing. https://doi.org/10.1109/TAFFC.2019.2899305

RStudio Team. (2021). *RStudio: Integrated Development Environment for R* [Computer software]. RStudio, PBC. http://www.rstudio.com/

Sagui-Henson, S. J., Welcome Chamberlain, C. E., Smith, B. J., Li, E. J., Castro Sweet, C., & Altman, M. (2022). Understanding Components of Therapeutic Alliance and Well-Being from Use of a Global Digital Mental Health Benefit During the COVID-19 Pandemic: Longitudinal Observational Study. *Journal of Technology in Behavioral Science*, 1–12. https://doi.org/10.1007/s41347-022-00263-5

Sawyer, S. F. (2009). Analysis of Variance: The Fundamental Concepts. *Journal of Manual & Manipulative Therapy*, *17*(2), 27E-38E. https://doi.org/10.1179/jmt.2009.17.2.27E

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments*, *10*(3), 266–281.

Schulte, D. (2008). Patients' outcome expectancies and their impression of suitability as predictors of treatment outcome. *Psychotherapy Research*, *18*(4), 481–494. https://doi.org/10.1080/10503300801932505

Steinman, S. A., & Teachman, B. A. (2011). Cognitive Processing and Acrophobia: Validating the Heights Interpretation Questionnaire. *Journal of Anxiety Disorders*, *25*(7), 896–902. https://doi.org/10.1016/j.janxdis.2011.05.001

ter Stal, S., Jongbloed, G., & Tabak, M. (2021). Embodied Conversational Agents in eHealth: How Facial and Textual Expressions of Positive and Neutral Emotions Influence Perceptions of Mutual Understanding. *Interacting with Computers*, *33*(2), 167–176. https://doi.org/10.1093/iwc/iwab019

Thompson-Hollands, J., Bentley, K. H., Gallagher, M. W., Boswell, J. F., & Barlow, D. H. (2014). Credibility and Outcome Expectancy in the Unified Protocol: Relationship to Outcomes. *Journal of Experimental Psychopathology*, *5*(1), 72–82. https://doi.org/10.5127/jep.033712

Trotta, A., Kang, J., Stahl, D., & Yiend, J. (2021). Interpretation Bias in Paranoia: A Systematic Review and Meta-Analysis. *Clinical Psychological Science*, *9*(1), 3–23. https://doi.org/10.1177/2167702620951552

Urbaniak, G., & Plous, S. (2013). *Research randomizer (version 4.0)[computer software]. 2013* [Computer software].

Volonte, M., Hsu, Y.-C., Liu, K.-Y., Mazer, J. P., Wong, S.-K., & Babu, S. V. (2020). Effects of Interacting with a Crowd of Emotional Virtual Humans on Users' Affective and Non-Verbal Behaviors. *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 293–302. https://doi.org/10.1109/VR46266.2020.00049

Wu, Y., Wang, Y., Jung, S., Hoermann, S., & Lindeman, R. W. (2021). Using a Fully Expressive Avatar to Collaborate in Virtual Reality: Evaluation of Task Performance, Presence, and Attraction. *Frontiers in Virtual Reality*, *2*. https://www.frontiersin.org/articles/10.3389/frvir.2021.641296

4 Physiological Responses to A Virtual Coach in A Virtual Reality Mental Health Application

ABSTRACT

Background: The design and implementation of virtual coaches in automated VR therapy can influence individuals' subjective perceptions of the coach and the treatment. It is unclear whether the virtual human design also affect people's physiological responses.

Aim: This chapter investigates whether VR and the detailed design of a virtual coach would influence participants' physiological responses.

Method: We analysed the physiological data from the experimental study described in Chapter 3. A total of 120 individuals with fear of heights interacted with a virtual coach that varied in warmth of facial expression (with/without) and affirmative nods (with/without). Physiological responses, including cardiac measures (blood volume pulse) and electrodermal activity (EDA), were recorded at baseline, during the indoor VR consultation and the outdoor VR heights exposure.

Results: Exposure to virtual heights elicited stronger physiological responses compared to the indoor consultation, evident in both cardiac responses and EDA. The VR coach's warm facial expressions and affirmative head nods independently increased EDA during the indoor consultation. Affirmative nods showed a trend to increase tonic skin conductance levels ($p = 0.059, \eta_p^2 = 0.610$), while warm facial expressions led larger skin conductance response peak amplitudes ($p = 0.043, \eta_p^2 = 0.574$). No significant effects from warm facial expressions or affirmative head nods were observed on any cardiac measures, including heart rate and heart rate variability.

Conclusions: The emotional attributes of a VR coach modulated physiological responses. Careful design and implementation of the VR coach should be considered to potentially enhance the therapeutic outcomes of automated VR mental health therapy.

4.1 INTRODUCTION

Self-reported subjective measures are the most common way of collecting data in mental health research to capture individuals' thoughts and feelings towards others, facilitating the understanding of social interactions, for example to assess trust towards an interaction partner (Fett et al., 2012; Rotenberg et al., 2023) and evaluating therapeutic relationships (Ardito & Rabellino, 2011; Jerome et al., 2020; Låver et al., 2024). These measures typically take the form of self-rated questionnaires and offer several advantages, including the direct capture of thoughts through explicit questions, ease of administration, and cost-effectiveness (Ciuk et al., 2015; Halbig & Latoschik, 2021). However, self-reports also face potential biases. Social desirability bias can lead individuals to respond in ways they believe will be viewed favourably by others (Braun et al., 2001). Additionally, recall bias may result in inaccurate memories of past experiences, and misinformation can arise from respondents' limited introspective ability to remember their feelings (Brenner & DeLamater, 2016).

Physiological signals, indicative of autonomic nervous system (ANS) activity, reflect the interplay between the sympathetic nervous system (SNS), which triggers the 'fight or flight' response, and the parasympathetic nervous system (PNS), which supports "rest and digest" functions (McCorry, 2007). Physiological monitoring provides objective data. It can detect subtle changes in affect not evident through self-reporting, and reveal both positive and negative emotions that people may not consciously recognise or disclose (Del Piccolo & Finset, 2018; C. Marci & Riess, 2005). In an early study, Di Mascio et al. (1957) explored the relationship between tension and physiological activity during a psychotherapy session. They found that higher tension was associated with increased heart rate (HR), while a decrease in HR corresponded to a sensation of tension release. There was also a tendency for the therapist's HR and lability to mirror those of the patient. More recently, Arza et al. (2019) presented 40 healthy volunteers with various stressor stimuli to monitor their acute stress levels using both subjective questionnaire and physiological signals (e.g. HR, skin temperature, and pulse). Their results showed that physiological biomarkers could effectively distinguish between relaxation and stress states and continuously estimate stress levels that were not apparent through self-reports. Although collecting physiological data can be more costly and the signals often contain noise and artefacts (Can et al., 2023), these measures supplement subjective self-reports, offering better understanding into real-time mental health states. Therefore, this chapter investigates whether physiological signals can reflect participants' perception of VR and emotional engagement with a virtual coach, providing insights into the psychophysiological impact of non-verbal cues during a consultation.

Numerous experimental studies in VR have incorporated physiological measures to assess affective states in mental health research (Halbig & Latoschik, 2021), particularly in phobia-related scenarios without social interaction. Mühlberger and colleagues (2001) examined the effects of repeated VR flight exposure on individuals with a fear of flying, showing that such exposures elicited subjective and physiological fear responses characterized by elevated HR and skin conductance levels (SCL). Each subsequent exposure led to decreases in both HR and SCL compared to the previous session. Martens et al. (2019) assessed participants' stress responses during an open-platform lift ride up the outside of a tall building, offering views of the entire city, compared to a controlled indoor lift condition. In the stressful scenario, participants were asked to step off the platform as if falling, which significantly increased SCL and pulse. Similarly, study about fear of spiders (Mertens et al., 2019) demonstrated that VR can evoke distinct fear responses that greater skin conductance responses (SCR) were associated with higher subjective ratings of fear. Other studies have also recreated traumatic scenes (e.g., war zones) (Maples-Keller et al., 2019; Ridout et al., 2017), addiction stimuli such as nicotine craving (Thompson-Lake et al., 2015), and gambling addiction (Bruder et al., 2020) to monitor people's HR and SCL responses to the virtual cues. These studies provide evidence that VR can trigger physiological responses related to fear and craving comparable to those in real life, with potential for assessment and treatment through virtual exposure.

Other VR mental health studies have focused on the physiological responses to VHs in social interactions, offering a relevant basis for investigating the emotional cues of a virtual coach. These studies primarily use VHs as stressor cues to study the changes in their physiological data (Brinkman et al., 2011; Cornwell et al., 2006; Reichenberger et al., 2022). Brinkman et al. (2011) exposed participants from the general population to varying crowd densities and different ethnic appearances of VHs. They found that exposure to dense crowds or characters of different ethnicities led to stronger fluctuations in HR and SCR. Also, Reichenberger et al. (2022) compared responses to VR human exposure between patients with social anxiety disorder (SAD) and healthy controls. While SAD group showed higher subjective fear, there were no significant group differences in SCL and HR. Furthermore, researchers have begun investigating the potential benefits of supportive VHs in stressful scenarios, with mixed results. Kothgassner et al. (2019) found that receiving social support from a VH in VR prior to a stressor could reduce stress indicators, such as lower HR. However, Stallmann et al. (2023) reported that while support from a virtual agent improved self-reported emotional states, it did not alter stress responses, as indicated by unchanged heart rate variability (HRV) and SCL. Although research has examined VHs as stressors, the effect of detailed characteristics in supportive VHs, such as a virtual coach, on physiological responses remains unexplored.

In real-life healthcare consultations, researchers have examined the impact of therapist communication styles on physiological responses, with mixed outcomes depending on context. Interacting with a supportive therapist could reduce arousal (lower HR and SCL), as a form of more calming and relaxed response during discussions of concerns or bad news (Reblin et al., 2012; Williams et al., 1975). Conversely, other researchers found that talking to an empathetic therapist could increase autonomic arousal when discussing about the patients' general states or non-concerning topics (Finset et al., 2011; C. D. Marci et al., 2004). Chapter 3 suggested that positive emotional attributes of a virtual coach enhanced therapy engagement and treatment confidence. Our aim was to understand whether these emotional attributes also change their physiological responses, and how a VR coach's emotional attributes affect people's autonomic response during consultation and exposure. This chapter looks into the physiological data recorded during the randomised controlled test described in Chapter 3 to analyse their changes through objective data. The research questions (RQs) are:

RQ1: Does the VR heights scenario elicit changes in physiological responses in individuals with a fear of heights?

RQ2 [primary]: Do the emotional attributes of a virtual coach affect participants' physiological responses during a VR consultation?

RQ3: Is there a correlation between participants' physiological responses during the VR consultation and their psychological responses (therapeutic alliance and treatment credibility/expectancy)?

RQ4: Do the emotional attributes of a virtual coach influence participants' physiological responses during a subsequent virtual heights task?

For RQ1, we hypothesise that exposure to VR heights will induce physiological arousal, characterised by an increase in HR, a decrease in HRV, and an increase in EDA. For RQ2 to RQ4, we do not make directional predictions due to the mixed findings in the literature (Reblin et al., 2012; Williams et al., 1975; Finset et al., 2011). However, we hypothesise that the addition of emotional attributes in the VR coach will result in changes in participants' physiological responses.

4.2 METHOD

4.2.1 Study Design

This chapter focuses on the physiological data we collected in the same experimental study described in the Chapter 3. The study design is summarised in Figure 4.1. The details of apparatus, measures, participants and procedure can be found in Chapter 3.

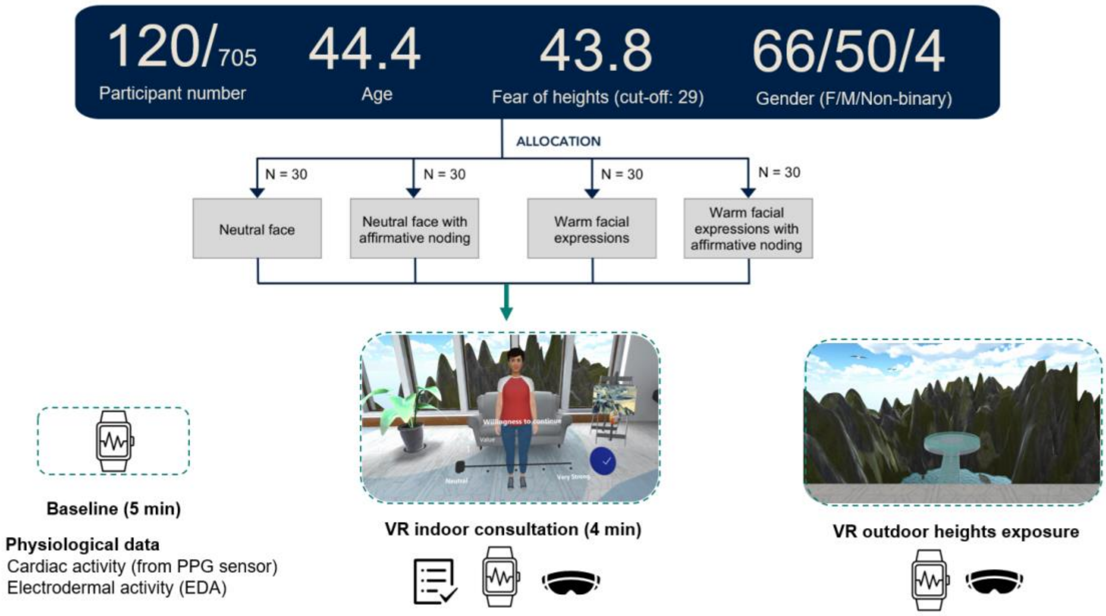


Figure 4.1. Summary of experimental design and participants.

4.2.2 Measures

Physiological data

As detailed Chapter 3, we recorded participants' cardiac activity (BVP) and EDA using the Empatica E4 wristband during the baseline, indoor VR, and outdoor VR sessions (see Figure 4.1). Data were collected via the E4 streaming server mode, with a Unity-integrated script, ensuring all signals were synchronised with the VR experience using millisecond-accurate timestamps.

Subjective measures

The details of subjective measures are also described in Chapter 3. In summary, we collected participants' subjective ratings of therapeutic alliance by the Virtual Therapist Alliance Scale (Miloff et al., 2020b) and treatment credibility/expectancy using the Credibility/Expectancy Questionnaire (Devilly & Borkovec, 2000) after the VR indoor consultation. The sense of presence was measured separately after both the VR indoor and VR outdoor sessions. Participants' demographic information (age, gender), experience with VR and their fear of heights scores were also collected during the screening stage.

4.2.3 Physiological Data Processing

Data pre-processing

The raw data from blood volume pulse (BVP) (64 Hz) and EDA (4 Hz) signals were cleaned using eyeballing, filtering, and outlier screening. We removed the data from a participant if more than 15% of their data were missing. For the EDA signals, we additionally removed their data if more than 15% of the values were zero or negative, which indicates loose skin contact of the device according to the manufacturer (Empatica, 2024). For both signals, we trimmed the first and last 3 seconds to reduce the likelihood of bad readings at either end.

PPG signals (recorded as blood volume pressure) were processed using 2nd order Butterworth bandpass filter (0.6875-10 Hz) and Stationary Wavelet Transform (SWT) 7th level Daubechies mother wavelet (Nason & Silverman, 1995) to remove drift and high-frequency artifacts. EDA data underwent pre-processing of 2nd order low-pass filter (0.25 Hz) for noise elimination.

Feature extractions

We utilised the Neurokit2 Python package (Makowski et al., 2021) to calculate and extract key features commonly used in related VR physiological studies (Gupta et al., 2024; Luong & Holz, 2022; Martens et al., 2019) from the processed cardiac and EDA signals, including both time-domain and frequency-domain features. For cardiac data, we extracted the mean HR and the main HRV metrics: the root mean square of successive differences (RMSSD) and the ratio of low-frequency to high-frequency signals (LF/HF). For EDA data, we used the continuous decomposition analysis method to separate it into the tonic SCL component (reflecting the background level of skin conductance) and the phasic SCR component (capturing rapid, transient responses to specific stimuli) (Benedek & Kaernbach, 2010; Makowski et al., 2021). We then calculated the mean for the tonic SCL, the SCR peaks per minute (PeaksPerMin), and the average SCR peak amplitude (PeakAmp).

Features were extracted for each participant both VR scenarios and baseline recording. The values from the indoor VR consultation and outdoor VR walk were normalised as a percentage of the baseline values (% Baseline) to account for individual differences at baseline.

4.2.4 Statistical Analysis

The extracted feature values from the physiological data were used for statistical testing. We tested the normality of the values across all conditions. As the Shapiro-Wilk test indicated all the data in the indoor VR and outdoor VR sessions deviated significantly from the normal distribution ($p < 0.01$), we used non-parametric statistical analysis methods for the main research questions.

For RQ1, a Wilcoxon Rank Sum Test was used to compare the pairwise differences of each participants' physiological data between the indoor VR consultation and outdoor VR. The effect size was calculated using rank-biserial correlation (r). For RQ2 and RQ4, we applied the Aligned Ranked Transform (ART) (Wobbrock et al., 2011) for nonparametric factorial data analysis to examine the effects of warm facial expressions and affirmative head nods on physiological responses. The effect size for each factor was calculated using Partial Eta Squared (η_p^2), and post-hoc contrast tests were performed using the ART-C tool (Elkin et al., 2021). For RQ3, we calculated the Spearman correlation to investigate the relationships between physiological and subjective measures, applying Bonferroni correction for multiple comparison.

All significance tests were made at the $\alpha = 0.05$ level. Data cleaning and processing was performed using *Python*'s *Pandas* and *NumPy* libraries (Harris et al., 2020; McKinney, 2010). The statistical analysis was done in *R* with the ART package.

4.3 RESULTS

4.3.1 RQ1: Physiological responses in Indoor VR VS Outdoor VR

For the pairwise comparison between the responses in indoor and outdoor VR sessions, we removed the participant's data from the pairwise statistical testing if one of the sessions was not available. This left us with 104 out of 120 pairs of data for within-subject comparison. The summary of the indoor vs. outdoor comparison is shown in Table 4.1.

	Indoor			**Outdoor**		
Physiological Measures	***Mean***	***Median***	***IQR***	***Mean***	***Median***	***IQR***
HR (% Baseline)	105%	104%	9%	106%	106%	8%
RMSSD (% Baseline)	106%	94%	58%	122%	105%	57%
LF/HF (% Baseline)	224%	102%	151%	166%	86%	120%
SCL (% Baseline)	297%	153%	157%	414%	215%	262%
PeaksPerMin (% Baseline)	124%	105%	87%	108%	79%	70%
PeakAmp (% Baseline)	191%	76%	140%	349%	103%	223%

Table 4.1. Comparison of physiological measures between indoor and outdoor VR environments, including mean, median, and interquartile range (IQR = Q3 − Q1).

Cardiac responses: Participants' HR was significantly higher (N = 104, W = 1639, p < 0.001, r = 0.347) in the outdoor VR (mean = 106%; median = 106%) than in the indoor VR (mean = 105%; median = 104%). Participants' HRV (measured by RMSSD) was significantly higher (N = 105, W = 1912, p = 0.005, r = 0.272) in the outdoor VR (mean = 122%; median = 105%) than in the indoor VR (mean = 106%; median = 94%).

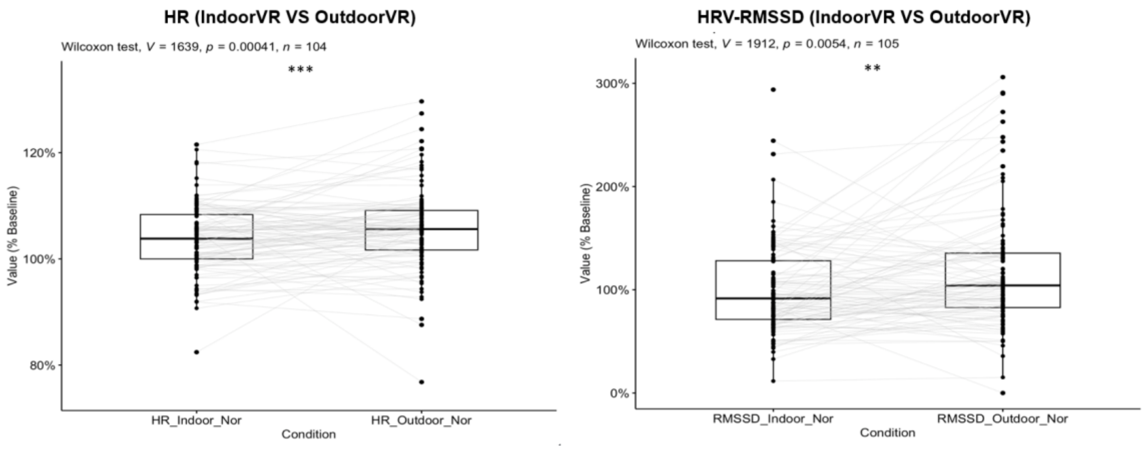


Figure 4.2. Indoor and Outdoor comparison of HR and HRV-RMSSD.

EDA: Participants' SCL was significantly higher (W = 1848, p = 0.003, r = 0.292) in the outdoor VR (mean = 414%; median = 215%) than in the indoor VR (mean = 297%; median = 153%). The PeaksPerMin was significantly smaller (W = 3124, p = 0.013, r = 0.244) in the outdoor VR (mean = 108%; median = 79%) than in the indoor VR (mean = 124%; median = 105%). The PeakAmp was significantly higher (W = 1859, p = 0.005, r = 0.277) in the outdoor VR (mean = 349%; median = 103%) than in the indoor VR (mean = 191%; median = 76%).

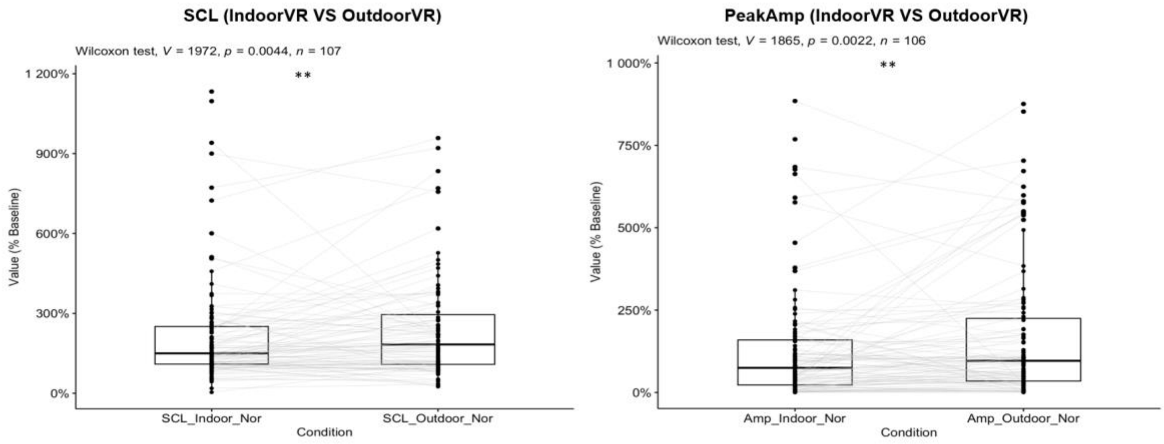


Figure 4.3. Indoor and Outdoor comparison of SCL and PeakAmp.

4.3.2 RQ2: Physiological responses in Indoor VR

We conducted the ART test to analyse the effects of warm facial expressions, affirmative head nods, and their interaction on physiological responses (HR, RMSSD, LF/HF, SCL, PeaksPerMin, PeakAmp). Table 4.2 summarises the results of statistical testing for indoor physiological data.

Cardiac responses indoor: There were no significant effects of warm facial expressions (F(1, 102) = 0.014, p = 0.904, η_p^2 < 0.001) or affirmative head nods (F(1, 102) = 0.087, p = 0.769, η_p^2 < 0.001) on HR. There was also no significant interaction between warm facial expressions and affirmative head nods on HR (F(1, 102) = 0.582, p = 0.447, η_p^2 = 0.006).

There were no significant effects from warm facial expressions (F(1, 102) = 0.773, p = 0.381, η_p^2 = 0.008) or affirmative head nods (F(1, 102) = 0.319, p = 0.574, η_p^2 = 0.003) on RMSSD. There was also no significant interaction between warm facial expressions and affirmative head nods on RMSSD (F(1, 102) = 0.041, p = 0.840, η_p^2 < 0.001).

There were no significant effects from warm facial expressions (F(1, 97) = 3.581, p = 0.061, η_p^2 = 0.036) or affirmative head nods (F(1, 97) = 1.098, p = 0.297, η_p^2 = 0.011) on LF/HF. There was also no significant interaction between warm facial expressions and affirmative head nods on LF/HF (F(1, 102) = 0.000, p = 0.991, η_p^2 < 0.001).

EDA indoor: Affirmative head nods (F(1, 104) = 3.638, p = 0.059, η_p^2 = 0.034) but not warm facial expressions (F(1, 104) = 2.879, p = 0.093, η_p^2 = 0.027) led to marginally non-significant increase in SCL. There was no significant interaction between warm facial expressions and affirmative head nods on SCL (F(1, 104) = 2.319, p = 0.131, η_p^2 = 0.022). Post-hoc pairwise comparison revealed a significant difference between the neutral face and warm face with nod condition (p-adj = 0.029), and between the warm face and warm face with nod condition (p-adj = 0.046).

Warm facial expressions (F(1, 104) = 3.889, p = 0.051, η_p^2 = 0.036) but not affirmative head nods (F(1, 104) = 1.115, p = 0.293, η_p^2 = 0.011) led to marginally non-significant decrease in PeaksPerMin. There was no significant interaction between warm facial expressions and affirmative head nods on PeaksPerMin (F(1, 104) = 0.010, p = 0.921, η_p^2 < 0.001). Post-hoc pairwise comparison revealed a significant difference between the neutral face and warm face with nod condition (p-adj = 0.031).

Warm facial expressions (F(1, 104) = 4.207, p = 0.043, η_p^2 = 0.040) but not affirmative head nods (F(1, 104) = 0.007, p = 0.935, η_p^2 < 0.001) led to a significantly higher PeakAmp. There was no significant interaction between warm facial expressions and affirmative head nods on PeakAmp (F(1, 103) = 3.329, p = 0.071, η_p^2 = 0.031). Post-hoc pairwise comparison revealed a significant difference between the neutral face with nod and warm face condition (p-adj < 0.001), and between the neutral face with nod and warm face with nod condition (p-adj = 0.033).

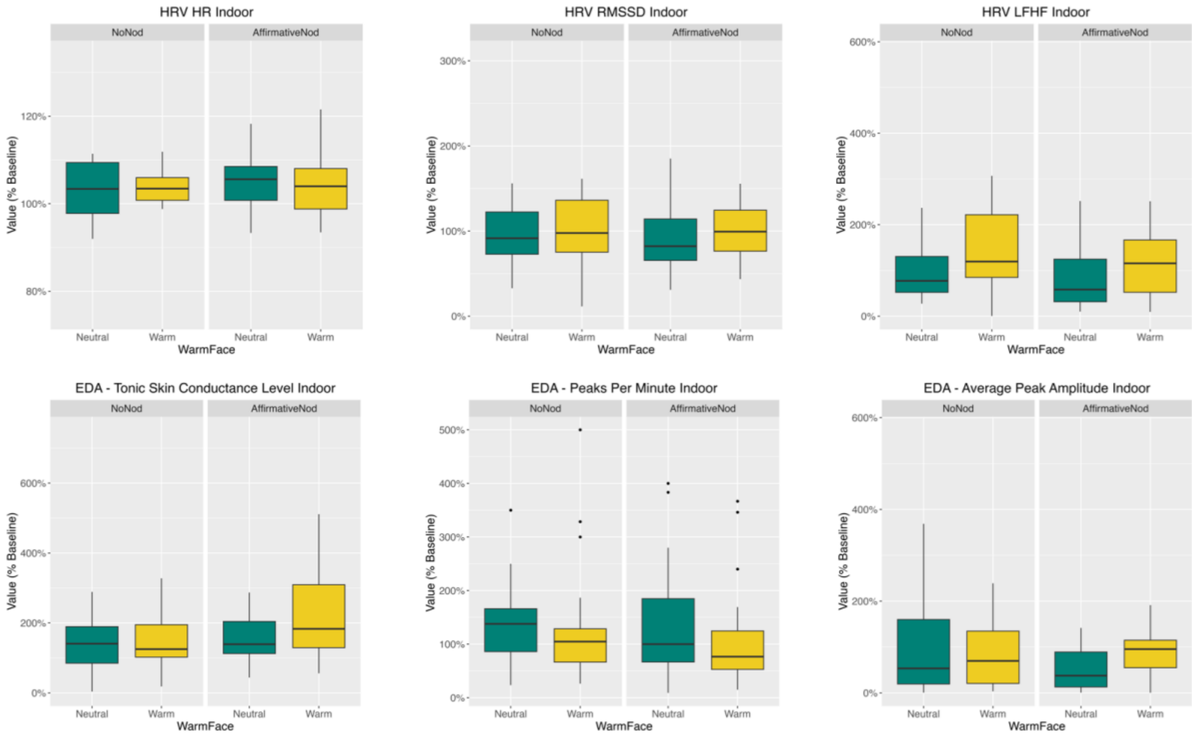


Figure 4.4. Cardiac response and EDA during indoor VR consultation (black dots indicate outliers)**.**

Indoor	HR		HRV-RMSSD		HRV-LFHF		SCL		PeakPerMin		PeakAmp	
	p	η_p^2	p	η_p^2	p	η_p^2	p	η_p^2	p	η_p^2	p	η_p^2
WarmFace	0.904	0.024	0.381	0.949	0.061	1.000	0.093	0.581	0.051 .	0.997	0.043 *	0.574
HeadNod	0.769	0.130	0.574	0.887	0.297	1.000	0.059 .	0.610	0.293	0.991	0.935	0.002
WarmFace:HeadNod	0.447	0.500	0.840	0.500	0.991	0.500	0.131	0.500	0.921	0.500	0.071	0.500

Table 4.2. Physiological responses during Indoor VR.

	HR (N =106)		RMSSD (N = 106)		SCL (N = 108)		PeakAmp (N =107)	
Subjective Measures	rho	Adj-p value	rho	Adj-p value	rho	Adj-p value	rho	Adj-p value
VirtualCoachAlliance	0.109	1.000	-0.014	1.000	0.277	0.015 *	0.158	0.412
Credibility	0.307	0.006 **	0.050	1.000	0.070	1.000	-0.052	1.000
Expectancy	0.361	0.001 **	-0.034	1.000	-0.005	1.000	-0.106	1.000
PresenceIndoor	0.216	0.107	-0.136	0.667	0.078	1.000	0.122	0.847

Table 4.3. Correlation between physiological and subjective measure during Indoor VR.

4.3.3 RQ3: Correlation between physiological and subjective measures

The results of the non-parametric correlation test (see in Figure 4.5) between physiological responses during indoor VR consultation and subjective measures were summarised in Table 4.3.

HR (N = 106) was positively correlated with treatment credibility (rho = 0.307, p-adj = 0.006) and expectancy (rho = 0.361, p = 0.001). SCL (N = 108) was positive correlated with the therapeutic alliance (rho = 0.277, p-adj = 0.015). There were no significant correlations between RMSSD/SCL and any of the other subjective measures.

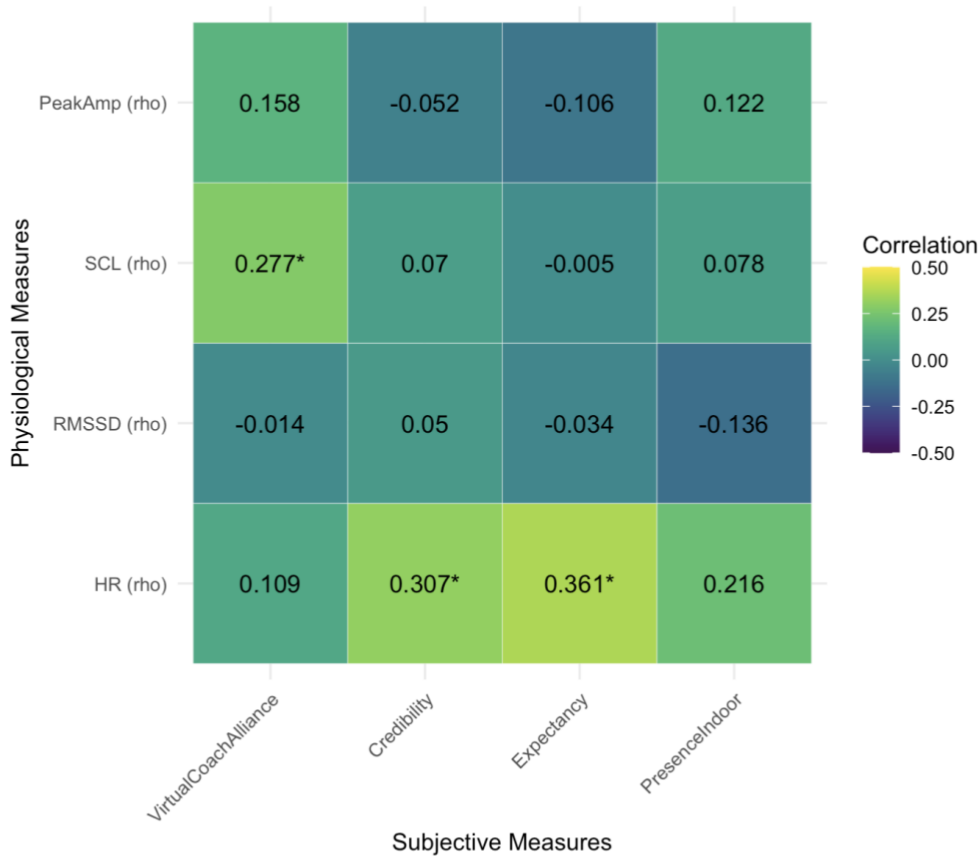


Figure 4.5. Correlation between physiological and subjective measures. (* indicates the significant correlation).

4.3.4 RQ4: Physiological responses in Outdoor VR

Table 4.4 summarises the results of statistical testing for outdoor physiological data.

Cardiac responses outdoor: There were no significant effects of warm facial expressions ($F(1, 107) = 0.081$, $p = 0.777$, $\eta_p^2 < 0.001$) or affirmative head nods ($F(1, 107) = 0.316$, $p = 0.575$, $\eta_p^2 = 0.003$) on HR. There was also no significant interaction between warm facial expressions and affirmative head nods on HR ($F(1, 107) = 1.340$, $p = 0.250$, $\eta_p^2 = 0.012$).

There were no significant effects from warm facial expressions ($F(1, 108) = 0.744$, $p = 0.390$, $\eta_p^2 = 0.007$) or affirmative head nods ($F(1, 108) = 0.119$, $p = 0.731$, $\eta_p^2 = 0.001$) on RMSSD. There was also no significant interaction between warm facial expressions and affirmative head nods on RMSSD ($F(1, 108) = 1.682$, $p = 0.197$, $\eta_p^2 = 0.015$).

Warm facial expressions ($F(1, 105) = 4.512$, $p = 0.036$, $\eta_p^2 = 0.041$) but not affirmative head nods ($F(1, 105) = 0.033$, $p = 0.856$, $\eta_p^2 < 0.001$) led to a significantly higher LF/HF value. There was no significant interaction between warm facial expressions and affirmative head nods on SCL ($F(1, 105) = 0.373$, $p = 0.542$, $\eta_p^2 = 0.004$). Post-hoc pairwise comparisons revealed a significant difference between the neutral face with nod and warm face condition (p-adj = 0.017).

EDA outdoor: There were no significant effects of warm facial expressions ($F(1, 106) = 0.159$, $p = 0.691$, $\eta_p^2 = 0.002$) or affirmative head nods ($F(1, 106) = 0.726$, $p = 0.396$, $\eta_p^2 = 0.007$) on SCL. There was also no significant interaction between warm facial expressions and affirmative head nods on SCL ($F(1, 106) = 0.197$, $p = 0.658$, $\eta_p^2 = 0.002$).

There were no significant effects of warm facial expressions (F(1, 107) = 1.686, p = 0.197, $\eta_p^2 = 0.016$) or affirmative head nods (F(1, 107) = 2.116, p = 0.149, $\eta_p^2 = 0.019$) on PeaksPerMin. There was also no significant interaction between warm facial expressions and affirmative head nods on PeaksPerMin (F(1, 107) = 0.832, p = 0.364, $\eta_p^2 = 0.008$).

There were no significant effects of warm facial expressions (F(1, 104) = 3.140, p = 0.079, $\eta_p^2 = 0.029$) or affirmative head nods (F(1, 104) = 0.362, p = 0.549, $\eta_p^2 = 0.003$) on PeakAmp. There was also marginally no significant interaction between warm facial expressions and affirmative head nods on PeakAmp (F(1, 104) = 3.503, p = 0.064, η_p^2 = 0.033). Post-hoc comparison revealed no significant pairwise difference.

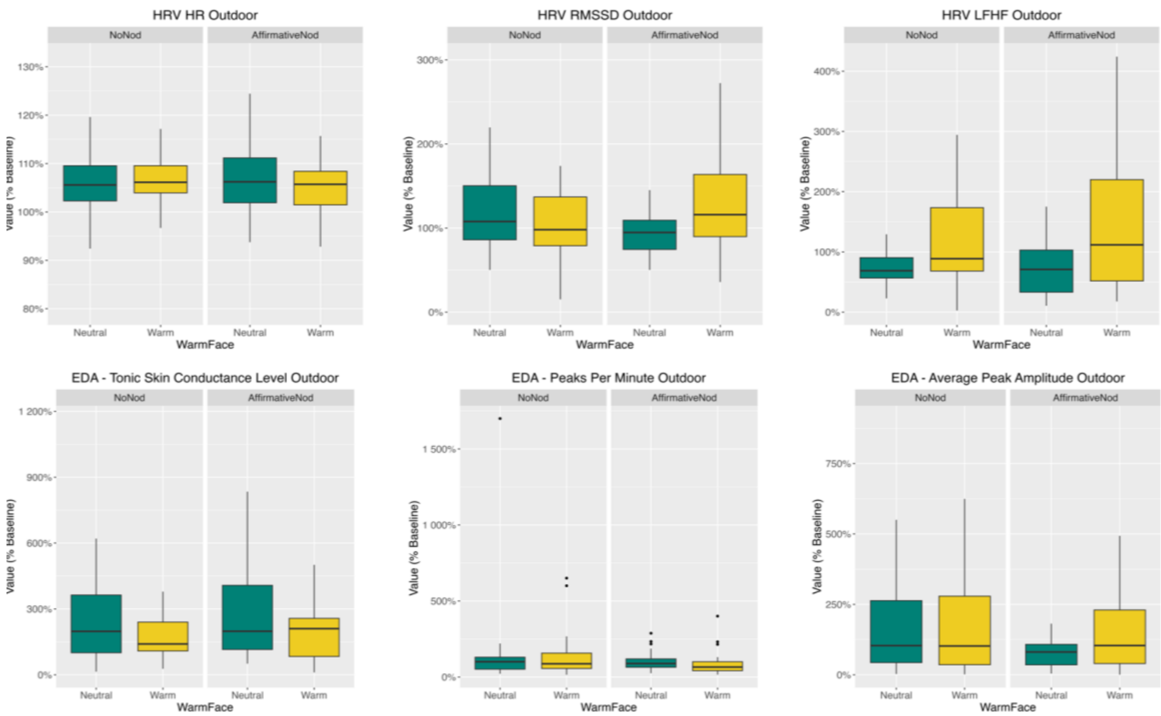


Figure 4.6. Cardiac response and EDA during outdoor VR heights exposure (black dots indicate outliers).

4.3.5 Additional Results

Indoor vs. outdoor VR presence: Given both the indoor presence and outdoor presence are non-normally distributed ($p < 0.01$), we used the Wilcoxon Rank Sum Test to compare the presence in indoor VS outdoor VR. Participants' presence was significantly higher ($N = 120$, $W = 269$, $p < 0.001$) in the outdoor VR (mean = 4.43, median = 5) than in the indoor VR (mean = 3.79, median = 4).

Fear of heights and outdoor physiological responses: The Spearman's correlation test indicated that there was no significant correlation between participants' fear of heights scores and HR (rho = 0.024, $p = 0.802$), RMSSD (rho = 0.031, $p = 0.744$), SCL (rho = -0.019, $p = 0.843$) or PeakAmp (rho = 0.052, $p = 0.588$) during the outdoor VR exposure.

Outdoor	HR		HRV-RMSSD		HRV-LFHF		SCL		PeakPerMin		PeakAmp	
	p	η_p^2	p	η_p^2	p	η_p^2	p	η_p^2	p	η_p^2	p	η_p^2
WarmFace	0.777	0.057	0.390	0.309	0.036 *	0.919	0.691	0.475	0.197	0.671	0.079	0.483
HeadNod	0.575	0.191	0.731	0.067	0.856	0.081	0.396	0.798	0.149	0.736	0.549	0.100
WarmFace:HeadNod	0.250	0.500	0.197	0.500	0.542	0.500	0.658	0.500	0.364	0.500	0.064	0.500

Table 4.4. Physiological responses in Outdoor VR.

4.4 DISCUSSION

In this chapter, we examined the physiological impacts of the VR and the detailed design of VHs through a VR mental health application, which consisted of an indoor consultation and an outdoor heights exposure. The study addressed several research questions. For VR heights exposure and physiological responses (RQ1), the outdoor heights exposure elicited higher HR and EDA, but also increased HRV (RMSSD) compared to the indoor VR. For the positive emotional attributes and physiological responses (RQ2), positive emotional attributes from the VR coach induced stronger physiological arousal in EDA during consultation. We also looked at the correlation between people's physiological responses and subjective perception (RQ3), finding that physiological activities were positively correlated with the therapeutic alliance and treatment credibility/expectancy during the consultation. Regarding the impact of a VR coach on subsequent exposure (RQ4), interacting with a virtual coach with positive emotional attributes during the consultation did not affect physiological responses in subsequent heights exposure, except for LF/HF. These findings offering further evidence of VR as an affective stimulus and highlight the importance of VHs' emotional attributes and their psychophysical influence in therapeutic settings.

As a validation test of the VR heights scenario, the results of RQ1 verified that the simulated environment effectively triggered physiological responses, most of which were similar to the stressful responses in real-life situations (Wonghempoom et al., 2023). Our findings partially align with previous VR studies on acrophobia, where individuals showed immediate physiological arousal, with increased HR and elevated EDA (tonic SCL and phasic SCR) during the exposure (Diemer et al., 2016; Peterson et al., 2018). These results provide further evidence that VR can elicit physiological responses that are compatible with fear and stress, which are common responses to the virtual heights scenario, though contrasting outcomes have been reported in other studies (Peterson et al., 2018; Raffegeau et al., 2020). Notably, we observed an unexpected increase in HRV, specifically RMSSD, during heights exposure. While higher RMSSD typically indicates parasympathetic dominance and relaxation (Shaffer & Ginsberg, 2017), this pattern contrasts with the heightened HR and EDA results, and the expected effects from the heights simulation. Similar findings have been reported when comparing calm VR environments to more stimulating ones (Martens et al., 2019), suggesting that RMSSD in very short-term recordings (<5 minutes) may have nuanced interpretations and potentially non-linear relationships with stress responses. Furthermore, when examining the potential association between the initial assessment of their fear of heights and their physiological responses during VR, we found no significant correlation between the fear of heights scores and physiological responses. This finding mirrors results from other studies, where individuals exhibited elevated HR and EDA despite reporting no subjective fear (Diemer et al., 2016) or showed no physiological arousal despite having a fear of heights (Wuehr et al., 2019). In summary, while the VR heights scenario effectively induces physiological responses, the dynamics in HRV patterns and the lack of correlation with subjective fear scores highlight the complexity of psychophysiological responses in VR. Further research is needed to clarify these relationships and explore the extent to which VR can reflect fear levels.

To our knowledge, this study is the first randomised controlled test to explore the physiological effects of detailed VR coach design. RQ2 results indicated that the coach's positive emotional attributes could increase EDA during the introductory consultation session, which focused on providing general information (e.g. rationale behind the fear of height). Specifically, affirmative head nods showed a trend toward increasing SCL, while warm facial expressions resulted in less frequent SCRs but with larger average peak amplitudes. In other words, the addition of affirmative nods produced a sustained increase in physiological arousal (tonic level), while the addition of warm facial expressions elicited more intense but less frequent bursts of arousal (phasic changes). Meanwhile, the positive emotional attributes did not influence cardiac responses.

Existing research in the relevant consultation scenarios has produced mixed results. Although some studies suggest that positive communication styles such as permissiveness and reassurance, decrease EDA due to the relief they provide (Reblin et al., 2012; Sep et al., 2014), our findings align with other observations where compassionate communication increased autonomic activities (e.g. higher SCL and larger SCR amplitude) (Finset et al., 2011; C. D. Marci et al., 2004). One possible explanation is that our VR consultation was more general and introductory, addressing a condition that is less severe, rather than deep personal concerns (e.g., life-threatening illness). This context suggests that the increased arousal may reflect positive appraisals or excitement, as indicated by the higher therapeutic alliance and confidence in the VR treatment, along with the positive correlation between therapeutic alliance and SCL (as found in RQ3 results). Another reason could be that the positive non-verbal behaviours of the VR coach, which were programmed with additional animation and details, heightened participants' attention focus on the coach and the consultation, leading to greater cognitive engagement, which consequently increased physiological activities. Research suggests that non-verbal cues, such as facial expressions and gestures, can significantly increase attention and cognitive loads, which are positively associated with physiological activation (Kleinsmith & Bianchi-Berthouze, 2013; Volonte et al., 2020). Regarding the non-significant effects on heart rate-related measures, a potential reason is that HR and HRV are often influenced by a broad range of factors, including respiratory patterns, motion artefacts and individual differences in autonomic regulation (Shaffer & Ginsberg, 2017). In contrast, EDA may be more directly responsive to the immediate emotional and attentional demands of the interaction, which might explain why EDA showed significant changes while cardiac measures did not.

Interestingly, the effects of positive non-verbal behaviours were not observed in participants' physiological responses during the outdoor heights exposure, except for an increase in the LF/HF ratio due to warm facial expressions. This ratio change indicates greater sympathetic activation, commonly associated with heightened emotional arousal or preparation for a 'fight or flight' response. However, other indices such as HR and EDA remained unaffected by the positive emotional attributes. This discrepancy could be attributed to the short duration of the indoor consultation and the lack of close support from the coach when participants walked on the virtual plank outdoors. As a result, the differences observed during the consultation (with a relatively small effect size) were not sustained during the exposure task. Another possible explanation relates to the complex interaction between the sympathetic and parasympathetic nervous systems. While the SNS activates the 'fight or flight' response, the PNS works to calm the body down. In highly stressful scenarios, such as heights exposure for people with a fear of heights, the dominant activation of the SNS may override or mask the effects of positive non-verbal behaviours effective during the indoor consultation (McCorry, 2007). This SNS dominance was especially pronounced in the group facing the VR coach with warm facial expressions. Future research could explore extending the consultation duration or incorporating additional support, such as having the VR coach accompany participants during the exposure task, to determine whether these adjustments influence physiological responses in challenging scenarios and improve therapeutic outcomes.

The study has several limitations in terms of understanding the physiological insights in VR. Firstly, we only measured arousal and did not assess valence (the extent to which an emotion is positive or negative) in response to the virtual coach. We also did not measure subjective stress or fear after the heights exposure in the outdoor VR scenario. Valence is a key metric in understanding affective states (Barrett & Russell, 1999; Citron et al., 2014), and its absence makes it less clear how to interpret participants' physiological changes, especially in the consultation scenario where feelings of tension and relief can be experienced simultaneously (Voutilainen et al., 2018). However, we used the therapeutic alliance as a proxy for participants' positive emotions in the consultation. Additionally, we did not track changes in physiological responses throughout the VR consultation, nor did we analyse responses to specific lines of the VR coach's scripts. This limitation prevented us from identifying the detailed affective changes and precise emotional responses to the meaningful content delivered by the coach (DeVault et al., 2013; Giron et al., 2023). Furthermore, we did not measure post-task physiological responses to assess recovery and the state following heights exposure, potentially missing insights into the after-effects of the coach design in a challenging task. Additionally, as noted in Chapter 3, we did not consider the relationship between participants and the VR coach either, or did we account for diversity in the virtual coach's representation, such as gender, ethnicity, and body shape similarity to the participant. These factors could influence the study outcomes, as the extent to which people associate the VH with themselves also affect their physiological responses (Kocur et al., 2021).

In this chapter, we explored the effects of a VR coach's detailed design on participants' physiological responses. The findings indicate that specific positive non-verbal cues, such as affirmative nods and warm facial expressions, significantly increased EDA, which was potentially linked to heightened emotional engagement during the VR consultation. However, these effects did not extend to cardiac measures, and the positive cues did not sustain their influence during the subsequent VR heights exposure apart from the HRV-LF/HF. These results highlight the psychophysical impacts of carefully considering the VR coach's emotional attributes in automated VR therapy. Future research could consider extending consultation duration with the coach and integrating continuous support to better understand the potential for enhancing therapeutic outcomes in VR therapy through character design.

References

Aburumman, N., Gillies, M., Ward, J. A., & Hamilton, A. F. de C. (2022). Nonverbal communication in virtual reality: Nodding as a social signal in virtual interactions. *International Journal of Human-Computer Studies*, *164*, 102819. https://doi.org/10.1016/j.ijhcs.2022.102819

Amaral, C., Mouga, S., Simões, M., Pereira, H. C., Bernardino, I., Quental, H., Playle, R., McNamara, R., Oliveira, G., & Castelo-Branco, M. (2018). A Feasibility Clinical Trial to Improve Social Attention in Autistic Spectrum Disorder (ASD) Using a Brain Computer Interface. *Frontiers in Neuroscience*, *12*, 477. https://doi.org/10.3389/fnins.2018.00477

Amestoy Alonso, B., Donegan, T., Calvis, I., Swidrak, J., Rodriguez, E., Vargas-Reverón, C. L., Combalia, A., Oliva Martinez, R., & Sanchez-Vives, M. V. (2024). Focus groups in the metaverse: Shared virtual spaces for patients, clinicians, and researchers. *Frontiers in Virtual Reality*, *5*. https://doi.org/10.3389/frvir.2024.1432282

APA. (2020). *Publication manual of the American Psychological Association, 7th ed* (pp. xxii, 427). American Psychological Association. https://doi.org/10.1037/0000165-000

Ardito, R. B., & Rabellino, D. (2011). Therapeutic Alliance and Outcome of Psychotherapy: Historical Excursus, Measurements, and Prospects for Research. *Frontiers in Psychology*, *2*, 270. https://doi.org/10.3389/fpsyg.2011.00270

Artiran, S., Bedmutha, P. S., & Cosman, P. (2024). Analysis of Gaze, Head Orientation, and Joint Attention in Autism With Triadic VR Interviews. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *32*, 759–769. https://doi.org/10.1109/TNSRE.2024.3363728

Arza, A., Garzón-Rey, J. M., Lázaro, J., Gil, E., Lopez-Anton, R., de la Camara, C., Laguna, P., Bailon, R., & Aguiló, J. (2019). Measuring acute stress response through physiological signals: Towards a quantitative assessment of stress. *Medical & Biological Engineering & Computing*, *57*(1), 271–287. https://doi.org/10.1007/s11517-018-1879-z

Ascione, M., Carulla-Roig, M., Miquel-Nabau, H., Porras-Garcia, B., Meschberger-Annweiler, F.-A., Serrano-Troncoso, E., Ferrer-Garcia, M., Moreno-Sánchez, M., & Gutierrez-Maldonado, J. (2023). Attentional Bias Modification Training Based on Virtual Reality and Eye Tracking in Anorexia Nervosa Patients. *Journal of Clinical Medicine*, *12*(18), 5932. https://doi.org/10.3390/jcm12185932

Atherton, S., Antley, A., Evans, N., Cernis, E., Lister, R., Dunn, G., Slater, M., & Freeman, D. (2016). Self-Confidence and Paranoia: An Experimental Study Using an Immersive Virtual Reality Social Situation. *Behavioural and Cognitive Psychotherapy*, *44*(1), 56–64. https://doi.org/10.1017/S1352465814000496

Aymerich-Franch, L., Kizilcec, R. F., & Bailenson, J. N. (2014). The Relationship between Virtual Self Similarity and Social Anxiety. *Frontiers in Human Neuroscience*, *8*, 944. https://doi.org/10.3389/fnhum.2014.00944

Aziz, S., & Komogortsev, O. (2022). An Assessment of the Eye Tracking Signal Quality Captured in the HoloLens 2. *2022 Symposium on Eye Tracking Research and Applications*, 1–6. https://doi.org/10.1145/3517031.3529626

Banakou, D., Beacco, A., Neyret, S., Blasco-Oliver, M., Seinfeld, S., & Slater, M. (2020). Virtual body ownership and its consequences for implicit racial bias are dependent on social context. *Royal Society Open Science*, *7*(12), 201848. https://doi.org/10.1098/rsos.201848

Banakou, D., Johnston, T., Beacco, A., Senel, G., & Slater, M. (2024). Desensitizing Anxiety Through Imperceptible Change: Feasibility Study on a Paradigm for Single-Session Exposure Therapy for Fear of Public Speaking. *JMIR Formative Research*, *8*(1), e52212. https://doi.org/10.2196/52212

Barisnikov, K., Thomasson, M., Stutzmann, J., & Lejeune, F. (2021). Sensitivity to Emotion Intensity and Recognition of Emotion Expression in Neurotypical Children. *Children*, *8*(12), 1108. https://doi.org/10.3390/children8121108

Barrett, L. F., & Russell, J. A. (1999). The Structure of Current Affect: Controversies and Emerging Consensus. *Current Directions in Psychological Science*, *8*(1), 10–14. https://doi.org/10.1111/1467-8721.00003

Bektas, S., Natali, L., Rowlands, K., Valmaggia, L., Di Pietro, J., Mutwalli, H., Himmerich, H., Treasure, J., & Cardi, V. (2023). Exploring Correlations of Food-Specific Disgust with Eating Disorder Psychopathology and Food Interaction: A Preliminary Study Using Virtual Reality. *Nutrients*, *15*(20), Article 20. https://doi.org/10.3390/nu15204443

Bell, I. H., Nicholas, J., Alvarez-Jimenez, M., Thompson, A., & Valmaggia, L. (2020). Virtual reality as a clinical tool in mental health research and practice. *Dialogues in Clinical Neuroscience*, *22*(2), 169–177. https://doi.org/10.31887/DCNS.2020.22.2/lvalmaggia

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, *190*(1), 80–91. https://doi.org/10.1016/j.jneumeth.2010.04.028

Bickmore, T. W., & Picard, R. W. (2004). Towards caring machines. *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, 1489–1492. https://doi.org/10.1145/985921.986097

Bolinski, F., Etzelmüller, A., De Witte, N. A. J., van Beurden, C., Debard, G., Bonroy, B., Cuijpers, P., Riper, H., & Kleiboer, A. (2021). Physiological and self-reported arousal in virtual reality versus face-to-face emotional activation and cognitive restructuring in university students: A crossover experimental study using wearable monitoring. *Behaviour Research and Therapy*, *142*, 103877. https://doi.org/10.1016/j.brat.2021.103877

Bombari, D., Schmid Mast, M., Canadas, E., & Bachmann, M. (2015). Studying social interactions through immersive virtual environment technology: Virtues, pitfalls, and future challenges. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00869

Bönsch, A., Radke, S., Ehret, J., Habel, U., & Kuhlen, T. W. (2020). The Impact of a Virtual Agent's Non-Verbal Emotional Expression on a User's Personal Space Preferences. *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*, 1–8. https://doi.org/10.1145/3383652.3423888

Brander, M., Egger, S. T., Hürlimann, N., Seifritz, E., Sumner, R. W., Vetter, S., & Magnenat, S. (2021). Virtual Reality Human–Human Interface to Deliver Psychotherapy to People Experiencing Auditory Verbal Hallucinations: Development and Usability Study. *JMIR Serious Games*, *9*(2). https://doi.org/10.2196/26820

Braun, H. I., Jackson, D. N., & Wiley, D. E. (Eds.). (2001). Socially Desirable Responding: The Evolution of a Construct. In *The Role of Constructs in Psychological and Educational Measurement* (0 ed., pp. 61–84). Routledge. https://doi.org/10.4324/9781410607454-10

Brenner, P. S., & DeLamater, J. (2016). Lies, Damned Lies, and Survey Self-Reports? Identity as a Cause of Measurement Bias. *Social Psychology Quarterly*, *79*(4), 333–354. https://doi.org/10.1177/0190272516628298

Brett, J., Staniszewska, S., Mockford, C., Herron-Marx, S., Hughes, J., Tysall, C., & Suleman, R. (2014). A Systematic Review of the Impact of Patient and Public Involvement on Service Users, Researchers and Communities. *The Patient - Patient-Centered Outcomes Research*, *7*(4), 387–395. https://doi.org/10.1007/s40271-014-0065-0

Brinkman, W.-P., Veling, W., Dorrestijn, E., Sandino, G., Vakili, V., & Gaag, M. (2011). Virtual Reality to study responses to social environmental stressors in individuals with and without psychosis. *Annual Review of CyberTherapy and Telemedicine*, 86–91. https://doi.org/10.3233/978-1-6-750-766-6-86

Brown, P., Waite, F., Lambe, S., Jones, J., Jenner, L., Diamond, R., & Freeman, D. (2022). Automated Virtual Reality Cognitive Therapy (gameChange) in Inpatient Psychiatric Wards: Qualitative Study of Staff and Patient Views Using an Implementation Framework. *JMIR Formative Research*, *6*(4), e34225. https://doi.org/10.2196/34225

Brown, P., Waite, F., Rovira, A., Nickless, A., & Freeman, D. (2020). Virtual reality clinical-experimental tests of compassion treatment techniques to reduce paranoia. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-64957-7

Bruder, L. R., Scharer, L., & Peters, J. (2020). *A Virtual Reality set-up for the combined assessment of behavioral and psychophysiological cue-reactivity in gambling disorder* (p. 2020.08.07.237826). bioRxiv. https://doi.org/10.1101/2020.08.07.237826

Burin, D., Cavanna, G., Rabellino, D., Kotozaki, Y., & Kawashima, R. (2022). Neuroendocrine Response and State Anxiety Due to Psychosocial Stress Decrease after a Training with Subject's Own (but Not Another) Virtual Body: An RCT Study. *International Journal of Environmental Research and Public Health*, *19*(10), 6340. https://doi.org/10.3390/ijerph19106340

Can, Y., Mahesh, B., & Andre, E. (2023). Approaches, Applications, and Challenges in Physiological Emotion Recognition—A Tutorial Overview. *Proceedings of the IEEE*, *PP*, 1–27. https://doi.org/10.1109/JPROC.2023.3286445

Cassell, J. (2001). Embodied Conversational Agents: Representation and Intelligence in User Interfaces. *AI Magazine*, *22*(4), Article 4. https://doi.org/10.1609/aimag.v22i4.1593

Cerritelli, F., Chiera, M., Abbro, M., Megale, V., Esteves, J., Gallace, A., & Manzotti, A. (2021). The Challenges and Perspectives of the Integration Between Virtual and Augmented Reality and Manual Therapies. *Frontiers in Neurology*, *12*, 700211. https://doi.org/10.3389/fneur.2021.700211

Chang, Z., Bai, H., Zhang, L., Gupta, K., He, W., & Billinghurst, M. (2022). The impact of virtual agents' multimodal communication on brain activity and cognitive load in Virtual Reality. *Frontiers in Virtual Reality*, *3*. https://doi.org/10.3389/frvir.2022.995090

Chattopadhyay, D., Ma, T., Sharifi, H., & Martyn-Nemeth, P. (2020). Computer-Controlled Virtual Humans in Patient-Facing Systems: Systematic Review and Meta-Analysis. *Journal of Medical Internet Research*, *22*(7), e18839. https://doi.org/10.2196/18839

Cho, S., Ku, J., Park, J., Han, K., Lee, H., Choi, Y. K., Jung, Y.-C., Namkoong, K., Kim, J.-J., Kim, I. Y., Kim, S. I., & Shen, D. F. (2008a). Development and Verification of an Alcohol Craving–Induction Tool Using Virtual Reality: Craving Characteristics in Social Pressure Situation. *CyberPsychology & Behavior*, *11*(3), Article 3. https://doi.org/10.1089/cpb.2007.0149

Cho, S., Ku, J., Park, J., Han, K., Lee, H., Choi, Y. K., Jung, Y.-C., Namkoong, K., Kim, J.-J., Kim, I. Y., Kim, S. I., & Shen, D. F. (2008b). Development and Verification of an Alcohol Craving–Induction Tool Using Virtual Reality: Craving Characteristics in Social Pressure Situation. *CyberPsychology & Behavior*, *11*(3), 302–309. https://doi.org/10.1089/cpb.2007.0149

Choudhary, Z., Norouzi, N., Erickson, A., Schubert, R., Bruder, G., & Welch, G. F. (2023). Exploring the Social Influence of Virtual Humans Unintentionally Conveying Conflicting Emotions. *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, 571–580. https://doi.org/10.1109/VR55154.2023.00072

Christinaki, E., Vidakis, N., & Triantafyllidis, G. (2013). Facial expression recognition teaching to preschoolers with autism: A natural user interface approach. *Proceedings of the 6th Balkan Conference in Informatics*, 141–148. https://doi.org/10.1145/2490257.2490262

Citron, F. M. M., Gray, M. A., Critchley, H. D., Weekes, B. S., & Ferstl, E. C. (2014). Emotional valence and arousal affect reading in an interactive way: Neuroimaging evidence for an approach-withdrawal framework. *Neuropsychologia*, *56*(100), 79–89. https://doi.org/10.1016/j.neuropsychologia.2014.01.002

Ciuk, D., Troy, A. K., & Jones, M. C. (2015). Measuring Emotion: Self-Reports vs. Physiological Indicators. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2595359

Clavel, C., Plessier, J., Martin, J.-C., Ach, L., & Morel, B. (2009). Combining Facial and Postural Expressions of Emotions in a Virtual Character. In Z. Ruttkay, M. Kipp, A. Nijholt, & H. H. Vilhjálmsson (Eds.), *Intelligent Virtual Agents* (pp. 287–300). Springer. https://doi.org/10.1007/978-3-642-04380-2_31

Cornwell, B. R., Johnson, L., Berardi, L., & Grillon, C. (2006). Anticipation of Public Speaking in Virtual Reality Reveals a Relationship Between Trait Social Anxiety and Startle Reactivity. *Biological Psychiatry*, *59*(7), 664–666. https://doi.org/10.1016/j.biopsych.2005.09.015

Del Piccolo, L., & Finset, A. (2018). Patients' autonomic activation during clinical interaction: A review of empirical studies. *Patient Education and Counseling*, *101*(2), 195–208. https://doi.org/10.1016/j.pec.2017.08.007

DeVault, D., Artstein, R., Benn, G., Dey, T., Fast, E., Gainer, A., Georgila, K., Gratch, J., Hartholt, A., Lhommet, M., Lucas, G., Marsella, S., Morbini, F., Nazarian, A., Scherer, S., Stratou, G., Suri, A., Traum, D., Wood, R., … Morency, L.-P. (2014). SimSensei kiosk: A virtual human interviewer for healthcare decision support. *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-Agent Systems*, 1061–1068.

DeVault, D., Georgila, K., Artstein, R., Morbini, F., Traum, D., Scherer, S., Rizzo, A. S., & Morency, L.-P. (2013). Verbal indicators of psychological distress in interactive dialogue with a virtual human. In M. Eskenazi, M. Strube, B. Di Eugenio, & J. D. Williams (Eds.), *Proceedings of the SIGDIAL 2013 Conference* (pp. 193–202). Association for Computational Linguistics. https://aclanthology.org/W13-4032

Devilly, G. J., & Borkovec, T. D. (2000). Psychometric properties of the credibility/expectancy questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*, *31*(2), 73–86. https://doi.org/10.1016/S0005-7916(00)00012-4

Dickinson, P., Gerling, K., Hicks, K., Murray, J., Shearer, J., & Greenwood, J. (2019). Virtual reality crowd simulation: Effects of agent density on user experience and behaviour. *Virtual Reality*, *23*(1), 19–32. https://doi.org/10.1007/s10055-018-0365-0

Diemer, J., Lohkamp, N., Mühlberger, A., & Zwanzger, P. (2016). Fear and physiological arousal during a virtual height challenge—Effects in patients with acrophobia and healthy controls. *Journal of Anxiety Disorders*, *37*, 30–39. https://doi.org/10.1016/j.janxdis.2015.10.007

Dimascio, A., Boyd, R. W., & Greenblatt, M. (1957). Physiological Correlates of Tension and Antagonism During Psychotherapy: A Study of 'Interpersonal Physiology'. *Psychosomatic Medicine*, *19*(2), 99.

Dobre, G. C., Gillies, M., & Pan, X. (2022). Immersive machine learning for social attitude detection in virtual reality narrative games. *Virtual Reality*, *26*(4), 1519–1538. https://doi.org/10.1007/s10055-022-00644-4

Dobs, K., Bülthoff, I., & Schultz, J. (2018). Use and Usefulness of Dynamic Face Stimuli for Face Perception Studies—A Review of Behavioral Findings and Methodology. *Frontiers in Psychology*, *9*, 1355. https://doi.org/10.3389/fpsyg.2018.01355

Dong, Y., & Peng, C. (2019). Real-Time Large Crowd Rendering with Efficient Character and Instance Management on GPU. *International Journal of Computer Games Technology*, *2019*, e1792304. https://doi.org/10.1155/2019/1792304

du Sert, O. P., Potvin, S., Lipp, O., Dellazizzo, L., Laurelli, M., Breton, R., Lalonde, P., Phraxayavong, K., O'Connor, K., Pelletier, J.-F., Boukhalfi, T., Renaud, P., & Dumais, A. (2018). Virtual reality therapy for refractory auditory verbal hallucinations in schizophrenia: A pilot clinical trial. *Schizophrenia Research*, *197*, 176–181. https://doi.org/10.1016/j.schres.2018.02.031

Dyer, M. L., Attwood, A. S., Penton-Voak, I. S., & Munafò, M. R. (2022). The role of state and trait anxiety in the processing of facial expressions of emotion. *Royal Society Open Science*, *9*(1), 210056. https://doi.org/10.1098/rsos.210056

Ekman, P. (2004). *Emotions Revealed: Understanding Faces and Feelings*. https://www.amazon.co.uk/Emotions-Revealed-Understanding-Faces-Feelings/dp/0753817659

El Ali, A., Ney, R., Van Berlo, Z. M. C., & Cesar, P. (2023). Is that My Heartbeat? Measuring and Understanding Modality-Dependent Cardiac Interoception in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, *29*(11), 4805–4815. https://doi.org/10.1109/TVCG.2023.3320228

Elkin, L. A., Kay, M., Higgins, J. J., & Wobbrock, J. O. (2021). An Aligned Rank Transform Procedure for Multifactor Contrast Tests. *The 34th Annual ACM Symposium on User Interface Software and Technology*, 754–768. https://doi.org/10.1145/3472749.3474784

Elshan, E., Zierau, N., Engel, C., Janson, A., & Leimeister, J. M. (2022). Understanding the Design Elements Affecting User Acceptance of Intelligent Agents: Past, Present and Future. *Information Systems Frontiers*, *24*(3), 699–730. https://doi.org/10.1007/s10796-021-10230-9

Empatica. (2024). *E4 data—EDA potential issues*. Empatica Support. https://support.empatica.com/hc/en-us/articles/204676479-E4-data-EDA-potential-issues

Faita, C., Vanni, F., Lorenzini, C., Carrozzino, M., Tanca, C., & Bergamasco, M. (2015). Perception of Basic Emotions from Facial Expressions of Dynamic Virtual Avatars. In L. T. De Paolis & A. Mongelli (Eds.), *Augmented and Virtual Reality* (pp. 409–419). Springer International Publishing. https://doi.org/10.1007/978-3-319-22888-4_30

Falconer, C. J., Rovira, A., King, J. A., Gilbert, P., Antley, A., Fearon, P., Ralph, N., Slater, M., & Brewin, C. R. (2016). Embodying self-compassion within virtual reality and its effects on patients with depression. *BJPsych Open*, *2*(1), 74–80. https://doi.org/10.1192/bjpo.bp.115.002147

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Felnhofer, A., Knaust, T., Weiss, L., Goinska, K., Mayer, A., & Kothgassner, O. D. (2024). A Virtual Character's Agency Affects Social Responses in Immersive Virtual Reality: A Systematic Review and Meta-Analysis. *International Journal of Human–Computer Interaction*, *40*(16), 4167–4182. https://doi.org/10.1080/10447318.2023.2209979

Ferrer-Garcia M., Porras-Garcia B., Gonzalez-Ibanez C., Gracia-Blanes M., Vilalta-Abella F., Pla-Sanjuanelo J., Riva G., Dakanalis A., Achotegui-Loizate J., Talarn-Caparros A., Ribas-Sabate J., Andreu-Gracia A., Diaz-Marsa M., Monras-Arnau M., Serrano-Troncoso E., Treasure J., & Gutierrez-Maldonado J. (2017). Does owning a 'fatter' virtual body increase body anxiety in college students? *Annual Review of CyberTherapy and Telemedicine*, *15*, 147–153.

Ferrer-Garcia, M., Porras-Garcia, B., Moreno, M., Bertomeu, P., & Maldonado, J. G. (2018). Embodiment in different size virtual bodies produces changes in women's body image distortion and dissatisfaction. *Annual Review of CyberTherapy and Telemedicine*, *16*(Human Factors Engineering [4010]), 111–117.

Fett, A.-K. J., Shergill, S. S., Joyce, D. W., Riedl, A., Strobel, M., Gromann, P. M., & Krabbendam, L. (2012). To trust or not to trust: The dynamics of social interaction in psychosis. *Brain*, *135*(3), 976–984. https://doi.org/10.1093/brain/awr359

Finset, A., Stensrud, T. L., Holt, E., Verheul, W., & Bensing, J. (2011). Electrodermal activity in response to empathic statements in clinical interviews with fibromyalgia patients. *Patient Education and Counseling*, *82*(3), 355–360. Scopus. https://doi.org/10.1016/j.pec.2010.12.029

Fisher, S., Abdullah, A., Charvin, I., Da Fonseca, D., & Bat-Pitault, F. (2020). Comparison of body image evaluation by virtual reality and paper-based figure rating scales in adolescents with anorexia nervosa: Retrospective study. *Eating and Weight Disorders: EWD*, *25*(3), 735–743. https://doi.org/10.1007/s40519-019-00680-1

Flavián, C., Ibáñez-Sánchez, S., & Orús, C. (2021). The influence of scent on virtual reality experiences: The role of aroma-content congruence. *Journal of Business Research*, *123*, 289–301. https://doi.org/10.1016/j.jbusres.2020.09.036

Fornells-Ambrojo, M., Elenbaas, M., Barker, C., Swapp, D., Navarro, X., Rovira, A., Sanahuja, J. M. T., & Slater, M. (2016). Hypersensitivity to Contingent Behavior in Paranoia: A New Virtual Reality Paradigm. *The Journal of Nervous and Mental Disease*, *204*(2), 148–152. https://doi.org/10.1097/NMD.0000000000000414

Fornells-Ambrojo, M., Freeman, D., Slater, M., Swapp, D., Antley, A., & Barker, C. (2015). How Do People with Persecutory Delusions Evaluate Threat in a Controlled Social Environment? A Qualitative Study Using Virtual Reality. *Behavioural and Cognitive Psychotherapy*, *43*(1), 89–107. https://doi.org/10.1017/S1352465813000830

Freeman, D., Bradley, J., Antley, A., Bourke, E., DeWeever, N., Evans, N., Černis, E., Sheaves, B., Waite, F., Dunn, G., Slater, M., & Clark, D. M. (2016). Virtual reality in the treatment of persecutory delusions: Randomised controlled experimental study testing how to reduce delusional conviction. *British Journal of Psychiatry*, *209*(1), 62–67. https://doi.org/10.1192/bjp.bp.115.176438

Freeman, D., Dunn, G., Murray, R. M., Evans, N., Lister, R., Antley, A., Slater, M., Godlewska, B., Cornish, R., Williams, J., Di Simplicio, M., Igoumenou, A., Brenneisen, R., Tunbridge, E. M., Harrison, P. J., Harmer, C. J., Cowen, P., & Morrison, P. D. (2015). How Cannabis Causes Paranoia: Using the Intravenous Administration of Δ 9 - Tetrahydrocannabinol (THC) to Identify Key Cognitive Mechanisms Leading to Paranoia. *Schizophrenia Bulletin*, *41*(2), 391–399. https://doi.org/10.1093/schbul/sbu098

Freeman, D., Freeman, J., Ahmed, M., Haynes, P., Beckwith, H., Rovira, A., Miguel, A. L., Ward, R., Bousfield, M., Riffiod, L., Kabir, T., Waite, F., & Rosebrock, L. (2024). Automated VR therapy for improving positive self-beliefs and psychological well-being in young patients with psychosis: A proof of concept evaluation of Phoenix VR self-confidence therapy. *Behavioural and Cognitive Psychotherapy*, *52*(3), 277–287. https://doi.org/10.1017/S1352465823000553

Freeman, D., Garety, P. A., Bebbington, P. E., Smith, B., Rollinson, R., Fowler, D., Kuipers, E., Ray, K., & Dunn, G. (2005). Psychological investigation of the structure of paranoia in a non-clinical population. *The British Journal of Psychiatry*, *186*(5), 427–435. https://doi.org/10.1192/bjp.186.5.427

Freeman, D., Garety, P. A., Bebbington, P., Slater, M., Kuipers, E., Fowler, D., Green, C., Jordan, J., Ray, K., & Dunn, G. (2005). The psychology of persecutory ideation II: A virtual reality experimental study. *The Journal of Nervous and Mental Disease*, *193*(5), 309–315. https://doi.org/10.1097/01.nmd.0000161686.53245.70

Freeman, D., Garety, P. A., Kuipers, E., Fowler, D., Bebbington, P. E., & Dunn, G. (2007). Acting on persecutory delusions: The importance of safety seeking. *Behaviour Research and Therapy*, *45*(1), 89–99. https://doi.org/10.1016/j.brat.2006.01.014

Freeman, D., Haselton, P., Freeman, J., Spanlang, B., Kishore, S., Albery, E., Denne, M., Brown, P., Slater, M., & Nickless, A. (2018a). Automated psychological therapy using immersive virtual reality for treatment of fear of heights: A single-blind, parallel-group, randomised controlled trial. *The Lancet Psychiatry*, *5*(8), 625–632. https://doi.org/10.1016/S2215-0366(18)30226-8

Freeman, D., Haselton, P., Freeman, J., Spanlang, B., Kishore, S., Albery, E., Denne, M., Brown, P., Slater, M., & Nickless, A. (2018b). Automated psychological therapy using immersive virtual reality for treatment of fear of heights: A single-blind, parallel-group, randomised controlled trial. *The Lancet. Psychiatry*, *5*(8), 625–632. https://doi.org/10.1016/S2215-0366(18)30226-8

Freeman, D., Haselton, P., Freeman, J., Spanlang, B., Kishore, S., Albery, E., Denne, M., Brown, P., Slater, M., & Nickless, A. (2018c). Automated psychological therapy using immersive virtual reality for treatment of fear of heights: A single-blind, parallel-group, randomised controlled trial. *The Lancet Psychiatry*, *5*(8), 625–632. https://doi.org/10.1016/S2215-0366(18)30226-8

Freeman, D., Lambe, S., Galal, U., Yu, L.-M., Kabir, T., Petit, A., Rosebrock, L., Dudley, R., Chapman, K., Morrison, A., O'Regan, E., Murphy, E., Aynsworth, C., Jones, J., Powling, R., Grabey, J., Rovira, A., Freeman, J., Clark, D. M., & Waite, F. (2022). Agoraphobic avoidance in patients with psychosis: Severity and response to automated VR therapy in a secondary analysis of a randomised controlled clinical trial. *Schizophrenia Research*, *250*, 50–59. https://doi.org/10.1016/j.schres.2022.10.008

Freeman, D., Lambe, S., Kabir, T., Petit, A., Rosebrock, L., Yu, L.-M., Dudley, R., Chapman, K., Morrison, A., O'Regan, E., Aynsworth, C., Jones, J., Murphy, E., Powling, R., Galal, U., Grabey, J., Rovira, A., Martin, J., Hollis, C., … West, J. (2022a). Automated virtual reality therapy to treat agoraphobic avoidance and distress in patients with psychosis (gameChange): A multicentre, parallel-group, single-blind, randomised, controlled trial in England with mediation and moderation analyses. *The Lancet Psychiatry*, *9*(5), 375–388. https://doi.org/10.1016/S2215-0366(22)00060-8

Freeman, D., Lambe, S., Kabir, T., Petit, A., Rosebrock, L., Yu, L.-M., Dudley, R., Chapman, K., Morrison, A., O'Regan, E., Aynsworth, C., Jones, J., Murphy, E., Powling, R., Galal, U., Grabey, J., Rovira, A., Martin, J., Hollis, C., … West, J. (2022b). Automated virtual reality therapy to treat agoraphobic avoidance and distress in patients with psychosis (gameChange): A multicentre, parallel-group, single-blind, randomised, controlled trial in England with mediation and moderation analyses. *The Lancet Psychiatry*, *9*(5), 375–388. https://doi.org/10.1016/S2215-0366(22)00060-8

Freeman, D., Lister, R., Waite, F., Galal, U., Yu, L.-M., Lambe, S., Beckley, A., Bold, E., Jenner, L., Diamond, R., Kirkham, M., Twivy, E., Causier, C., Carr, L., Saidel, S., Day, R., Beacco, A., Rovira, A., Ivins, A., … Rosebrock, L. (2023). *Automated virtual reality cognitive therapy versus virtual reality mental relaxation therapy for the treatment of persistent persecutory delusions in patients with psychosis (THRIVE): A parallel-group, single-blind, randomised controlled trial in England with mediation analyses. 10.*

Freeman, D., & Loe, B. S. (2023). Explaining paranoia: Cognitive and social processes in the occurrence of extreme mistrust. *BMJ Ment Health*, *26*(1). https://doi.org/10.1136/bmjment-2023-300880

Freeman, D., Loe, B. S., Kingdon, D., Startup, H., Molodynski, A., Rosebrock, L., Brown, P., Sheaves, B., Waite, F., & Bird, J. C. (2021). The revised Green *et al.* , Paranoid Thoughts Scale (R-GPTS): Psychometric properties, severity ranges, and clinical cut-offs. *Psychological Medicine*, *51*(2), 244–253. https://doi.org/10.1017/S0033291719003155

Freeman, D., Pugh, K., Antley, A., Slater, M., Bebbington, P., Gittins, M., Dunn, G., Kuipers, E., Fowler, D., & Garety, P. (2008). Virtual reality study of paranoid thinking in the general population. *The British Journal of Psychiatry*, *192*(4), 258–263. https://doi.org/10.1192/bjp.bp.107.044677

Freeman, D., Pugh, K., Green, C., Valmaggia, L., Dunn, G., & Garety, P. (2007). A Measure of State Persecutory Ideation for Experimental Studies. *Journal of Nervous & Mental Disease*, *195*(9), 781–784. https://doi.org/10.1097/NMD.0b013e318145a0a9

Freeman, D., Pugh, K., Vorontsova, N., Antley, A., & Slater, M. (2010). Testing the continuum of delusional beliefs: An experimental study using virtual reality. *Journal of Abnormal Psychology*, *119*(1), 83–92. https://doi.org/10.1037/a0017514

Freeman, D., Rosebrock, L., Waite, F., Loe, B. S., Kabir, T., Petit, A., Dudley, R., Chapman, K., Morrison, A., O'Regan, E., Aynsworth, C., Jones, J., Murphy, E., Powling, R., Peel, H., Walker, H., Byrne, R., Freeman, J., Rovira, A., … Lambe, S. (2023). Virtual reality (VR) therapy for patients with psychosis: Satisfaction and side effects. *Psychological Medicine*, *53*(10), 4373–4384. https://doi.org/10.1017/S0033291722001167

Freeman, D., Slater, M., Bebbington, P. E., Garety, P. A., Kuipers, E., Fowler, D., Met, A., Read, C. M., Jordan, J., & Vinayagamoorthy, V. (2003a). Can Virtual Reality be Used to Investigate Persecutory Ideation? *The Journal of Nervous and Mental Disease*, *191*(8), 509–514. https://doi.org/10.1097/01.nmd.0000082212.83842.fe

Freeman, D., Slater, M., Bebbington, P. E., Garety, P. A., Kuipers, E., Fowler, D., Met, A., Read, C. M., Jordan, J., & Vinayagamoorthy, V. (2003b). Can Virtual Reality be Used to Investigate Persecutory Ideation?: *The Journal of Nervous and Mental Disease*, *191*(8), 509–514. https://doi.org/10.1097/01.nmd.0000082212.83842.fe

Frith, C. (2009). Role of facial expressions in social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3453–3458. https://doi.org/10.1098/rstb.2009.0142

Fujihara, Y., Guo, K., & Liu, C. H. (2023). Relationship between types of anxiety and the ability to recognize facial expressions. *Acta Psychologica*, *241*, 104100. https://doi.org/10.1016/j.actpsy.2023.104100

Fusaro, M., Fanti, V., & Chakrabarti, B. (2023). Greater interpersonal distance in adults with autism. *Autism Research*, *16*(10), 2002–2007. https://doi.org/10.1002/aur.3013

Gamberini, L., Chittaro, L., Spagnolli, A., & Carlesso, C. (2015). Psychological response to an emergency in virtual reality: Effects of victim ethnicity and emergency type on helping behavior and navigation. *Computers in Human Behavior*, *48*, 104–113. https://doi.org/10.1016/j.chb.2015.01.040

Geraets, C. N. W., Klein Tuente, S., Lestestuiver, B. P., van Beilen, M., Nijman, S. A., Marsman, J. B. C., & Veling, W. (2021). Virtual reality facial emotion recognition in social environments: An eye-tracking study. *Internet Interventions*, *25*, 100432. https://doi.org/10.1016/j.invent.2021.100432

Giguère, S., Potvin, S., Beaudoin, M., Dellazizzo, L., Giguère, C.-É., Furtos, A., Gilbert, K., Phraxayavong, K., & Dumais, A. (2023). Avatar Intervention for Cannabis Use Disorder in Individuals with Severe Mental Disorders: A Pilot Study. *Journal of Personalized Medicine*, *13*(5), Article 5. https://doi.org/10.3390/jpm13050766

Gillies, M. (2018). Creating Virtual Characters. *Proceedings of the 5th International Conference on Movement and Computing*, 1–8. https://doi.org/10.1145/3212721.3212835

Giron, J., Mladenovic, J., Hanrieder, M., Golland, Y., & Friedman, D. (2023). Physiological correlates of stress induced by virtual humans in a naturalistic virtual reality scenario. *2023 11th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, 1–8. https://doi.org/10.1109/ACIIW59127.2023.10388148

Gonzalez-Franco, M., Egan, Z., Peachey, M., Antley, A., Randhavane, T., Panda, P., Zhang, Y., Wang, C. Y., Reilly, D. F., Peck, T. C., Won, A. S., Steed, A., & Ofek, E. (2020). MoveBox: Democratizing MoCap for the Microsoft Rocketbox Avatar Library. *2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, 91–98. https://doi.org/10.1109/AIVR50618.2020.00026

Gonzalez-Franco, M., Ofek, E., Pan, Y., Antley, A., Steed, A., Spanlang, B., Maselli, A., Banakou, D., Pelechano, N., Orts-Escolano, S., Orvalho, V., Trutoiu, L., Wojcik, M., Sanchez-Vives, M. V., Bailenson, J., Slater, M., & Lanier, J. (2020). The Rocketbox Library and the Utility of Freely Available Rigged Avatars. *Frontiers in Virtual Reality*, *1*. https://doi.org/10.3389/frvir.2020.561558

Gorisse, G., Senel, G., Banakou, D., Beacco, A., Oliva, R., Freeman, D., & Slater, M. (2021). Self-observation of a virtual body-double engaged in social interaction reduces persecutory thoughts. *Scientific Reports*, *11*(1), 23923. https://doi.org/10.1038/s41598-021-03373-x

Green, C. E. L., Freeman, D., Kuipers, E., Bebbington, P., Fowler, D., Dunn, G., & Garety, P. A. (2011). Paranoid Explanations of Experience: A Novel Experimental Study. *Behavioural and Cognitive Psychotherapy*, *39*(1), 21–34. https://doi.org/10.1017/S1352465810000457

Guillén, V., Baños, R. M., & Botella, C. (2018). Users' Opinion About a Virtual Reality System as an Adjunct to Psychological Treatment for Stress-Related Disorders: A Quantitative and Qualitative Mixed-Methods Study. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01038

Guldager, J. D., Kjær, S. L., Grittner, U., & Stock, C. (2022). Efficacy of the Virtual Reality Intervention VR FestLab on Alcohol Refusal Self-Efficacy: A Cluster-Randomized Controlled Trial. *International Journal of Environmental Research and Public Health*, *19*(6), 3293. https://doi.org/10.3390/ijerph19063293

Gupta, K., Zhang, Y., Gunasekaran, T. S., Krishna, N., Pai, Y. S., & Billinghurst, M. (2024a). CAEVR: Biosignals-Driven Context-Aware Empathy in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, 1–11. https://doi.org/10.1109/TVCG.2024.3372130

Gupta, K., Zhang, Y., Gunasekaran, T. S., Krishna, N., Pai, Y. S., & Billinghurst, M. (2024b). CAEVR: Biosignals-Driven Context-Aware Empathy in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, 1–11. https://doi.org/10.1109/TVCG.2024.3372130

Gupta, T., Osborne, K. J., Nadig, A., Haase, C. M., & Mittal, V. A. (2023). Alterations in facial expressions in individuals at risk for psychosis: A facial electromyography approach using emotionally evocative film clips. *Psychological Medicine*, *53*(12), 5829–5838. https://doi.org/10.1017/S0033291722003087

Gürerk, Ö., Bönsch, A., Kittsteiner, T., & Staffeldt, A. (2019). Virtual humans as co-workers: A novel methodology to study peer effects. *Journal of Behavioral and Experimental Economics*, *78*, 17–29. https://doi.org/10.1016/j.socec.2018.11.003

Halbig, A., & Latoschik, M. E. (2021). A Systematic Review of Physiological Measurements, Factors, Methods, and Applications in Virtual Reality. *Frontiers in Virtual Reality*, *2*. https://doi.org/10.3389/frvir.2021.694567

Hale, J., & Hamilton, A. F. D. C. (2016). Testing the relationship between mimicry, trust and rapport in virtual reality conversations. *Scientific Reports*, *6*(1), Article 1. https://doi.org/10.1038/srep35295

Halim, I., Stemmet, L., Hach, S., Porter, R., Liang, H.-N., Vaezipour, A., Henry, J. D., & Baghaei, N. (2023). Individualized Virtual Reality for Increasing Self-Compassion: Evaluation Study. *JMIR Mental Health*, *10*(1), e47617. https://doi.org/10.2196/47617

Han, K., Shin, J., Yoon, S. Y., Jang, D.-P., & Kim, J.-J. (2014). Deficient gaze pattern during virtual multiparty conversation in patients with schizophrenia. *Computers in Biology and Medicine*, *49*, 60–66. https://doi.org/10.1016/j.compbiomed.2014.03.012

Hariton, E., & Locascio, J. J. (2018). Randomised controlled trials—The gold standard for effectiveness research. *BJOG : An International Journal of Obstetrics and Gynaecology*, *125*(13), 1716. https://doi.org/10.1111/1471-0528.15199

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), Article 7825. https://doi.org/10.1038/s41586-020-2649-2

Hartanto, D., Kampmann, I. L., Morina, N., Emmelkamp, P. G. M., Neerincx, M. A., & Brinkman, W.-P. (2014). Controlling Social Stress in Virtual Reality Environments. *PLOS ONE*, *9*(3), e92804. https://doi.org/10.1371/journal.pone.0092804

Hidding, M., Veling, W., Pijnenborg, G. H. M., & van der Stouwe, E. C. D. (2024). A single-session VR intervention addressing self-compassion and self-criticism with and without perspective change: Results of a randomized controlled experiment. *Behaviour Research and Therapy*, *173*, 104466. https://doi.org/10.1016/j.brat.2023.104466

Holmqvist, K., Nyström, M., & Mulvey, F. (2012). Eye tracker data quality: What it is and how to measure it. *Proceedings of the Symposium on Eye Tracking Research and Applications*, 45–52. https://doi.org/10.1145/2168556.2168563

Horigome, T., Yoshida, S., Tanikawa, T., Mimura, M., & Kishimoto, T. (2023). Modification of the therapist's facial expressions using virtual reality technology during the treatment of social anxiety disorder: A case series. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1030050

Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology*, *38*(2), 139–149. https://doi.org/10.1037/0022-0167.38.2.139

Inoue, M., Irino, T., Furuyama, N., & Hanada, R. (2021). Observational and Accelerometer Analysis of Head Movement Patterns in Psychotherapeutic Dialogue. *Sensors*, *21*, 3162. https://doi.org/10.3390/s21093162

Jack, R. E., & Schyns, P. G. (2015). The Human Face as a Dynamic Tool for Social Communication. *Current Biology*, *25*(14), R621–R634. https://doi.org/10.1016/j.cub.2015.05.052

Jain, D., Jain, U., Rao, K., Dhingra, D., Nagrath, P., Varshney, A. R., & Nagrath, P. (2024). A Real-Time Crowd Tracking and Control System Using Deep Learning. In A. E. Hassanien, S. Anand, A. Jaiswal, & P. Kumar (Eds.), *Innovative Computing and Communications* (pp. 587–601). Springer Nature. https://doi.org/10.1007/978-981-97-3817-5_42

Jebb, A. T., Tay, L., Wang, W., & Huang, Q. (2015). Time series analysis for psychological research: Examining and forecasting change. *Frontiers in Psychology*, *6*, 727. https://doi.org/10.3389/fpsyg.2015.00727

Jeong, H. S., Lee, J. H., Kim, H. E., & Kim, J.-J. (2021). Appropriate Number of Treatment Sessions in Virtual Reality-Based Individual Cognitive Behavioral Therapy for Social Anxiety Disorder. *Journal of Clinical Medicine*, *10*(5), 915. https://doi.org/10.3390/jcm10050915

Jerome, L., Feduccia, A. A., Wang, J. B., Hamilton, S., Yazar-Klosinski, B., Emerson, A., Mithoefer, M. C., & Doblin, R. (2020). Long-term follow-up outcomes of MDMA-assisted psychotherapy for treatment of PTSD: A longitudinal pooled analysis of six phase 2 trials. *Psychopharmacology*, *237*(8), 2485–2497. https://doi.org/10.1007/s00213-020-05548-2

Jorjafki, E. M., Sagarin, B. J., & Butail, S. (2018). Drawing power of virtual crowds. *Journal of The Royal Society Interface*, *15*(145), 20180335. https://doi.org/10.1098/rsif.2018.0335

Kaminska, O. K., Magnuski, M., Olszanowski, M., Gola, M., Brzezicka, A., & Winkielman, P. (2020). Ambiguous at the second sight: Mixed facial expressions trigger late electrophysiological responses linked to lower social impressions. *Cognitive, Affective, & Behavioral Neuroscience*, *20*(2), 441–454. https://doi.org/10.3758/s13415-020-00778-5

Kane, H. S., McCall, C., Collins, N. L., & Blascovich, J. (2012). Mere presence is not enough: Responsive support in a virtual world. *Journal of Experimental Social Psychology*, *48*(1), 37–44. https://doi.org/10.1016/j.jesp.2011.07.001

Kang, W., Kim, G., Kim, H., & Lee, S.-H. (2019). The Influence of Anxiety on the Recognition of Facial Emotion Depends on the Emotion Category and Race of the Target Faces. *Experimental Neurobiology*, *28*(2), 261–269. https://doi.org/10.5607/en.2019.28.2.261

Kao, D., Ratan, R., Mousas, C., & Magana, A. J. (2021). The Effects of a Self-Similar Avatar Voice in Educational Games. *Proc. ACM Hum.-Comput. Interact.*, *5*(CHI PLAY), 238:1-238:28. https://doi.org/10.1145/3474665

Kappas, A., Krumhuber, E., & Küster, D. (2013). 6 Facial behavior. In J. A. Hall & M. L. Knapp (Eds.), *Nonverbal Communication* (pp. 131–166). De Gruyter Mouton. https://doi.org/10.1515/9783110238150.131

Kilteni, K., Groten, R., & Slater, M. (2012). The Sense of Embodiment in Virtual Reality. *Presence: Teleoperators and Virtual Environments*, *21*(4), 373–387. https://doi.org/10.1162/PRES_a_00124

Kim, J., Jung, Y. H., Shin, Y.-B., Kim, M.-K., Eom, H., Kim, E., Kim, J., & Kim, J.-J. (2020). Development and Validation of a Virtual Reality-Based Training Program for Promoting Subjective Well-Being. *Psychiatry Investigation*, *17*(12), 1207–1215. https://doi.org/10.30773/pi.2020.0311

Kimmel, S., Jung, F., Matviienko, A., Heuten, W., & Boll, S. (2023). Let's Face It: Influence of Facial Expressions on Social Presence in Collaborative Virtual Reality. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–16. https://doi.org/10.1145/3544548.3580707

Kleinsmith, A., & Bianchi-Berthouze, N. (2013). Affective Body Expression Perception and Recognition: A Survey. *IEEE Transactions on Affective Computing*, *4*(1), 15–33. IEEE Transactions on Affective Computing. https://doi.org/10.1109/T-AFFC.2012.16

Knight, I., West, J., Matthews, E., Kabir, T., Lambe, S., Waite, F., & Freeman, D. (2021). Participatory design to create a VR therapy for psychosis. *Design for Health*, *5*(1), 98–119. https://doi.org/10.1080/24735132.2021.1885889

Kocur, M., Habler, F., Schwind, V., Woźniak, P. W., Wolff, C., & Henze, N. (2021). Physiological and Perceptual Responses to Athletic Avatars while Cycling in Virtual Reality. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–18. https://doi.org/10.1145/3411764.3445160

Kothgassner, O. D., Goreis, A., Glenk, L. M., Kafka, J. X., Beutl, L., Kryspin-Exner, I., Hlavacs, H., Palme, R., & Felnhofer, A. (2021). Virtual and real-life ostracism and its impact on a subsequent acute stressor. *Physiology & Behavior*, *228*, 113205. https://doi.org/10.1016/j.physbeh.2020.113205

Kothgassner, O. D., Goreis, A., Kafka, J. X., Kaufmann, M., Atteneder, K., Beutl, L., Hennig-Fast, K., Hlavacs, H., & Felnhofer, A. (2019). Virtual social support buffers stress response: An experimental comparison of real-life and virtual support prior to a social stressor. *Journal of Behavior Therapy and Experimental Psychiatry*, *63*, 57–65. https://doi.org/10.1016/j.jbtep.2018.11.003

Kroczek, L. O. H., Pfaller, M., Lange, B., Müller, M., & Mühlberger, A. (2020). Interpersonal Distance During Real-Time Social Interaction: Insights From Subjective Experience, Behavior, and Physiology. *Frontiers in Psychiatry*, *11*. https://www.frontiersin.org/articles/10.3389/fpsyt.2020.00561

Krumhuber, E. G., Skora, L. I., Hill, H. C. H., & Lander, K. (2023). The role of facial movements in emotion recognition. *Nature Reviews Psychology*, *2*(5), Article 5. https://doi.org/10.1038/s44159-023-00172-1

Kullmann, P., Menzel, T., Botsch, M., & Latoschik, M. E. (2023). An Evaluation of Other-Avatar Facial Animation Methods for Social VR. *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–7. https://doi.org/10.1145/3544549.3585617

Kwon, J. H., Powell, J., & Chalmers, A. (2013). How level of realism influences anxiety in virtual reality environments for a job interview. *International Journal of Human-Computer Studies*, *71*(10), 978–987. https://doi.org/10.1016/j.ijhcs.2013.07.003

Lambe, S., Knight, I., Kabir, T., West, J., Patel, R., Lister, R., Rosebrock, L., Rovira, A., Garnish, B., Freeman, J., M. Clark, D., Waite, F., & Freeman, D. (2020). Developing an automated VR cognitive treatment for psychosis: gameChange VR therapy. *Journal of Behavioral and Cognitive Therapy*, *30*(1), 33–40. https://doi.org/10.1016/j.jbct.2019.12.001

Låver, J., McAleavey, A., Valaker, I., Castonguay, L. G., & Moltu, C. (2024). Therapists' and patients' experiences of using patients' self-reported data in ongoing psychotherapy processes—A systematic review and meta-analysis of qualitative studies. *Psychotherapy Research*, *34*(3), 293–310. https://doi.org/10.1080/10503307.2023.2222896

Lawson, A. P., & Mayer, R. E. (2022). Does the emotional stance of human and virtual instructors in instructional videos affect learning processes and outcomes? *Contemporary Educational Psychology*, *70*, 102080. https://doi.org/10.1016/j.cedpsych.2022.102080

Lee, H.-S., Hong, S.-J. J., Baxter, T., Scott, J., Shenoy, S., Buck, L., Bodenheimer, B., & Park, S. (2021). Altered Peripersonal Space and the Bodily Self in Schizophrenia: A Virtual Reality Study. *Schizophrenia Bulletin*, *47*(4), 927–937. https://doi.org/10.1093/schbul/sbab024

Lee, J., Lim, Y., Graham, S. J., Kim, G., Wiederhold, B. K., Wiederhold, M. D., Kim, I. Y., & Kim, S. I. (2004). Nicotine craving and cue exposure therapy by using virtual environments. *Cyberpsychology & Behavior : The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society*, *7*(6), 705–713. https://doi.org/10.1089/cpb.2004.7.705

Leonardis, D., Frisoli, A., Barsotti, M., Carrozzino, M., & Bergamasco, M. (2014). Multisensory Feedback Can Enhance Embodiment Within an Enriched Virtual Walking Scenario. *Presence: Teleoperators and Virtual Environments*, *23*(3), 253–266. https://doi.org/10.1162/PRES_a_00190

Liew, T. W., Mat Zin, N. A., Ashaari, N., & Tan, S.-M. (2016). The Effects of a Pedagogical Agent's Smiling Expression on the Learner's Emotions and Motivation in a Virtual Learning Environment. *The International Review of Research in Open and Distributed Learning*, *17*, 248–266. https://doi.org/10.19173/irrodl.v17i5.2350

Lindner, P., Miloff, A., Fagernäs, S., Andersen, J., Sigeman, M., Andersson, G., Furmark, T., & Carlbring, P. (2019). Therapist-led and self-led one-session virtual reality exposure therapy for public speaking anxiety with consumer hardware and software: A randomized controlled trial. *Journal of Anxiety Disorders*, *61*, 45–54. https://doi.org/10.1016/j.janxdis.2018.07.003

Lisetti, C., Amini, R., Yasavur, U., & Rishe, N. (2013). I Can Help You Change! An Empathic Virtual Agent Delivers Behavior Change Health Interventions. *ACM Transactions on Management Information Systems*, *4*(4), 1–28. https://doi.org/10.1145/2544103

Liu, F., Lijn, F. van der, Schurmann, C., Zhu, G., Chakravarty, M. M., Hysi, P. G., Wollstein, A., Lao, O., Bruijne, M. de, Ikram, M. A., Lugt, A. van der, Rivadeneira, F., Uitterlinden, A. G., Hofman, A., Niessen, W. J., Homuth, G., Zubicaray, G. de, McMahon, K. L., Thompson, P. M., … Kayser, M. (2012). A Genome-Wide Association Study Identifies Five Loci Influencing Facial Morphology in Europeans. *PLOS Genetics*, *8*(9), e1002932. https://doi.org/10.1371/journal.pgen.1002932

Llanes-Jurado, J., Gómez-Zaragozá, L., Minissi, M. E., Alcañiz, M., & Marín-Morales, J. (2024). Developing conversational Virtual Humans for social emotion elicitation based on large language models. *Expert Systems with Applications*, *246*, 123261. https://doi.org/10.1016/j.eswa.2024.123261

Llanes-Jurado, J., Marín-Morales, J., Guixeres, J., & Alcañiz, M. (2020). Development and Calibration of an Eye-Tracking Fixation Identification Algorithm for Immersive Virtual Reality. *Sensors*, *20*(17), Article 17. https://doi.org/10.3390/s20174956

Loth, E., Garrido, L., Ahmad, J., Watson, E., Duff, A., & Duchaine, B. (2018). Facial expression recognition as a candidate marker for autism spectrum disorder: How frequent and severe are deficits? *Molecular Autism*, *9*(1), 7. https://doi.org/10.1186/s13229-018-0187-7

Lundin, R. M., Yeap, Y., & Menkes, D. B. (2023). Adverse Effects of Virtual and Augmented Reality Interventions in Psychiatry: Systematic Review. *JMIR Mental Health*, *10*, e43240. https://doi.org/10.2196/43240

Luo, L., Weng, D., Ding, N., Hao, J., & Tu, Z. (2022). The effect of avatar facial expressions on trust building in social virtual reality. *The Visual Computer*. https://doi.org/10.1007/s00371-022-02700-1

Luong, T., & Holz, C. (2022). Characterizing Physiological Responses to Fear, Frustration, and Insight in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, *28*(11), 3917–3927. IEEE Transactions on Visualization and Computer Graphics. https://doi.org/10.1109/TVCG.2022.3203113

Ma, F., & Pan, X. (2022). Visual Fidelity Effects on Expressive Self-avatar in Virtual Reality: First Impressions Matter. *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 57–65. https://doi.org/10.1109/VR51125.2022.00023

Maestri, G. (2006). Facial and Dialogue Animation. In *Digital Character Animation 3* (pp. 196–219). New Riders.

Magnenat-Thalmann, N., & Thalmann, D. (2005). Virtual humans: Thirty years of research, what next? *The Visual Computer*, *21*(12), 997–1015. https://doi.org/10.1007/s00371-005-0363-6

Makowski, D., Pham, T., Lau, Z. J., Brammer, J. C., Lespinasse, F., Pham, H., Schölzel, C., & Chen, S. H. A. (2021). NeuroKit2: A Python toolbox for neurophysiological signal processing. *Behavior Research Methods*, *53*(4), 1689–1696. https://doi.org/10.3758/s13428-020-01516-y

Maples-Keller, J. L., Rauch, S. A. M., Jovanovic, T., Yasinski, C. W., Goodnight, J. M., Sherrill, A., Black, K., Michopoulos, V., Dunlop, B. W., Rothbaum, B. O., & Norrholm, S. D. (2019). Changes in trauma-potentiated startle, skin conductance, and heart rate within prolonged exposure therapy for PTSD in high and low treatment responders. *Journal of Anxiety Disorders*, *68*, 102147. https://doi.org/10.1016/j.janxdis.2019.102147

Marci, C. D., Moran, E. K., & Orr, S. P. (2004). Physiologic Evidence for the Interpersonal Role of Laughter During Psychotherapy. *The Journal of Nervous and Mental Disease*, *192*(10), 689. https://doi.org/10.1097/01.nmd.0000142032.04196.63

Marci, C., & Riess, H. (2005). The clinical relevance of psychophysiology: Support for the psychobiology of empathy and psychodynamic process. *American Journal of Psychotherapy*, *59*(3), 213–226. https://doi.org/10.1176/appi.psychotherapy.2005.59.3.213

Martens, M. A., Antley, A., Freeman, D., Slater, M., Harrison, P. J., & Tunbridge, E. M. (2019). It feels real: Physiological responses to a stressful virtual reality environment and its impact on working memory. *Journal of Psychopharmacology*, *33*(10), 1264–1273. https://doi.org/10.1177/0269881119860156

McCorry, L. K. (2007). Physiology of the Autonomic Nervous System. *American Journal of Pharmaceutical Education*, *71*(4), 78.

McKinney, W. (2010). *Data Structures for Statistical Computing in Python*. 56–61. https://doi.org/10.25080/Majora-92bf1922-00a

Mendoza-Medialdea, M. T., Meschberger-Annweiler, F.-A., Ascione, M., Rueda-Pina, A., Rabarbari, E., Porras-Garcia, B., Ferrer-Garcia, M., & Gutiérrez-Maldonado, J. (2023). Body Dissatisfaction and Body-Related Attentional Bias: Is There a Causal Relationship? *Journal of Clinical Medicine*, *12*(17), Article 17. https://doi.org/10.3390/jcm12175659

Mertens, G., Wagensveld, P., & Engelhard, I. M. (2019). Cue conditioning using a virtual spider discriminates between high and low spider fearful individuals. *Computers in Human Behavior*, *91*, 192–200. https://doi.org/10.1016/j.chb.2018.10.006

Miloff, A., Carlbring, P., Hamilton, W., Andersson, G., Reuterskiöld, L., & Lindner, P. (2020a). Measuring Alliance Toward Embodied Virtual Therapists in the Era of Automated Treatments With the Virtual Therapist Alliance Scale (VTAS): Development and Psychometric Evaluation. *Journal of Medical Internet Research*, *22*(3), e16660. https://doi.org/10.2196/16660

Miloff, A., Carlbring, P., Hamilton, W., Andersson, G., Reuterskiöld, L., & Lindner, P. (2020b). Measuring Alliance Toward Embodied Virtual Therapists in the Era of Automated Treatments With the Virtual Therapist Alliance Scale (VTAS): Development and Psychometric Evaluation. *Journal of Medical Internet Research*, *22*(3), e16660. https://doi.org/10.2196/16660

Mitruț, O., Moldoveanu, A., Petrescu, L., Petrescu, C., & Moldoveanu, F. (2021). A Review of Virtual Therapists in Anxiety and Phobias Alleviating Applications. In J. Y. C. Chen & G. Fragomeni (Eds.), *Virtual, Augmented and Mixed Reality* (pp. 71–79). Springer International Publishing. https://doi.org/10.1007/978-3-030-77599-5_6

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine*, *19*(2), 98–100. IEEE Robotics & Automation Magazine. https://doi.org/10.1109/MRA.2012.2192811

Mountford, V. A., Tchanturia, K., & Valmaggia, L. (2016). 'What Are You Thinking When You Look at Me?' A Pilot Study of the Use of Virtual Reality in Body Image. *Cyberpsychology, Behavior and Social Networking*, *19*(2), 93–99. https://doi.org/10.1089/cyber.2015.0169

Mousas, C., Anastasiou, D., & Spantidi, O. (2018). The effects of appearance and motion of virtual characters on emotional reactivity. *Computers in Human Behavior*, *86*, 99–108. https://doi.org/10.1016/j.chb.2018.04.036

Mühlberger, A., Herrmann, M. J., Wiedemann, G., Ellgring, H., & Pauli, P. (2001). Repeated exposure of flight phobics to flights in virtual reality. *Behaviour Research and Therapy*, *39*(9), 1033–1050. https://doi.org/10.1016/S0005-7967(00)00076-0

Nason, G. P., & Silverman, B. W. (1995). The Stationary Wavelet Transform and some Statistical Applications. In A. Antoniadis & G. Oppenheim (Eds.), *Wavelets and Statistics* (pp. 281–299). Springer. https://doi.org/10.1007/978-1-4612-2544-7_17

Natali, L., Meregalli, V., Rowlands, K., Di Pietro, J., Treasure, J., Collantoni, E., Meneguzzo, P., Tenconi, E., Favaro, A., Fontana, F., Ceccato, E., Sala, A., Valmaggia, L., & Cardi, V. (2024). Virtual food exposure with positive mood induction or social support to reduce food anxiety in anorexia nervosa: A feasibility study. *International Journal of Eating Disorders*, *57*(3), 703–715. https://doi.org/10.1002/eat.24155

Nijman, S. A., Veling, W., Greaves-Lord, K., Vos, M., Zandee, C. E. R., Aan het Rot, M., Geraets, C. N. W., & Pijnenborg, G. H. M. (2020). Dynamic Interactive Social Cognition

Training in Virtual Reality (DiSCoVR) for People With a Psychotic Disorder: Single-Group Feasibility and Acceptability Study. *JMIR Mental Health*, *7*(8). https://doi.org/10.2196/17808

Nowak, K., & Biocca, F. (2003). The Effect of the Agency and Anthropomorphism on Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments. *Presence Teleoperators & Virtual Environments*, *12*, 481–494. https://doi.org/10.1162/105474603322761289

Oh Kruzic, C., Kruzic, D., Herrera, F., & Bailenson, J. (2020). Facial expressions contribute more than body movements to conversational outcomes in avatar-mediated virtual environments. *Scientific Reports*, *10*(1), 20626. https://doi.org/10.1038/s41598-020-76672-4

Oh, S. Y., Bailenson, J., Krämer, N., & Li, B. (2016). Let the Avatar Brighten Your Smile: Effects of Enhancing Facial Expressions in Virtual Environments. *PLOS ONE*, *11*(9), e0161794. https://doi.org/10.1371/journal.pone.0161794

Osugi, T., & Kawahara, J. I. (2018). Effects of Head Nodding and Shaking Motions on Perceptions of Likeability and Approachability. *Perception*, *47*(1), 16–29. https://doi.org/10.1177/0301006617733209

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ*, *372*, n160. https://doi.org/10.1136/bmj.n160

Pan, X., Gillies, M., & Slater, M. (2015). Virtual Character Personality Influences Participant Attitudes and Behavior – An Interview with a Virtual Human Character about Her Social Anxiety. *Frontiers in Robotics and AI*, *2*. https://doi.org/10.3389/frobt.2015.00001

Pan, X., & Hamilton, A. F. de C. (2018a). Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology*, *109*(3), 395–417. https://doi.org/10.1111/bjop.12290

Pan, X., & Hamilton, A. F. de C. (2018b). Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology*, *109*(3), Article 3. https://doi.org/10.1111/bjop.12290

Park, I. H., Kim, J.-J., Ku, J., Jang, H. J., Park, S.-H., Kim, C.-H., Kim, I. Y., & Kim, S. I. (2009). Characteristics of social anxiety from virtual interpersonal interactions in patients with schizophrenia. *Psychiatry*, *72*(1), 79–93. https://doi.org/10.1521/psyc.2009.72.1.79

Parmar, D., Olafsson, S., Utami, D., Murali, P., & Bickmore, T. (2022). Designing empathic virtual agents: Manipulating animation, voice, rendering, and empathy to create persuasive agents. *Autonomous Agents and Multi-Agent Systems*, *36*(1). https://doi.org/10.1007/s10458-021-09539-1

Perpiñá, C., Botella, C., Baños, R., Marco, H., Alcañiz, M., & Quero, S. (1999). Body Image and Virtual Reality in Eating Disorders: Is Exposure to Virtual Reality More Effective than the Classical Body Image Treatment? *CyberPsychology & Behavior*, *2*(2), 149–155. https://doi.org/10.1089/cpb.1999.2.149

Pertaub, D.-P., Slater, M., & Barker, C. (2002). An Experiment on Public Speaking Anxiety in Response to Three Different Types of Virtual Audience. *Presence: Teleoperators and Virtual Environments*, *11*(1), 68–78. https://doi.org/10.1162/105474602317343668

Peterson, S. M., Furuichi, E., & Ferris, D. P. (2018). Effects of virtual reality high heights exposure during beam-walking on physiological stress and cognitive loading. *PLoS ONE*, *13*(7), e0200306. https://doi.org/10.1371/journal.pone.0200306

Ponton, J. L., Yun, H., Andujar, C., & Pelechano, N. (2022). Combining Motion Matching and Orientation Prediction to Animate Avatars for Consumer-Grade VR Devices. *Computer Graphics Forum*, *41*(8), 107–118. https://doi.org/10.1111/cgf.14628

Porras Garcia, B., Ferrer Garcia, M., Olszewska, A., Yilmaz, L., González Ibañez, C., Gracia Blanes, M., Gültekin, G., Serrano Troncoso, E., & Gutiérrez Maldonado, J. (2019). Is This My Own Body? Changing the Perceptual and Affective Body Image Experience among College Students Using a New Virtual Reality Embodiment-Based Technique. *Journal of Clinical Medicine*, *8*(7). https://doi.org/10.3390/jcm8070925

Porras-Garcia, B., Ferrer-Garcia, M., Ghita, A., Moreno, M., López-Jiménez, L., Vallvé-Romeu, A., Serrano-Troncoso, E., & Gutiérrez-Maldonado, J. (2019). The influence of gender and body dissatisfaction on body-related attentional bias: An eye-tracking and virtual reality study. *International Journal of Eating Disorders*, *52*(10), 1181–1190. https://doi.org/10.1002/eat.23136

Porras-Garcia, B., Ferrer-Garcia, M., Yilmaz, L., Sen, Y. O., Olszewska, A., Ghita, A., Serrano-Troncoso, E., Treasure, J., & Gutiérrez-Maldonado, J. (2020). Body-related attentional bias as mediator of the relationship between body mass index and body dissatisfaction. *European Eating Disorders Review : The Journal of the Eating Disorders Association*, *28*(4), 454–464. https://doi.org/10.1002/erv.2730

Porras-Garcia, B., Ghita, A., Moreno, M., Ferrer-Garcia, M., Bertomeu-Panisello, P., Serrano-Troncoso, E., Riva, G., Dakanalis, A., Achotegui-Loizate, J., Talarn-Caparros, A., Andreu-Gracia, A., Treasure, J., Martinez-Mallen, E., Moreno-Perez, E., & Gutierrez-Maldonado, J. (2018). Gender differences in attentional bias after owning a virtual avatar with increased weight. *Annual Review of CyberTherapy and Telemedicine*, *16*(Human Factors Engineering [4010]), 73–79.

Post, R., Haberman, J., Iwaki, L., & Whitney, D. (2012). The Frozen Face Effect: Why Static Photographs May Not Do You Justice. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00022

Pot-Kolder, R. M. C. A., Geraets, C. N. W., Veling, W., Beilen, M. van, Staring, A. B. P., Gijsman, H. J., Delespaul, P. A. E. G., & Gaag, M. van der. (2018a). Virtual-reality-based cognitive behavioural therapy versus waiting list control for paranoid ideation and social avoidance in patients with psychotic disorders: A single-blind randomised controlled trial. *The Lancet Psychiatry*, *5*(3), 217–226. https://doi.org/10.1016/S2215-0366(18)30053-1

Pot-Kolder, R. M. C. A., Geraets, C. N. W., Veling, W., Beilen, M. van, Staring, A. B. P., Gijsman, H. J., Delespaul, P. A. E. G., & Gaag, M. van der. (2018b). Virtual-reality-based cognitive behavioural therapy versus waiting list control for paranoid ideation and social avoidance in patients with psychotic disorders: A single-blind randomised controlled trial. *The Lancet Psychiatry*, *5*(3), 217–226. https://doi.org/10.1016/S2215-0366(18)30053-1

Powers, M. B., Briceno, N. F., Gresham, R., Jouriles, E. N., Emmelkamp, P. M. G., & Smits, J. A. J. (2013). Do conversations with virtual avatars increase feelings of social anxiety? *Journal of Anxiety Disorders*, *27*(4), 398–403. https://doi.org/10.1016/j.janxdis.2013.03.003

Poyatos, F. (2010). Forms and Functions of Nonverbal Communication in the Novel: A New Perspective of the Author-Character-Reader Relationship. In A. Kendon (Ed.), *Nonverbal Communication, Interaction, and Gesture: Selections from SEMIOTICA* (pp. 107–150). De Gruyter Mouton. https://doi.org/10.1515/9783110880021.107

Provenzano, L., Porciello, G., Ciccarone, S., Lenggenhager, B., Tieri, G., Marucci, M., Dazzi, F., Loriedo, C., & Bufalari, I. (2020). Characterizing Body Image Distortion and Bodily Self-Plasticity in Anorexia Nervosa via Visuo-Tactile Stimulation in Virtual Reality. *Journal of Clinical Medicine*, *9*(1), Article 1. https://doi.org/10.3390/jcm9010098

Quintana, P., Nolet, K., Baus, O., & Bouchard, S. (2019). The Effect of Exposure to Fear-Related Body Odorants on Anxiety and Interpersonal Trust Toward a Virtual Character. *Chemical Senses*, *44*(9), 683–692. https://doi.org/10.1093/chemse/bjz063

Raffegeau, T. E., Fawver, B., Young, W. R., Williams, A. M., Lohse, K. R., & Fino, P. C. (2020). The direction of postural threat alters balance control when standing at virtual elevation. *Experimental Brain Research*, *238*(11), 2653–2663. https://doi.org/10.1007/s00221-020-05917-5

Rahman, M. A., Brown, D. J., Mahmud, M., Harris, M., Shopland, N., Heym, N., Sumich, A., Turabee, Z. B., Standen, B., Downes, D., Xing, Y., Thomas, C., Haddick, S., Premkumar, P., Nastase, S., Burton, A., & Lewis, J. (2023). Enhancing biofeedback-driven self-guided virtual reality exposure therapy through arousal detection from multimodal data using machine learning. *Brain Informatics*, *10*(1), 14. https://doi.org/10.1186/s40708-023-00193-9

Raimbaud, P., Jovane, A., Zibrek, K., Pacchierotti, C., Christie, M., Hoyet, L., Pettré, J., & Olivier, A.-H. (2022). The Stare-in-the-Crowd Effect in Virtual Reality. *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 281–290. https://doi.org/10.1109/VR51125.2022.00047

Ranjbartabar, H., Richards, D., Bilgin, A. A., & Kutay, C. (2021). First Impressions Count! The Role of the Human's Emotional State on Rapport Established with an Empathic versus Neutral Virtual Therapist. *IEEE Transactions on Affective Computing*, *12*(3), 788–800. IEEE Transactions on Affective Computing. https://doi.org/10.1109/TAFFC.2019.2899305

Reblin, M., Ellington, L., Latimer, S., Uchino, B. N., Roter, D., & Maxwell, A. (2012). Communication Style Affects Physiological Response in Simulated Cancer Risk Reduction Interactions. *Journal of Applied Biobehavioral Research*, *17*(2), 129–156. https://doi.org/10.1111/j.1751-9861.2012.00080.x

Recio, G., Sommer, W., & Schacht, A. (2011). Electrophysiological correlates of perceiving and evaluating static and dynamic facial emotional expressions. *Brain Research*, *1376*, 66–75. https://doi.org/10.1016/j.brainres.2010.12.041

Reichenberger, J., Pfaller, M., Forster, D., Gerczuk, J., Youssef Shiban, & Mühlberger, A. (2019). Men Scare Me More: Gender Differences in Social Fear Conditioning in Virtual Reality. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.01617

Reichenberger, J., Pfaller, M., & Mühlberger, A. (2020). Gaze Behavior in Social Fear Conditioning: An Eye-Tracking Study in Virtual Reality. *Frontiers in Psychology*, *11*. https://www.frontiersin.org/articles/10.3389/fpsyg.2020.00035

Reichenberger, J., Wechsler, T. F., Diemer, J., Mühlberger, A., & Notzon, S. (2022). Fear, psychophysiological arousal, and cognitions during a virtual social skills training in social anxiety disorder while manipulating gaze duration. *Biological Psychology*, *175*, 108432. https://doi.org/10.1016/j.biopsycho.2022.108432

Ridout, S. J., Spofford, C. M., van't Wout-Frank, M., Philip, N. S., Unger, W. S., Carpenter, L. L., Tyrka, A. R., & Shea, M. T. (2017). Heart Rate Variability Responses to a Standardized Virtual Reality Exposure in Veterans with PTSD. *Current Treatment Options in Psychiatry*, *4*(3), 271–280. https://doi.org/10.1007/s40501-017-0118-9

Riva, G., & Melis, L. (1997). Virtual Reality for the Treatment of Body Image Disturbances. *Studies in Health Technology and Informatics*, 95–112.

Rizzo, A. S., Difede, J., Rothbaum, B. O., Reger, G., Spitalnick, J., Cukor, J., & McLay, R. (2010). Development and early evaluation of the Virtual Iraq/Afghanistan exposure therapy system for combat-related PTSD. *Annals of the New York Academy of Sciences*, *1208*, 114–125. https://doi.org/10.1111/j.1749-6632.2010.05755.x

Rizzo, A. 'Skip', & Shilling, R. (2017). Clinical Virtual Reality tools to advance the prevention, assessment, and treatment of PTSD. *European Journal of Psychotraumatology*, *8*(sup5). https://doi.org/10.1080/20008198.2017.1414560

Robitaille, N., Jackson, P. L., Hébert, L. J., Mercier, C., Bouyer, L. J., Fecteau, S., Richards, C. L., & McFadyen, B. J. (2017). A Virtual Reality avatar interaction (VRai) platform to assess residual executive dysfunction in active military personnel with previous mild traumatic brain injury: Proof of concept. *Disability and Rehabilitation. Assistive Technology*, *12*(7), 758–764. https://doi.org/10.1080/17483107.2016.1229048

Rotenberg, K. J., Sharp, C., & Venta, A. (2023). Trust Beliefs in Significant Others, Interpersonal Stress, and Internalizing Psychopathology of Adolescents with Psychiatric Disorders. *Child Psychiatry & Human Development*, *54*(2), 450–459. https://doi.org/10.1007/s10578-021-01255-x

Rothbaum, B. O., Price, M., Jovanovic, T., Norrholm, S. D., Gerardi, M., Dunlop, B., Davis, M., Bradley, B., Duncan, E. J., Rizzo, A., & Ressler, K. J. (2014). A Randomized, Double-Blind Evaluation of d-Cycloserine or Alprazolam Combined With Virtual Reality Exposure Therapy for Posttraumatic Stress Disorder in Iraq and Afghanistan War Veterans. *American Journal of Psychiatry*, *171*(6), 640–648. https://doi.org/10.1176/appi.ajp.2014.13121625

Rovira, A., & Slater, M. (2022a). Encouraging bystander helping behaviour in a violent incident: A virtual reality study using reinforcement learning. *Scientific Reports*, *12*(1), 3843. https://doi.org/10.1038/s41598-022-07872-3

Rovira, A., & Slater, M. (2022b). Encouraging bystander helping behaviour in a violent incident: A virtual reality study using reinforcement learning. *Scientific Reports*, *12*(1), Article 1. https://doi.org/10.1038/s41598-022-07872-3

Rovira, A., Southern, R., Swapp, D., Campbell, C., Zhang, J. J., Levine, M., & Slater, M. (2021). Bystander Affiliation Influences Intervention Behavior: A Virtual Reality Study. *SAGE Open*, *11*(3), 21582440211040076. https://doi.org/10.1177/21582440211040076

Rovira, A., Swapp, D., Spanlang, B., & Slater, M. (2009). The use of virtual reality in the study of people's responses to violent incidents. *Frontiers in Behavioral Neuroscience*, *3*. https://doi.org/10.3389/neuro.08.059.2009

RStudio Team. (2021). *RStudio: Integrated Development Environment for R* [Computer software]. RStudio, PBC. http://www.rstudio.com/

Rubin, M., Minns, S., Muller, K., Tong, M. H., Hayhoe, M. M., & Telch, M. J. (2020). Avoidance of social threat: Evidence from eye movements during a public speaking challenge using 360°- video. *Behaviour Research and Therapy*, *134*, 103706. https://doi.org/10.1016/j.brat.2020.103706

Rubin, M., Muller, K., Hayhoe, M. M., & Telch, M. J. (2022). Attention guidance augmentation of virtual reality exposure therapy for social anxiety disorder: A pilot randomized controlled trial. *Cognitive Behaviour Therapy*, *51*(5), 371–387. https://doi.org/10.1080/16506073.2022.2053882

Ryan, V., & Griffin, R. (2016). An Investigation Into Anxiety In Virtual Reality Following A Self-Compassion Induction. *Annual Review of CyberTherapy and Telemedicine*, *14*, pp 109-114.

Saberi, M., Bernardet, U., & DiPaola, S. (2014). *An Architecture for Personality-based, Nonverbal Behavior in Affective Virtual Humanoid Character | Elsevier Enhanced Reader*. https://doi.org/10.1016/j.procs.2014.11.104

Sagui-Henson, S. J., Welcome Chamberlain, C. E., Smith, B. J., Li, E. J., Castro Sweet, C., & Altman, M. (2022). Understanding Components of Therapeutic Alliance and Well-Being from Use of a Global Digital Mental Health Benefit During the COVID-19 Pandemic: Longitudinal Observational Study. *Journal of Technology in Behavioral Science*, 1–12. https://doi.org/10.1007/s41347-022-00263-5

Salber, D., & Coutaz, J. (1993). Applying the Wizard of Oz technique to the study of multimodal systems. In L. J. Bass, J. Gornostaev, & C. Unger (Eds.), *Human-Computer Interaction* (pp. 219–230). Springer. https://doi.org/10.1007/3-540-57433-6_51

Sanchez-Vives, M. V., & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience*, *6*(4), 332–339. https://doi.org/10.1038/nrn1651

Savulich, G., Freeman, D., Shergill, S., & Yiend, J. (2015). Interpretation Biases in Paranoia. *Behavior Therapy*, *46*(1), 110–124. https://doi.org/10.1016/j.beth.2014.08.002

Sawyer, S. F. (2009). Analysis of Variance: The Fundamental Concepts. *Journal of Manual & Manipulative Therapy*, *17*(2), 27E-38E. https://doi.org/10.1179/jmt.2009.17.2.27E

Scarpina, F., Serino, S., Keizer, A., Chirico, A., Scacchi, M., Castelnuovo, G., Mauro, A., & Riva, G. (2019). The Effect of a Virtual-Reality Full-Body Illusion on Body Representation in Obesity. *Journal of Clinical Medicine*, *8*(9). https://doi.org/10.3390/jcm8091330

Schmuckler, M. (2001). What Is Ecological Validity? A Dimensional Analysis. *Infancy*, *2*. https://doi.org/10.1207/S15327078IN0204_02

Schroeder, P. A., Gehrer, N. A., Reents, M., Reimer, N., Vagedes, J., & Svaldi, J. (2023). Body Dissatisfaction Directs Avatar Perception: Embodiment and Selective Visual Attention in Body Mass-Modified Self-Avatars. *Cyberpsychology, Behavior, and Social Networking*, *26*(11), 850–860. https://doi.org/10.1089/cyber.2022.0385

Schroeder, R. (2002). Social Interaction in Virtual Environments: Key Issues, Common Themes, and a Framework for Research. In R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments* (pp. 1–18). Springer. https://doi.org/10.1007/978-1-4471-0277-9_1

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments*, *10*(3), 266–281.

Schuetz, I., & Fiehler, K. (2022). Eye tracking in virtual reality: Vive pro eye spatial accuracy, precision, and calibration reliability. *Journal of Eye Movement Research*, *15*(3), Article 3. https://doi.org/10.16910/jemr.15.3.3

Schulte, D. (2008). Patients' outcome expectancies and their impression of suitability as predictors of treatment outcome. *Psychotherapy Research*, *18*(4), 481–494. https://doi.org/10.1080/10503300801932505

Seiferth, N. Y., Pauly, K., Habel, U., Kellermann, T., Jon Shah, N., Ruhrmann, S., Klosterkötter, J., Schneider, F., & Kircher, T. (2008). Increased neural response related to neutral faces in individuals at risk for psychosis. *NeuroImage*, *40*(1), 289–297. https://doi.org/10.1016/j.neuroimage.2007.11.020

Selaskowski, B., Wiebe, A., Kannen, K., Asché, L., Pakos, J., Philipsen, A., & Braun, N. (2024). Clinical adoption of virtual reality in mental health is challenged by lack of high-quality research. *Npj Mental Health Research*, *3*(1), 1–5. https://doi.org/10.1038/s44184-024-00069-8

Seo, S., Kim, E., Mundy, P., Heo, J., & Kim, K. (Kenny). (2019). Joint Attention Virtual Classroom: A Preliminary Study. *Psychiatry Investigation*, *16*(4), 292–299. https://doi.org/10.30773/pi.2019.02.08

Sep, M. S. C., van Osch, M., van Vliet, L. M., Smets, E. M. A., & Bensing, J. M. (2014). The power of clinicians' affective communication: How reassurance about non-abandonment can reduce patients' physiological arousal and increase information recall in bad news consultations. An experimental study using analogue patients. *Patient Education and Counseling*, *95*(1), 45–52. https://doi.org/10.1016/j.pec.2013.12.022

Shadiev, R., & Li, D. (2023). A review study on eye-tracking technology usage in immersive virtual reality learning environments. *Computers & Education*, *196*, 104681. https://doi.org/10.1016/j.compedu.2022.104681

Shaffer, F., & Ginsberg, J. P. (2017). An Overview of Heart Rate Variability Metrics and Norms. *Frontiers in Public Health*, *5*. https://www.frontiersin.org/articles/10.3389/fpubh.2017.00258

Shiban, Y., Reichenberger, J., Neumann, I. D., & Mühlberger, A. (2015). Social conditioning and extinction paradigm: A translational study in virtual reality. *Frontiers in Psychology*, *6*, 400. https://doi.org/10.3389/fpsyg.2015.00400

Shin, Y.-B., Kim, J.-J., Kim, M.-K., Kyeong, S., Jung, Y. H., Eom, H., & Kim, E. (2018a). Development of an effective virtual environment in eliciting craving in adolescents and young adults with internet gaming disorder. *PLOS ONE*, *13*(4), Article 4. https://doi.org/10.1371/journal.pone.0195677

Shin, Y.-B., Kim, J.-J., Kim, M.-K., Kyeong, S., Jung, Y. H., Eom, H., & Kim, E. (2018b). Development of an effective virtual environment in eliciting craving in adolescents and young adults with internet gaming disorder. *PLOS ONE*, *13*(4), e0195677. https://doi.org/10.1371/journal.pone.0195677

Silva, N., Schreck, T., Veas, E., Sabol, V., Eggeling, E., & Fellner, D. W. (2018). Leveraging eye-gaze and time-series features to predict user interests and build a recommendation model for visual analysis. *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*, 1–9. https://doi.org/10.1145/3204493.3204546

Slater, M., Neyret, S., Johnston, T., Iruretagoyena, G., Crespo, M. Á. de la C., Alabèrnia-Segura, M., Spanlang, B., & Feixas, G. (2019). An experimental study of a virtual reality counselling paradigm using embodied self-dialogue. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-46877-3

Slater, M., Pertaub, D.-P., & Steed, A. (1999). Public Speaking in Virtual Reality: Facing an Audience of Avatars. *Computer Graphics and Applications, IEEE*, *19*, 6–9. https://doi.org/10.1109/38.749116

Slater, M., & Wilbur, S. (1997). A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, *6*(6), 603–616. https://doi.org/10.1162/pres.1997.6.6.603

Souto, T., Baptista, A., Tavares, D., Queirós, C., & António, M. (2013). Facial emotional recognition in schizophrenia: Preliminary results of the virtual reality program for facial emotional recognition. *Archives of Clinical Psychiatry (São Paulo)*, *40*(4), 129–134. https://doi.org/10.1590/S0101-60832013000400001

Spanlang, B., Nierula, B., Haffar, M., & Debruille, J. B. (2018). Mimicking Schizophrenia: Reducing P300b by Minimally Fragmenting Healthy Participants' Selves Using Immersive Virtual Reality Embodiment. *Frontiers in Human Neuroscience*, *12*, 504. https://doi.org/10.3389/fnhum.2018.00504

Spanlang, B., Normand, J.-M., Borland, D., Kilteni, K., Giannopoulos, E., Pomés, A., González-Franco, M., Perez-Marcos, D., Arroyo-Palacios, J., Muncunill, X. N., & Slater, M. (2014). How to Build an Embodiment Lab: Achieving Body Representation Illusions in Virtual Reality. *Frontiers in Robotics and AI*, *1*, 9. https://doi.org/10.3389/frobt.2014.00009

Spyridonis, F., Daylamani-Zad, D., & Nightingale, J. (2024). PublicVR: A virtual reality exposure therapy intervention for adults with speech anxiety. *Virtual Reality*, *28*. https://doi.org/10.1007/s10055-024-00998-x

Stallmann, L., Tran, M., Rudrauf, D., Dukes, D., & Samson, A. C. (2023). Simulating Social Emotion Regulation in Virtual Reality: Effect of Virtual Social Support Following Ostracism in a Cyberball Game. *International Journal of Emerging Technologies in Learning (iJET)*, *18*(07), 4–27. https://doi.org/10.3991/ijet.v18i07.29419

Stein, J.-P., & Ohler, P. (2017). Venturing into the uncanny valley of mind—The influence of mind attribution on the acceptance of human-like characters in a virtual reality setting. *Cognition*, *160*, 43–50. https://doi.org/10.1016/j.cognition.2016.12.010

Steinman, S. A., & Teachman, B. A. (2011). Cognitive Processing and Acrophobia: Validating the Heights Interpretation Questionnaire. *Journal of Anxiety Disorders*, *25*(7), 896–902. https://doi.org/10.1016/j.janxdis.2011.05.001

Szekely, R., Mason, O., Frohlich, D., & Barley, E. (2024). 'It's not everybody's snapshot. It's just an insight into that world': A qualitative study of multiple perspectives towards understanding the mental health experience and addressing stigma in healthcare students through virtual reality. *DIGITAL HEALTH*, *10*, 20552076231223801. https://doi.org/10.1177/20552076231223801

Takac, M., Collett, J., Blom, K. J., Conduit, R., Rehm, I., & De Foe, A. (2019). Public speaking anxiety decreases within repeated virtual reality training sessions. *PloS One*, *14*(5), e0216288. https://doi.org/10.1371/journal.pone.0216288

Takemoto, A., Aispuriete, I., Niedra, L., & Dreimane, L. F. (2023). Differentiating depression using facial expressions in a virtual avatar communication system. *Frontiers in Digital Health*, *5*. https://doi.org/10.3389/fdgth.2023.1080023

Talaat, F. M. (2023). Real-time facial emotion recognition system among children with autism based on deep learning and IoT. *Neural Computing and Applications*, *35*(17), 12717–12728. https://doi.org/10.1007/s00521-023-08372-9

Tarnanas, loannis, Tsoukalas, loannis, & Stogiannidou, A. (2003). Virtual reality as a psychosocial coping environment. *Annual Review of CyberTherapy and Telemedicine*, *1*(Bowers, D., Meichenbaum, D. (1994). The unconscious reconsidered. New York: John Willey.), 77–81.

ter Stal, S., Jongbloed, G., & Tabak, M. (2021). Embodied Conversational Agents in eHealth: How Facial and Textual Expressions of Positive and Neutral Emotions Influence Perceptions of Mutual Understanding. *Interacting with Computers*, *33*(2), 167–176. https://doi.org/10.1093/iwc/iwab019

Thalmann, D. (2000). *Challenges for the Research in Virtual Humans*. 9.

Thompson-Hollands, J., Bentley, K. H., Gallagher, M. W., Boswell, J. F., & Barlow, D. H. (2014a). Credibility and Outcome Expectancy in the Unified Protocol: Relationship to Outcomes. *Journal of Experimental Psychopathology*, *5*(1), 72–82. https://doi.org/10.5127/jep.033712

Thompson-Hollands, J., Bentley, K. H., Gallagher, M. W., Boswell, J. F., & Barlow, D. H. (2014b). Credibility and Outcome Expectancy in the Unified Protocol: Relationship to Outcomes. *Journal of Experimental Psychopathology*, *5*(1), 72–82. https://doi.org/10.5127/jep.033712

Thompson-Lake, D. G. Y., Cooper, K. N., Mahoney, J. J., Bordnick, P. S., Salas, R., Kosten, T. R., Dani, J. A., & De La Garza, R. (2015). Withdrawal Symptoms and Nicotine Dependence Severity Predict Virtual Reality Craving in Cigarette-Deprived Smokers. *Nicotine & Tobacco Research*, *17*(7), 796–802. https://doi.org/10.1093/ntr/ntu245

Tinwell, A., Grimshaw, M., Nabi, D. A., & Williams, A. (2011). Facial expression of emotion and perception of the Uncanny Valley in virtual characters. *Computers in Human Behavior*, *27*(2), 741–749. https://doi.org/10.1016/j.chb.2010.10.018

Treal, T., Jackson, P. L., Jeuvrey, J., Vignais, N., & Meugnot, A. (2021). Natural human postural oscillations enhance the empathic response to a facial pain expression in a virtual character. *Scientific Reports*, *11*(1), 12493. https://doi.org/10.1038/s41598-021-91710-5

Trivedi, H., & Mousas, C. (2023). Human-virtual crowd interaction: Towards understanding the effects of crowd avoidance proximity in an immersive virtual environment. *Computer Animation and Virtual Worlds*, *34*(3–4), e2169. https://doi.org/10.1002/cav.2169

Trotta, A., Kang, J., Stahl, D., & Yiend, J. (2021). Interpretation Bias in Paranoia: A Systematic Review and Meta-Analysis. *Clinical Psychological Science*, *9*(1), 3–23. https://doi.org/10.1177/2167702620951552

Tsiourti, C., Weiss, A., Wac, K., & Vincze, M. (2019). Multimodal Integration of Emotional Signals from Voice, Body, and Context: Effects of (In)Congruence on Emotion Recognition and Attitudes Towards Robots. *International Journal of Social Robotics*, *11*(4), 555–573. https://doi.org/10.1007/s12369-019-00524-z

Tsoli, A., Mahmood, N., & Black, M. J. (2014). Breathing life into shape: Capturing, modeling and animating 3D human breathing. *ACM Trans. Graph.*, *33*(4), 52:1-52:11. https://doi.org/10.1145/2601097.2601225

Ulicny, B., & Thalmann, D. (2001). Crowd simulation for interactive virtual environments and VR training systems. In N. Magnenat-Thalmann & D. Thalmann (Eds.), *Computer Animation and Simulation 2001* (pp. 163–170). Springer Vienna. https://doi.org/10.1007/978-3-7091-6240-8_15

Urbaniak, G., & Plous, S. (2013). *Research randomizer (version 4.0)[computer software]. 2013* [Computer software].

Vahle, N., & Tomasik, M. J. (2022). The Embodiment of an Older Avatar in a Virtual Reality Setting Impacts the Social Motivation of Young Adults. *Experimental Aging Research*, *48*(2), 164–176. https://doi.org/10.1080/0361073X.2021.1943793

Valmaggia, L. R., Day, F., & Rus-Calafell, M. (2016). Using virtual reality to investigate psychological processes and mechanisms associated with the onset and maintenance of psychosis: A systematic review. *Social Psychiatry and Psychiatric Epidemiology*, *51*(7), 921–936. https://doi.org/10.1007/s00127-016-1245-0

Valmaggia, L. R., Freeman, D., Green, C., Garety, P., Swapp, D., Antley, A., Prescott, C., Fowler, D., Kuipers, E., Bebbington, P., Slater, M., Broome, M., & McGuire, P. K. (2007). Virtual reality and paranoid ideations in people with an 'at-risk mental state' for psychosis. *The British Journal of Psychiatry*, *191*(S51), s63–s68. https://doi.org/10.1192/bjp.191.51.s63

Van Gelder, J.-L., Cornet, L. J. M., Zwalua, N. P., Mertens, E. C. A., & van der Schalk, J. (2022). Interaction with the future self in virtual reality reduces self-defeating behavior in a sample of convicted offenders. *Scientific Reports*, *12*(1), Article 1. https://doi.org/10.1038/s41598-022-06305-5

Veldmeijer, L., Terlouw, G., Van Os, J., Van Dijk, O., Van 't Veer, J., & Boonstra, N. (2023). The Involvement of Service Users and People With Lived Experience in Mental Health Care Innovation Through Design: Systematic Review. *JMIR Mental Health*, *10*, e46590. https://doi.org/10.2196/46590

Venning, A., Herd, M. C., Oswald, T. K., Razmi, S., Glover, F., Hawke, T., Quartermain, V., & Redpath, P. (2021). Exploring the acceptability of a digital mental health platform incorporating a virtual coach: The good, the bad, and the opportunities. *Health Informatics Journal*, *27*(1), 1460458221994873. https://doi.org/10.1177/1460458221994873

Vinayagamoorthy, V., Gillies, M., Steed, A., Tanguy, E., Pan, X., Loscos, C., & Slater, M. (2006). *Building Expression into Virtual Characters*. 41.

Volonte, M., Hsu, Y.-C., Liu, K.-Y., Mazer, J. P., Wong, S.-K., & Babu, S. V. (2020). Effects of Interacting with a Crowd of Emotional Virtual Humans on Users' Affective and Non-Verbal Behaviors. *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 293–302. https://doi.org/10.1109/VR46266.2020.00049

Volonte, M., Ofek, E., Jakubzak, K., Bruner, S., & Gonzalez-Franco, M. (2022). HeadBox: A Facial Blendshape Animation Toolkit for the Microsoft Rocketbox Library. *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 39–42. https://doi.org/10.1109/VRW55335.2022.00015

Voutilainen, L., Henttonen, P., Kahri, M., Ravaja, N., Sams, M., & Peräkylä, A. (2018). Empathy, Challenge, and Psychophysiological Activation in Therapist–Client Interaction. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00530

Wang, Z., Giannopoulos, E., Slater, M., Peer, A., & Buss, M. (2011). Handshake: Realistic Human-Robot Interaction in Haptic Enhanced Virtual Reality. *Presence: Teleoperators and Virtual Environments*, *20*(4), 371–392. https://doi.org/10.1162/PRES_a_00061

Wei, S., Freeman, D., & Rovira, A. (2023). A randomised controlled test of emotional attributes of a virtual coach within a virtual reality (VR) mental health treatment. *Scientific Reports*, *13*(1), Article 1. https://doi.org/10.1038/s41598-023-38499-7

Wienrich, C., & Latoschik, M. E. (2021). eXtended Artificial Intelligence: New Prospects of Human-AI Interaction Research. *Frontiers in Virtual Reality*, *2*. https://doi.org/10.3389/frvir.2021.686783

Wieser, M., Pauli, P., Grosseibl, M., Molzow, I., & Mühlberger, A. (2010a). Virtual Social Interactions in Social Anxiety-The Impact of Sex, Gaze, and Interpersonal Distance. *Cyberpsychology, Behavior and Social Networking*, *13*, 547–554. https://doi.org/10.1089/cyber.2009.0432

Wieser, M., Pauli, P., Grosseibl, M., Molzow, I., & Mühlberger, A. (2010b). Virtual Social Interactions in Social Anxiety-The Impact of Sex, Gaze, and Interpersonal Distance. *Cyberpsychology, Behavior and Social Networking*, *13*, 547–554. https://doi.org/10.1089/cyber.2009.0432

Williams, J. G. L., Jones, J. R., Workhoven, M. N., & Williams, B. (1975). The Psychological Control of Preoperative Anxiety. *Psychophysiology*, *12*(1), 50–54. https://doi.org/10.1111/j.1469-8986.1975.tb03059.x

Witmer, B., & Singer, M. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, *7*(3), 225–240.

Wobbrock, J. O., Findlater, L., Gergle, D., & Higgins, J. J. (2011). The aligned rank transform for nonparametric factorial analyses using only anova procedures. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 143–146. https://doi.org/10.1145/1978942.1978963

Wolf, E., Merdan, N., Dölinger, N., Mal, D., Wienrich, C., Botsch, M., & Latoschik, M. E. (2021). The Embodiment of Photorealistic Avatars Influences Female Body Weight

Perception in Virtual Reality. *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, 65–74. https://doi.org/10.1109/VR50410.2021.00027

Wonghempoom, A., Chuangchai, W., & Selanon, P. (2023). Physiological Stress Responses to Fear and Anxiety in a Height Change Experiment among Non-Labor Teenagers. *Safety*, *9*(2), Article 2. https://doi.org/10.3390/safety9020019

Wu, J., Sun, Y., Zhang, G., Zhou, Z., & Ren, Z. (2021). Virtual Reality-Assisted Cognitive Behavioral Therapy for Anxiety Disorders: A Systematic Review and Meta-Analysis. *Frontiers in Psychiatry*, *12*. https://doi.org/10.3389/fpsyt.2021.575094

Wu, Y., Wang, Y., Jung, S., Hoermann, S., & Lindeman, R. W. (2021). Using a Fully Expressive Avatar to Collaborate in Virtual Reality: Evaluation of Task Performance, Presence, and Attraction. *Frontiers in Virtual Reality*, *2*. https://www.frontiersin.org/articles/10.3389/frvir.2021.641296

Wuehr, M., Breitkopf, K., Decker, J., Ibarra, G., Huppert, D., & Brandt, T. (2019). Fear of heights in virtual reality saturates 20 to 40 m above ground. *Journal of Neurology*, *266*(Suppl 1), 80–87. https://doi.org/10.1007/s00415-019-09370-5

Yamamoto, K., & Nakao, T. (2022). Top-Down Influence Leads to a Reduced Sense of Body Ownership in Individuals With Depersonalization Tendencies: A Focus on Full Body Illusion. *Frontiers in Psychology*, *13*. https://www.frontiersin.org/articles/10.3389/fpsyg.2022.790960

Zell, E., Zibrek, K., & McDonnell, R. (2019). Perception of virtual characters. *ACM SIGGRAPH 2019 Courses*, 1–17. https://doi.org/10.1145/3305366.3328101

Zibrek, K., Kokkinara, E., & Mcdonnell, R. (2018). The Effect of Realistic Appearance of Virtual Characters in Immersive Environments—Does the Character's Personality Play a Role? *IEEE Transactions on Visualization and Computer Graphics*, *24*(4), 1681–1690. IEEE Transactions on Visualization and Computer Graphics. https://doi.org/10.1109/TVCG.2018.2794638

5 A Randomised Controlled Test In VR of The Effects on Paranoid Thoughts of Virtual Human's Facial Animation And Expression

ABSTRACT

Background: VR is used in the study and treatment of paranoia. This is based on the finding that people who mistakenly perceive hostile intent from other people also perceive similar threat from virtual characters. However, very few studies have explored the specific characteristics of virtual characters that may influence such interpretations.

Aim: In this chapter, we investigate how the animation of virtual humans (VHs) and their expression may affect paranoia.

Method: The study was designed as a two-by-two factor, between-group, randomised experiment. The experimental variables were facial animation (static or animated) and expression (neutral or positive). 122 individuals with elevated paranoia rated their perceptions of VHs, that varied in facial animation and expression. Participants' eye-tracking data were also collected throughout the VR lift ride.

Results: Both facial animation (group difference=102.328 [51.783, 152.872], $p<0.001$, η_p^2 =0.125) and positive expressions (group difference=53.016 [0.054, 105.979], $p=0.049$, η_p^2=0.033) led to less triggering of paranoid thoughts about the VHs. Facial animation (group difference=2.442 [-4.161, -0.724], $p=0.006$, η_p^2 =0.063) but not positive expressions (group difference=0.344 [-1.429, 2.110], $p=0.681$, η_p^2 =0.001) significantly increased the likelihood of neutral thoughts about the characters.

Conclusion: The results show that both VHs' facial animation and expression affected their appraisal by people vulnerable to paranoia. These details require careful consideration in the design of VR scenarios for mental health applications depending on the purpose of their use.

This chapter is adapted from the papers:

Wei, S., Freeman, D., Harris, V., & Rovira, A. (2024). A randomised controlled test in virtual reality of the effects on paranoid thoughts of virtual humans' facial animation and expression. *Sci Rep* **14**, 17102. https://doi.org/10.1038/s41598-024-67534-4

Wei, S., Bloemers, D., & Rovira, A. (2023). A Preliminary Study of the Eye Tracker in the Meta Quest Pro. Proceedings of the 2023 ACM International Conference on Interactive Media Experiences, 216–221. https://doi.org/10.1145/3573381.3596467

5.1 INTRODUCTION

Paranoia – perceiving hostile intent where there is none – is prevalent in the general population. People with paranoia are more likely to interpret neutral events or comments negatively or react defensively around them (Garety et al., 2001). Many individuals occasionally experience paranoid thoughts, while a smaller number of people experienced them frequently (Freeman, Garety, Bebbington, Smith, et al., 2005). A recent survey of a representative group of ten thousand UK adults indicated that approximately one in six people wanted help to be more trusting of other people (Freeman & Loe, 2023). VR has been used to both study (Freeman et al., 2003, 2008; Freeman, Garety, Bebbington, Slater, et al., 2005; Gorisse et al., 2021; Valmaggia et al., 2007) and treat (Freeman et al., 2016, 2022, 2023; Pot-Kolder et al., 2018) paranoia. Freeman et al (2003) pioneered the use of VR to assess and understand paranoia by examining people's appraisals of neutral VHs. The conclusion was that if the characters are neutral but hostile intent is perceived, then this is clear evidence of paranoid thinking. Studies have shown that higher levels of paranoia in daily life are associated with experiencing higher levels of paranoia about virtual characters (Freeman et al., 2010). Qualitative findings indicate that close observations of the VHs may contribute to the occurrence of paranoid interpretations. One participant in the study from Fornells-Ambrojo et al. described: "I was just looking around, looking at people, just observing them..." (2015). This paper reports, for the first time, on the detailed characteristics of VHs that may affect their appraisal by people vulnerable to paranoia.

Previous VR studies outside the topic of paranoia have shown that facial expressions and animations of VHs can significantly influence people's behavioural and psychological responses. For example, Geraets et al. (2021) suggested that facial emotion cues are beneficial for non-clinical populations to accurately identify the emotions of virtual characters, with the recognition accuracy in VR comparable to that in photographs and videos. Additionally, they stated that participants tended to focus more on the eye and nose areas than the rest parts of face when interpreting the emotions of the VHs. Bönsch et al. (2020) looked at how the emotional expressions of VHs affect personal space preference, by examining responses to approaches by virtual men exhibiting happy, angry, or neutral expressions on static faces. The study showed that participants maintained larger distances from a virtual man with an angry face compared to a happy or neutral face. Kimmel et al. (2023) found that integrating facial animations, such as mouth and eye movements, not only enhanced the social presence in the scenario felt by participants but also made participants feel that the VHs had a better understanding of their emotions and attitudes.

Notably, recent investigations have shown the importance of faces in building trust towards VHs (Choudhary et al., 2023; Luo et al., 2022; Wei et al., 2023). Luo et al. (2022) suggested that, compared to neutral or negative facial expressions, positive facial expressions promoted trust and willingness to cooperate in a VR game. Choudhary et al. (2023) explored the impact of conflicting facial and vocal emotional expressions. They found that VHs with happy faces were perceived as happier and more trustworthy than VHs with unhappy faces. Although appraisals of trust became less predictable for mismatched expressions (e.g., a happy face with unhappy voice), facial expressions had a stronger impact than vocal tone. Chapter 3 concluded that adding positive facial expressions in a virtual coach in a VR phobia treatment significantly improved therapeutic alliance with the coach (Wei et al., 2023). Furthermore, animated faces of VR characters have been found to be perceived as more natural and believable than static faces in VR social experiences (Kimmel et al., 2023; Kullmann et al., 2023).

Despite the evidence of VH characteristics influencing user perceptions, the impact of VH facial details on paranoia remains untested. In this chapter, we focused on two key elements – facial animation and expressions – within the context of virtual crowds as part of the VR stimuli. These two aspects of facial features were chosen as they are fundamental elements in modelling and animating VH faces. Previous research indicated that facial animation can make VHs appear more empathetic and less strange (Parmar et al., 2022), while positive facial expressions can promote trust (Luo et al., 2022). Therefore, our primary hypothesis was that using facial animation or positive facial expression would reduce the likelihood of paranoia appraisals.

An eye tracker enables the capture of eye gaze data to understand where a person is looking at (Geraets et al., 2021; Shadiev & Li, 2023). In VR environments, where individuals can freely explore their surroundings, eye-gaze patterns could reveal attentional biases, such as tendencies to fixate on or avoid specific areas, which may reflect underlying cognitive or emotional processes (Günther et al., 2021). For individuals with paranoia, these patterns could indicate heightened sensitivity to perceived threats or avoidance of distressing stimuli. Therefore, we examined participants' eye gaze behaviour to explore how individuals vulnerable to paranoia visually engage with virtual characters, providing behavioural insights into their mental state.

5.2 METHODS

5.2.1 Experimental Design

We used a two-by-two factorial, between-groups, randomised design to examine two aspects of facial programming: animation (static or animated) and expression (neutral or positive). Participants were randomised to one of the four experimental conditions: VHs with either (1) static neutral, (2) animated neutral, (3) static positive, or (4) animated positive faces. In all experimental conditions, the VHs had the same body animation and regular eye binks. The study was single-blind – participants were unaware of the study hypotheses and that they were randomized into one of the different versions before the VR experience. The randomization was conducted by an independent researcher using Research Randomizer (Urbaniak & Plous, 2013).

Figure 5.1. VR lift scenario. (a) VR lift environment: view of the lift when the door opened with five VHs standing inside. (b) Factorial design with VH faces varied in animation (static or animated) and expressions (neutral or positive).

5.2.2 Apparatus and VR Scenario

We used a Windows 10 computer (Intel i7-8700K, Nvidia GeForce GTX 1080Ti, 32 GB RAM) to run the VR scenario and render it on a Meta Quest Pro (Meta, 2022) via Meta Air Link. The headset has a resolution of 1832*1920 pixels per eye and a field of view of 106° (horizontal) × 96° (vertical). The refresh rate was set at 90 Hz refresh rate.

The VR software was developed in Unity 2021.3.15 with Oculus' Movement SDK V1.3.2. In the virtual lift experience, each participant began at the lobby waiting for the lift. The lift door opened automatically upon arrival, and participants were instructed to step in the lift and stay in it until they reach the top floor labelled as "sky lounge". Inside the lift, there were five people (three men and two women of different ages and ethnicities), as shown in Figure 5.1-a. The ride lasted three minutes, and the

VR scenario ended when the lift reached the destination floor and the doors opened. A video of the VR playthrough experience can be found via the provided link[12].

Similar to the facial animation process described in Chapter 3, the facial animations were recorded using an iPhone 11 depth camera and refined using Iclone7[13] with the LiveFace[14] plugin. The VHs were programmed to occasionally make eye contact with participants by looking towards and away from them. The customised gaze duration was randomly set between 6 and 10 seconds, with a delayed start time ranging from 2 to 5 seconds, to simulate the natural eye contact of passengers in a lift. Each VH made eye contact with participants for roughly 25% of the total ride duration.

5.2.3 Accuracy test of the Meta Quest Pro Eye Tracker

We used the eye tracker integrated in the Meta Quest Pro to record eye gaze data. A set of pre-defined regions of interest (ROIs) was established in the VR lift ride environment to track and count how often participants looked at each ROI. The ROIs included the faces of all five VHs and other specific areas where participants might look to avoid eye contact – the floor, the display panel showing the current floor, and the door. To ensure accurate detection of eye fixation on any ROI, we adjusted the software to minimise false negatives by setting the confidence threshold to 99% within the Movement SDK's eye-tracking settings.

At the time of designing and developing the VR software for this study, no publicly available information or manufacturer references regarding the accuracy of the Meta Quest Pro's eye tracker were available. Therefore, we conducted user testing to estimate its hardware capabilities and inform the VR scenario design, ensuring reliable eye gaze data. Following standard practice, we aimed to measure the spatial accuracy

12 https://youtu.be/Rj1aZZrg77Q
13 https://www.reallusion.com/iclone/
14 https://mocap.reallusion.com/iclone-motion-live-mocap/iphone-live-face.html

and precision under both head-fixed and head-free conditions. Spatial accuracy refers to the average angular offsets between the measured fixation positions and target positions, while spatial precision measures the stability of individual measured gaze samples over time for a fixation target (Holmqvist et al., 2012).

A 3D spherical object (Figure 5.2-a) was designed to place the target points so that all the targets have the same distance from the VR camera. A set of 13 targets were placed in a visual field spanning ±15° horizontally and vertically° with a 5° interval (Figure 5.2-b). Each target was shown as a red dot measuring around 0.7° at a 1-meter viewing distance on a non-reflective dark environmental background, which helped reduce visual distraction. For the head-free condition, the target points were fixed in the world coordinate system, and volunteers could rotate their heads freely to position the targets at the centre of their visual fields. For the head-restrained condition, the target points were attached to the camera and headset positional tracking was disabled so that the targets remained at a fixed position relative to the people's heads.

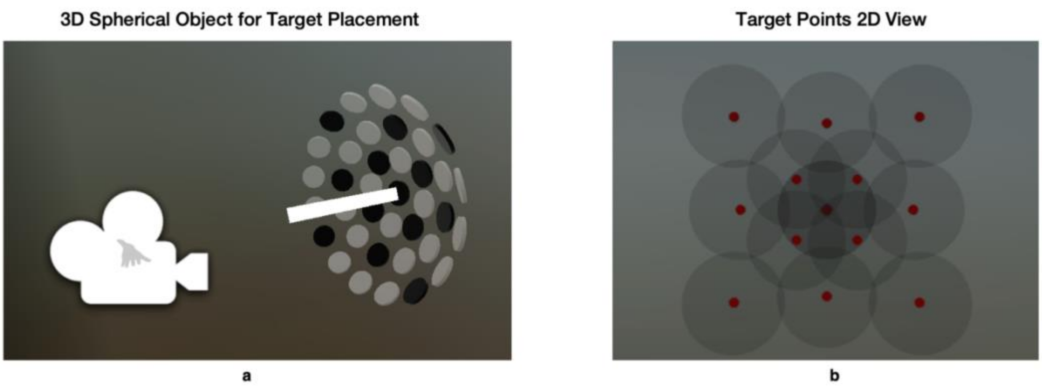


Figure 5.2. Eye tracking targets. (a) 3D spherical object to place the target points, all of which have the same distance with the camera; (b) The 13 targets points in 2D view (in a visual field spanning ±15°).

Twelve volunteers performed a visual focus task on these 13 targets under both head-free and head-restrained (head position restrained by a chinrest) conditions in a within-subjects design. Each volunteer experienced the VR eye-tracking test in both conditions with a counterbalanced order. The results (see Figure 5.3) were average accuracy = 1.652°, precision = 0.699° (standard deviation) and 0.849° (root mean square) for the head-free condition; accuracy = 2.162°, precision = 0.673° (standard deviation) and 0.772° (root mean square) for the head-strained condition. We concluded that the accuracy of the data from the Meta Quest Pro eye tracker is comparable to other existing AR/VR headsets like HTC Vive Pro Eye (mean accuracy of 1.08°) (Schuetz & Fiehler, 2022) and HoloLens 2 (a nominal spatial accuracy of 1.5°) (Aziz & Komogortsev, 2022).

Considering our design of life-size VH faces (width range: 0.13-0.16m) (Liu et al., 2012), we calculated that ROIs needed to be at a distance no larger than 3.5 meters using the visual angle formula (1) (where the visual angle represents the spatial accuracy of the eye tracker in this context), to ensure reliable eye tracking during the VR experience.

$$\text{Object Distance} = \frac{\text{ObjectSize}}{2 \cdot \tan\left(\frac{\text{VisualAngle}}{2}\right)} \qquad (1)$$

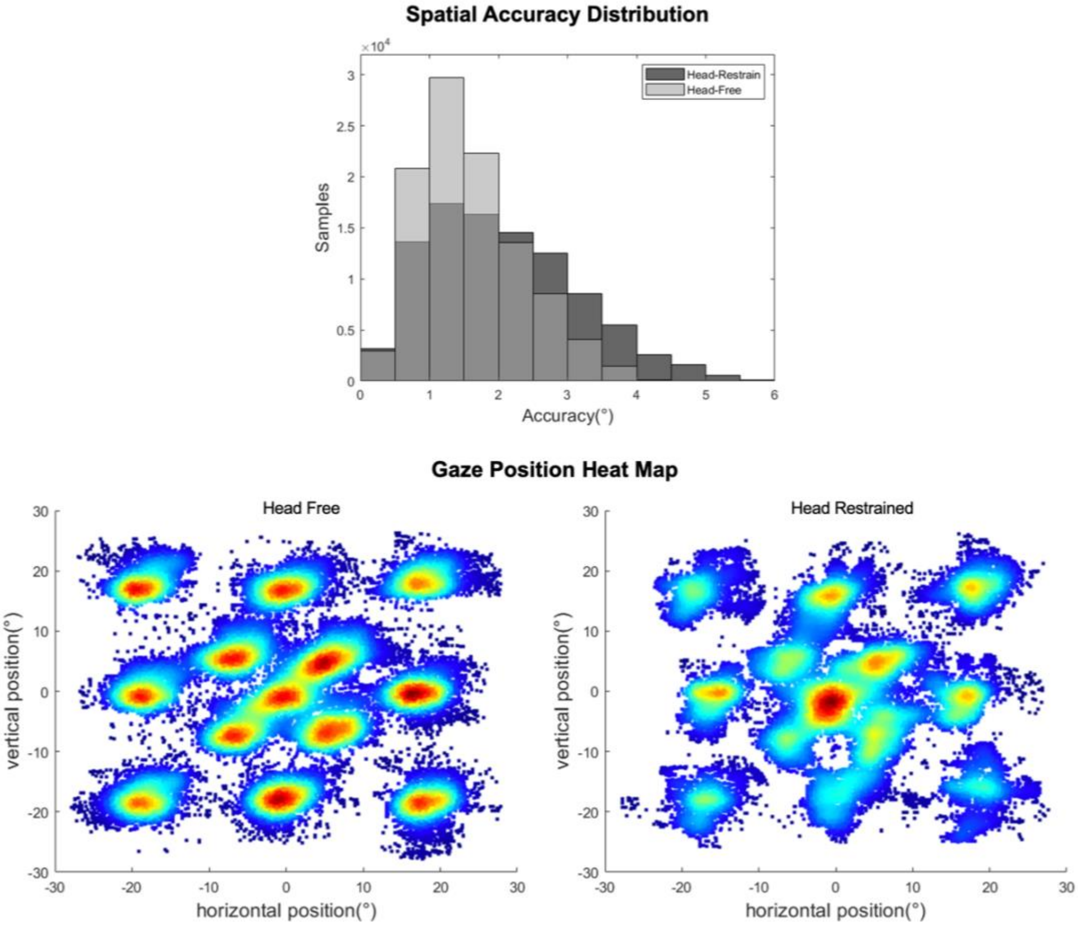


Figure 5.3. Eye tracking accuracy result. Spatial accuracy distribution gaze position heatmap for the 13 targets.

5.2.4 Measures

Baseline Paranoia (screening survey)**.** During screening participants completed The Revised Green et al., Paranoid Thoughts Scale (R-GPTS) (Freeman et al., 2021). We used Part B to assess ideas of persecution. There are 10 items such as "I was sure someone wanted to hurt me" and "People have been hostile towards me on purpose". All items are scored from 0 (Not at all) to 4 (Totally), with total scores ranging from 0 to 40 (Cronbach's α = 0.86 in the current study, N = 122). Higher scores reflect higher levels of paranoia in real life.

Experience with VR (screening survey). During the screening participants also rated a single item "I am very experienced with virtual reality (VR)." to indicate their experience with VR. It is a 5-point scale, where 1 indicates "never tried VR" and 5 indicates "very experienced".

Paranoid Thoughts Visual Analogue Scales (VAS) (Freeman et al., 2015). This was the primary paranoia outcome measure to assess participants' appraisals of the VHs in the virtual lift. After the VR experience, participants rated six visual analogue scales regarding VHs ("Right now I feel suspicious of the people in the lift", "Right now I feel that people in the lift wanted to harm me", "Right now I feel like the people in the lift wanted to upset me", "Right now I feel like the people in the lift were against me", "Right now I am thinking that the people in the lift were trying to persecute me" and "Right now I feel like the people in the lift were hostile towards me"). Participants marked each item on a standard 10cm long VAS on a paper from 0 (not at all anxious) to 100 (extremely), with total scores ranging from 0 to 600 (Cronbach's $\alpha = 0.935$ in the current study, N = 122). Higher scores indicate higher levels of paranoia about the VHs.

State Social Paranoia Scale (SSPS) (Freeman, Pugh, et al., 2007). This provided a further assessment of paranoid thoughts about the VHs and also neutral and positive appraisals. In the scale, each item is scored from 1 (Do not agree) to 5 (Totally agree). There are 10 items measuring paranoid thoughts (SSPSPersecutory) (range: 10, 50, Cronbach's $\alpha = 0.950$ in the current study, N = 122), and 5 items each measuring neutral views (SSPSNeutral) (range: 5, 25, Cronbach's $\alpha = 0.825$) and positive views (SSPSPositive) (range: 5, 25, Cronbach's $\alpha = 0.736$) of the people in the VR social situation. Higher SSPS scores on each subscale indicate greater levels of persecutory or neutral or positive thinking.

Eye Gaze Data. We recorded the raw data from the eye tracker and the occurrences of intersections between the ray representing the eye gaze direction and each ROI, including the start and end times for each occurrence. Data were first screened for missing records by examining gaps in the raw data output file. We excluded the data from a participant if more than 15% of the data were missing, following the suggested practices in Holmqvist et al. (2012) and Schuetz & Fiehler (2022). We processed the data to calculate the duration for each eye gaze event, and identified fixations using a time threshold of 0.275 seconds (Llanes-Jurado et al., 2020). We then aggregated these fixations for each ROI object per participant. The following variables were calculated for analysis:

- Visual Attention to VHs: total fixation time on VHs as the percentage of total fixation time on all ROI objects.

- Visual Attention to Exit/Floor/Lift Screen (displays the current floor): total fixation time on lift exit door/floor/screen as the percentage of total fixation time on all ROI objects.

- First Fixation Target: the object participants looked at on the first fixation after entering the lift.

5.2.5 Participants and Recruitment

Participants were recruited via social media advertisements in Oxfordshire, United Kingdom. We screened for adults vulnerable to paranoia using The Revised Green et al., Paranoid Thoughts Scale (R-GPTS) (Freeman et al., 2021), with a Part B score greater than 5. This cut-off score captures elevated or higher levels of persecutory ideation. Exclusion criteria were individuals (a) under 18 years old, (b) reported photosensitive epilepsy in the past or a significant visual, hearing, or mobility impairment that would prevent them from using VR, or (c) currently under medication that could induce motion sicknesses. Participants requiring correction-to-normal lenses were requested to use contact lenses instead of eyeglasses to avoid any potential discomfort wearing the VR headset and to avoid interfering with the eye tracker. Ethical approval was received from the University of Oxford Medical Sciences Interdivisional Research Ethics Committee (R85111/RE001).

1581 individuals completed the screening questionnaire, 296 were eligible (i.e. adults with elevated paranoia). Details of the recruitment process are shown in Figure 5.4.

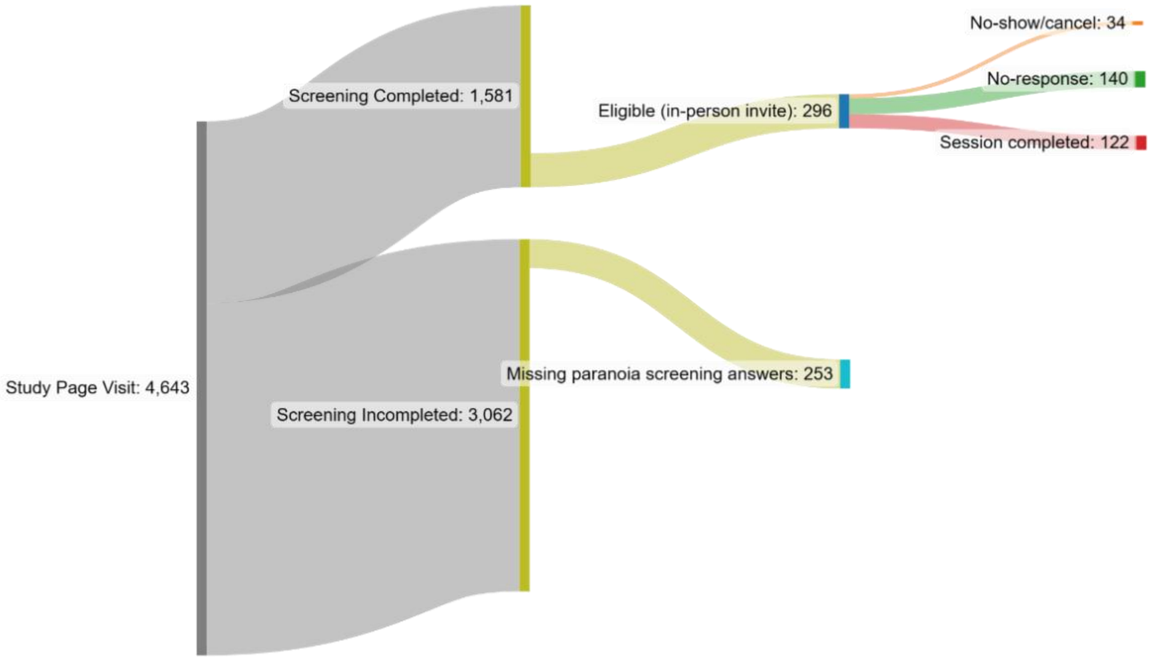


Figure 5.4. Participant recruitment data.

122 participants (female = 70, male = 52) were recruited. We booked two extra participants over the target sample size to account for potential cancellations. The average age of the participants was 36.2 years (SD = 14.8, range: 18, 76). The mean R-GPTS Part B score was 13.30 (SD = 7.70, range: 6, 37). The average previous experience of VR was 2.00 (SD = 1.13). Table 5.1 provides a summary of the participants' demographic information and study-relevant data.

	Static Neutral (n=31)	Static Positive (n=30)	Animated Neutral (n=30)	Animated Positive (n=31)
Mean age in years (SD)	36.2 (16.3)	36.1 (13.6)	36.8 (13.9)	35.9 (15.9)
Gender				
Male (%)	12 (38.7%)	12 (40.0%)	15 (50.0%)	13 (41.9%)
Female (%)	19 (61.3%)	18 (60.0%)	15 (50.0%)	18 (58.1%)
Ethnicity				
White	25	18	22	25
Black/African American	0	1	0	0
Asian	6	7	6	6
Others	0	4	2	0
R-GPTS Part B (i.e. baseline paranoia score) (SD)	13.35 (7.71)	13.07 (7.31)	14.00 (8.18)	12.81 (7.89)
Previous experience of VR (SD)	2.00 (0.93)	2.07 (1.28)	2.40 (1.19)	2.19 (1.17)

Table 5.1. Participant information per group.

5.2.6 Experimental Procedures

Each participant was invited for a single session at our VR lab. The researcher provided an overview of the study procedure and informed participants that they would try out a VR social experience, during which eye gaze data would be collected. They were informed that no pictures or videos of their eyes would be recorded. Participants gave written informed consent to participate. The researcher then helped the participants fit the VR headset and guided them through the eye tracker calibration process. After a successful calibration, the researcher selected the parameters according to the experimental group each participant had been allocated randomly, and they went through the virtual lift experience. When the VR scenario ended, participants took the

VR headset off and completed the paranoid thoughts visual analogue scales (Freeman et al., 2015) and State Social Paranoia Scale (Freeman, Pugh, et al., 2007). Finally, they were debriefed about the full purpose of the study. The session lasted approximately 45 minutes, and participants were reimbursed for their time.

5.2.7 Statistical Methods

We used two-way ANCOVA models, examining the effects on participant views of the VHs of facial animation and positive facial expression while controlling for baseline paranoia. We first checked the data against the assumptions of the ANCOVA model, using Levene's test for homogeneity of variance and Shapiro-Wilk test of normality. We also applied a log transformation on heavily skewed data (visual attention to the floor and screen) before the analysis. Details of the assumption checking results are included in the supplementary materials. A similar approach was taken to analysing the eye gaze data.

All significance tests were made at the $\alpha = 0.05$ level, and we calculated the partial eta-squared (η_p^2) to measure the effect sizes. Tukey's honest significant difference test (Tukey's HSD) was used for multiple pairwise comparisons with the adjusted p value. We report the results as mean group differences and 95% confidence interval (95% CI). Additionally, we conducted contrast tests in cases where a significant interaction was detected, assessing the impact of each factor at different levels of another factor, with estimates reported alongside their 95% confidence interval. Data cleaning and processing was performed using Python's Pandas and NumPy libraries (Harris et al., 2020; McKinney, 2010). The statistical analysis was done in R.

To determine the target sample size for our experimental design, we aimed to detect a medium effect size of $\eta_p^2 = 0.06$ and conventional values of power = 0.80 and $\alpha = 0.05$

for a between-factors ANOVA using G*power 3.124. Thus, a total of 120 participants (30 per condition) would be required.

5.3 RESULTS

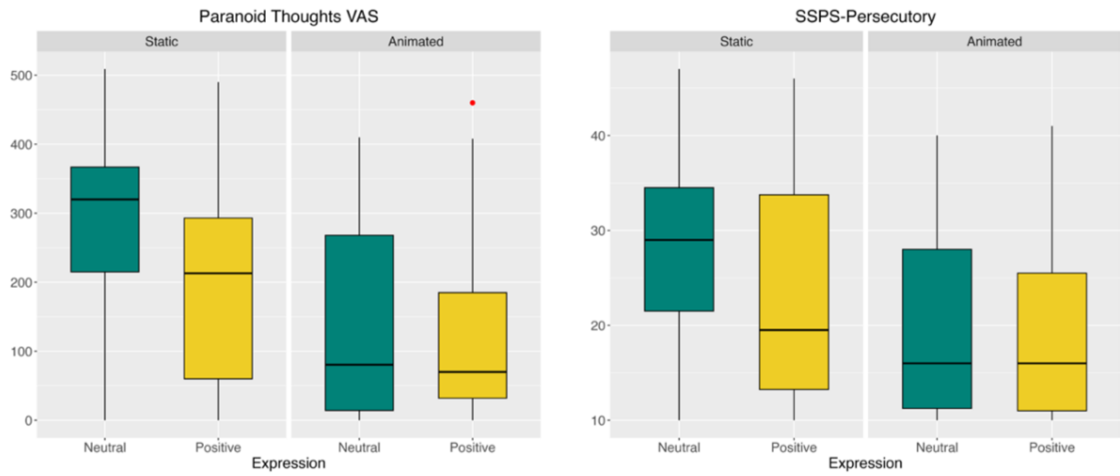


Figure 5.5. Box plots of the ParanoiaVAS and SSPS-Persecutory. Red point indicates the outlier.

Table 5.2 summarises scores for the paranoid thoughts VAS and the three SSPS subscales by randomised group. There were no missing data in these measures. Both the paranoid thoughts VAS and SSPS persecutory were used to assess participants' paranoid ideation and were positively correlated (Spearman $\varrho = 0.82$, $p < 0.001$). Figure 5.5 shows summary scores for these two measures in the randomised groups.

	Static Neutral (n=31)	**Static Positive (n=30)**	**Animated Neutral (n=30)**	**Animated Positive (n=31)**
ParanoiaVAS mean (SD)	285.00 (126.15)	195.33 (146.47)	145.20 (141.88)	132.16 (138.66)
SSPSPersecutory mean (SD)	28.71 (10.86)	20.20 (10.60)	23.03 (10.89)	19.52 (9.80)
SSPSNeutral mean (SD)	11.03 (4.56)	11.87 (4.26)	14.70 (5.08)	13.10 (5.19)

SSPSPositive mean (SD)	11.03 (3.69)	10.87 (4.36)	11.27 (4.17)	14.23 (4.28)

Table 5.2. Descriptive statistics of the appraisals of the VHs by randomisation group.

5.3.1 Paranoid Thoughts VAS (Paranoia VAS)

A two-way ANCOVA model was used to assess the impact of facial animation and expressions while controlling for baseline paranoia. Simple main effects analysis showed that facial animation (group difference = 102.328, 95% CI = [51.783, 152.872], $F(1, 117) = 17.071$, $p < 0.001$, $\eta_p^2 = 0.125$) and positive expression (group difference = 53.016, 95% CI = [0.054, 105.979], $F(1, 117) = 3.938$, $p = 0.049$, $\eta_p^2 = 0.033$) led to less paranoid thinking about the VHs. There was no significant interaction between animation and positive expression ($F(1, 117) = 2.519$, $p = 0.115$, $\eta_p^2 = 0.021$). The effect of baseline paranoia was not significant ($F(1, 117) = 2.678$, $p = 0.104$, $\eta_p^2 = 0.022$). Tukey's HSD test for multiple comparisons showed there was a significant difference between the static neutral and animated neutral group (p-adj < 0.001) and between the static neutral and animated positive group (p-adj < 0.001).

5.3.2 SSPS-Persecutory Thoughts

Simple main effects analysis showed that facial animation (group difference = 6.066, 95% CI = [2.247, 9.885], $F(1, 117) = 10.464$, $p = 0.002$, $\eta_p^2 = 0.081$) but not positive expressions (group difference = 3.279, 95% CI = [-0.650, 7.207], $F(1, 117) = 2.534$, $p = 0.114$, $\eta_p^2 = 0.021$) led to significantly lower levels of paranoia. There was no significant interaction between animation and positive expressions ($F(1, 117) = 1.897$, $p = 0.171$, $\eta_p^2 = 0.016$). The effect of baseline paranoia was not significant ($F(1, 117) = 3.624$, $p = 0.059$, $\eta_p^2 = 0.030$). Tukey's HSD test for multiple comparisons showed a statistically significant difference between the static neutral group and animated neutral group (p-adj = 0.008), and between the static neutral and animated positive group (p-adj = 0.005).

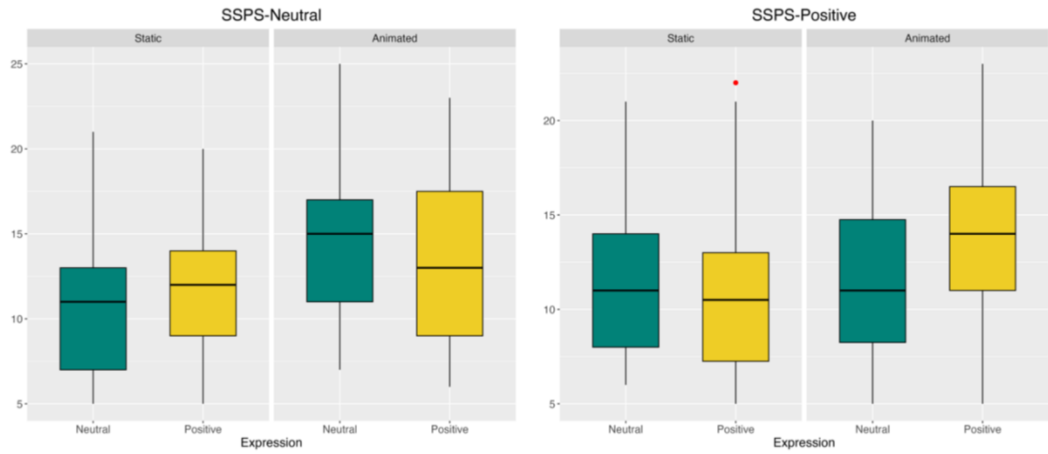


Figure 5.6. Box plots of the SPSS Neutral and Positive Thoughts scores. Red point indicates the outlier.

5.3.3 SSPS-Neutral Thoughts

Simple main effects analysis showed that facial animation (group difference = 2.442, 95% CI = [-4.161, -0.724], F(1, 117) = 7.843, p = 0.006, η_p^2 = 0.063) but not positive expressions (group difference = 0.344, 95% CI = [-1.429, 2.110], F(1, 117) = 0.17, p = 0.681, η_p^2 = 0.001) led to a more neutral interpretation of the VHs. There was no significant interaction between animation and positive expressions (F(1, 117) = 1.914, p = 0.169, η_p^2 = 0.016). The effect of baseline paranoia was not significant (F(1, 117) = 0.442, p = 0.508, η_p^2 = 0.002). Tukey's HSD test for multiple comparisons showed there was a statistically significant difference between the static neutral group and animated neutral group (p-adj = 0.019).

5.3.4 SSPS-Positive Thoughts

There was a significant interaction between animation and positive expressions (F(1, 117) = 4.297, p = 0.040, η_p^2 = 0.035). Facial animation led to more positive thoughts when the expressions were positive (estimate = 3.358 [1.250, 5.460], SE = 1.060, p = 0.002), but not when expressions were neutral (estimate = 0.236 [-1.870, 2.340], SE = 1.060, p = 0.825). Positive expressions led to more positive thoughts when the faces were animated (estimate = 2.956 [0.849, 5.060], SE = 1.060, p = 0.006), but not when faces were static (estimate = -0.166 [-2.271, 1.940], SE = 1.060, p = 0.876). Tukey's HSD test for multiple comparisons showed that there was a statistically significant difference between the static neutral and animated positive group (p-adj = 0.016), the static positive and animated positive group (p-adj = 0.011), and between the animated neutral and animated positive group (p-adj = 0.032).

5.3.5 Paranoid thinking in VR and Baseline Paranoia

	Spearman Correlation (ϱ)	**p Value**
Static Neutral (n=31)	-0.110	0.554
Static Positive (n=30)	0.000	0.999
Animated Neutral (n=30)	0.385	0.036 *
Animated Positive (n=31)	0.194	0.296

Table 5.3. Correlation between VAS Paranoia and baseline paranoia.

A moderate positive correlation (ϱ = 0.385) was found between paranoid thinking in VR and baseline paranoia within the animated neutral group.

5.3.6 Visual Attention

Five datasets from the eye-tracking analysis were excluded as more than 15% of the raw data were missing due to technical issues. Analysis of the remaining 117 participants during the lift ride showed an average total fixation duration of 112.59 seconds (SD = 28.42) and an average of 1.86 seconds (SD = 1.84) per fixation. Participants spent 31.95% of the fixation time on the VHs (SD = 24.68%). The average duration for these fixations was 5.87 seconds (SD = 5.84). The most common initial fixation targets were the male VH directly facing the lift entrance (27.4%), the exit (17.9%), and the floor (17.9%) (see Figure 5.7). Descriptive statistics for the visual attention allocation are shown in Table 5.4.

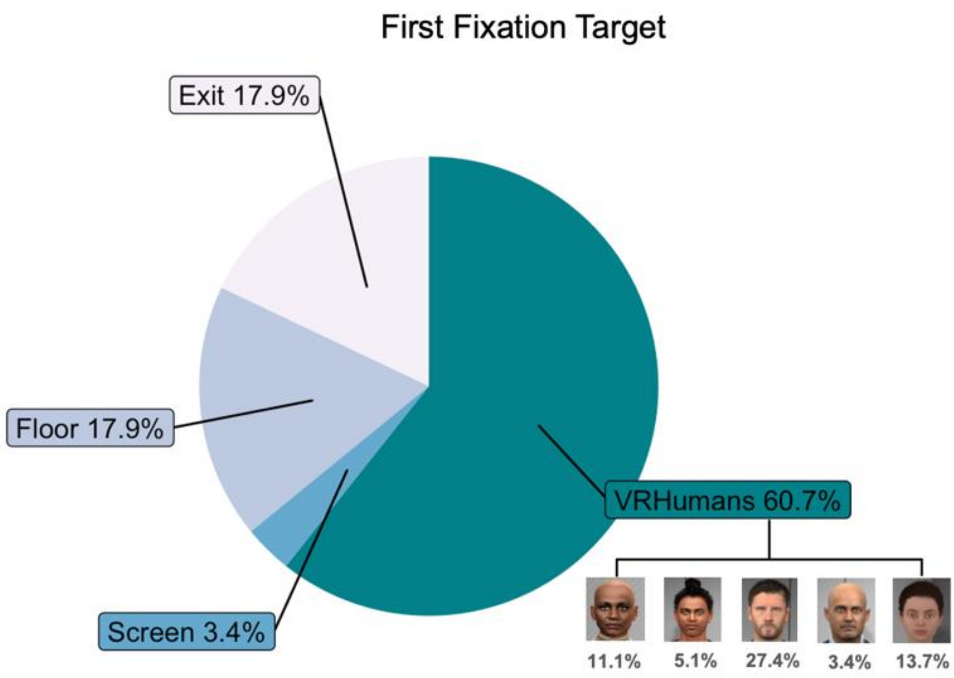


Figure 5.7. Distribution of the target of first fixation.

Visual Attention Allocation (%)	Static Neutral (n=30)	Static Positive (n=28)	Animated Neutral (n=30)	Animated Positive (n=29)
VR Humans (SD)	38.8 (20.6)	21.9 (19.7)	32.6 (27.3)	33.8 (27.9)
Exit (SD)	23.8 (21.2)	41.8 (30.8)	41.9 (36.6)	26.7 (29.3)
Floor (SD)	9.7 (18.3)	8.1 (15.5)	7.4 (14.5)	9.0 (11.6)
Screen (SD)	25.7 (19.7)	26.5 (23.4)	16.8 (16.1)	29.1 (26.0)

Table 5.4. Visual attention to VHs, lift exit, lift floor, and lift screen. (values represent the average percentage of fixation time to ROIs from participants within each condition group; numbers are rounded to one decimal place)

Visual attention to the VHs was tested using a two-way ANCOVA model controlling for baseline paranoia. There was a marginally non-significant interaction between animation and positive expressions ($F(1, 112) = 3.591$, $p = 0.061$, $\eta_p^2 = 0.031$), suggesting a trend where animation and positive expression might jointly influence the amount of visual attention allocated to the VHs. Specifically, positive expressions led to a lower amount of visual attention on VHs when faces were static (estimate = -0.169 [-0.294, -0.043], SE = 0.064, p = 0.010), but not when faces were animated (estimate = 0.001 [-0.124, 0.127], SE = 0.063, p = 0.983). Facial animation did not affect visual attention to VHs either when the facial expressions were neutral (estimate = -0.060 [-0.184, 0.064], SE = 0.062, p = 0.338) or positive (estimate = 0.110 [-0.017, 0.238], SE = 0.064, p = 0.090). Tukey's HSD test indicated a significant difference between the static neutral and static positive group (p-adj = 0.044).

Visual attention to the environment. The same two-way ANCOVA model was used to examine the extent of visual attention on the lift exit door, floor, and screen separately. For the exit, there was a significant interaction between animation and positive expressions ($F(1, 112) = 8.026$, $p = 0.005$, $\eta_p^2 = 0.067$). Facial animation led to a higher amount of visual attention on the exit when the expressions were neutral (estimate = 0.176 [0.027, 0.325], SE = 0.075, p = 0.021), but not when expressions were positive (estimate = -0.131 [-0.285, 0.023], SE = 0.078, p = 0.095). Positive expressions led to a higher amount of visual attention on the exit when faces were static (estimate = 0.182 [0.030, 0.334], SE = 0.077, p = 0.020), but not when faces were animated (estimate = -0.125 [-0.277, 0.027], SE = 0.077, p = 0.105). Tukey's HSD test indicated no statistically significant differences in pairwise comparisons.

The analysis was performed on log-transformed data for visual attention to the floor and the lift screen. For the floor, there was no significant interaction between animation and positive expressions ($F(1, 112) = 0.306$, $p = 0.581$, $\eta_p^2 = 0.003$) and there were no main effects from animation ($F(1, 112) = 0.025$, $p = 0.874$, $\eta_p^2 < 0.001$) or positive expression ($F(1, 112) = 0.011$, $p = 0.916$, $\eta_p^2 < 0.001$). The effect of baseline mistrust was not significant ($F(1, 112) = 2.493$, $p = 0.117$, $\eta_p^2 = 0.020$). For the lift screen, there was no significant interaction between animation and positive expressions ($F(1, 112) = 1.895$, $p = 0.171$, $\eta_p^2 = 0.017$) and there were no main effects from animation ($F(1, 112) = 0.846$, $p = 0.360$, $\eta_p^2 = 0.009$) or positive expression ($F(1, 112) = 2.496$, $p = 0.117$, $\eta_p^2 = 0.019$). The effect of baseline mistrust was not significant ($F(1, 112) = 1.250$, $p = 0.266$, $\eta_p^2 = 0.009$).

Correlation between visual attention and paranoia. We computed Spearman correlations between visual attention and paranoia using all the retained eye-tracking data across different condition groups (N = 117). There was a positive correlation between the amount of visual attention to the VHs and the severity of paranoid thoughts in VR (VAS: $r = 0.19$, $p = 0.040$) and a positive correlation between the amount of visual attention to the lift exit and baseline paranoia (Baseline Paranoia: $r = 0.21$, $p = 0.023$). There were no significant correlations between the paranoid thoughts in VR/ baseline paranoia and the visual attention to other ROIs. Additional correlation details between the eye-tracking data and other measures are presented in Figure 5.8.

	Correlation with VAS Paranoia		**Correlation with Baseline Paranoia**	
Visual Attention Allocation (%)	**Spearman ϱ**	**p Value**	**Spearman ϱ**	**p Value**
Virtual humans	0.190	0.040	-0.046	0.621
Exit	-0.152	0.101	0.210	0.023
Floor	0.117	0.208	-0.101	0.277
Screen	-0.040	0.668	-0.079	0.398

Table 5.5. Correlation between visual attention and paranoia measures (N=117).

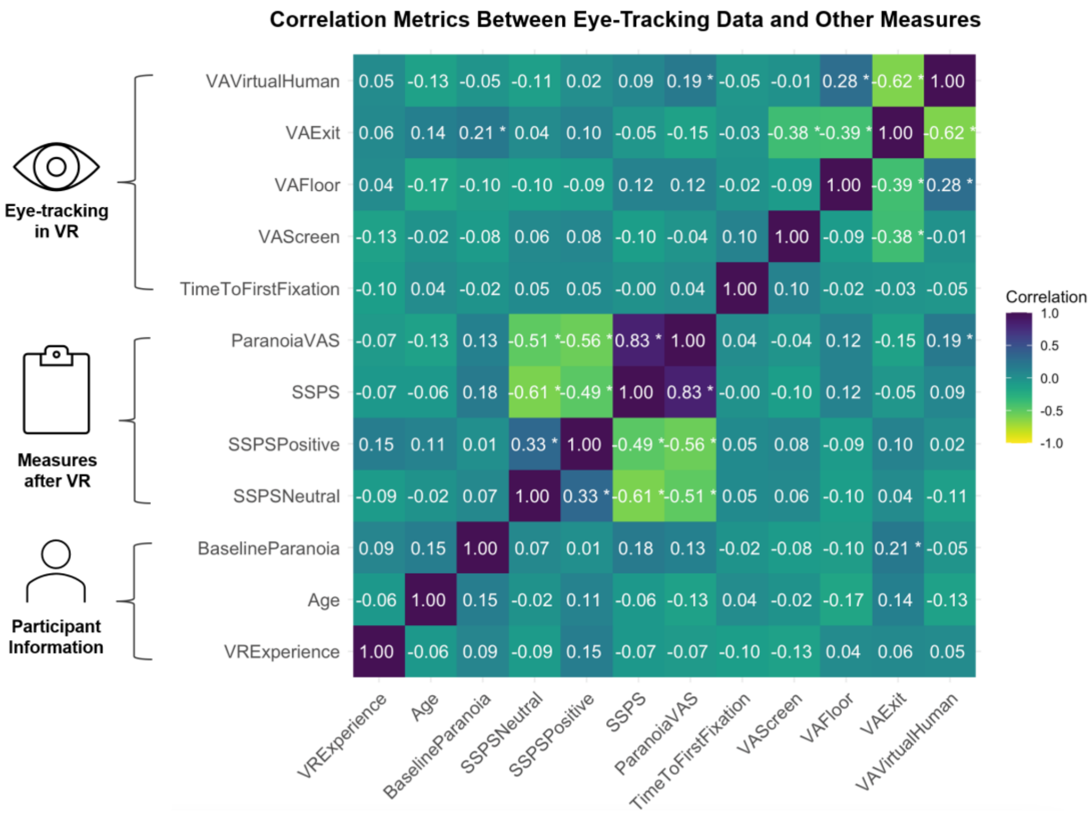


Figure 5.8. Correlation matrix between variables. The number in each cell shows the spearman correlation coefficient, with a star (*) indicating that the correlation is significant at the 0.05 level.

5.4 DISCUSSION

In this chapter, we presented the results of our second experimental study, which examined how the animations and facial expressions of VHs affect paranoid interpretations when VHs are part of the stimuli. The results confirmed the hypothesis that facial animation and positive facial expressions of VHs both reduce the likelihood of paranoid appraisals. In our study, facial animation and positive expression each independently led to people vulnerable to paranoia perceiving the virtual characters as less hostile. In contrast, paranoid thoughts were more likely to occur when faces were static or the expression was neutral. The sizes of the effects were moderate to large.

Facial animation fostered more neutral perceptions of VHs too. An examination of the correlation between paranoid thinking in VR and baseline paranoia indicated that this correlation was strongest in the animated neutral condition. This means that animated neutral characters in VR may provide the most effective assessment test for paranoia. Overall, the study highlights the importance of how facial expressions are designed and animated when assessing or treating paranoia in VR.

The findings align with prior research showing that facial-animated virtual characters appear more natural and believable (Kullmann et al., 2023), and people are better at recognizing emotions from them (Faita et al., 2015). Similarly, the addition of positive emotion led to the perception of less negative intention or attributes from the virtual characters regardless of whether their faces were animated (Geraets et al., 2021; Wei et al., 2023). Notably, the effect of participants' baseline paranoia was not significant, indicating that their appraisals of the VHs were shaped primarily by the facial animation and expressions rather than their initial levels of paranoia. Among the two features, animation had a stronger effect (accounting for 13% of the variance in paranoia) compared to facial expressions (accounting for 3% of the variance). This could be attributed to our implementation of positive expressions as friendly faces with gentle, subtle smiles, to fit the neutral context of a lift ride. Such nuanced emotional expression typically requires accurate delivery with dynamic movement (Recio et al., 2011); the absence of animations may lead to the "frozen face" effect, where a static human face appears less flattering than one with motion (Post et al., 2012). Additionally, the lack of dynamic information in the VH faces could render their expressions more ambiguous, leading individuals with elevated paranoia lean towards interpreting this ambiguity negatively and perceive the VHs as potentially hostile (Green et al., 2011; Savulich et al., 2015).

The design and implementation of VH faces also influenced individuals' neutral and positive perceptions. According to Krumhuber et al. (2023), dynamic information (e.g. animation) enhances emotion recognition, particularly when facial expressions are subtle or convey a neutral emotion. Consistent with this, animation led to a more neutral interpretation of the characters, and the animated neutral faces were rated as the most neutral. Interestingly, regarding positive thoughts about the characters, an interaction effect suggested that animation was important for positive expressions to lead to stronger positive interpretations, while the static positive faces scored the lowest. This reduced likelihood of eliciting positive thoughts from static positive faces likely stems from the mismatch between expressed emotions and the absence of movement, making the characters appear less lively and emotionally inconsistent (Kleinsmith & Bianchi-Berthouze, 2013).

Examining visual attention to VHs, we found that positive expressions led to less visual attention when VH faces were static. The discrepancy between positive facial expressions and the absence of animation could have led to the characters being perceived as anomalous. Particularly, the lack of eye movements, such as eye blinking or gaze, could cause smiles to seem eerie or ungenuine (Liew et al., 2016; Tinwell et al., 2011). Exploring the link between visual attention and paranoia, we found a positive relationship between attention to VHs and paranoid thoughts in VR. This might imply that closely observing VHs could foster the development of paranoid ideations (Fornells-Ambrojo et al., 2015), or that individuals with heightened paranoia are more inclined to concentrate on these characters. In addition, there was also a positive relationship between people's focus on the lift exit and their paranoid tendency in real life, measured by paranoia scores at baseline. This behaviour aligns with the use of safety-seeking strategies in response to persecutory thoughts (Freeman, Garety, et al., 2007). Such behaviour also coincides with the patterns found in social anxiety studies, where individuals often avoid eye contact with VHs and shift their attention to other areas of the virtual scene under distress (Reichenberger et al., 2020; Rubin et al., 2020). The connection between visual attention, character animation, and people's mental health state can be complex and requires further exploration, and the study of VHs eye gaze behaviour (e.g. whether they make eye contact with participants) could provide additional information to help understand paranoia.

There are a number of limitations and further directions to the study. First, we focused on two features of VHs - animation and positive expressions – but other characteristics such as VH's eye gaze behaviour patterns and facial mimicry are likely to be important too (Hale & Hamilton, 2016; Raimbaud et al., 2022). Second, we did not consider demographic (e.g. age, gender, ethnicity) similarities or differences between participants and the characters, nor their spatial arrangement. Third, our visual attention analysis focused only on spatial allocation and fixation-related metrics, excluding other potentially relevant measures such as the frequency of eye saccades and the angle at which participants looked at the VH, to understand whether they looked at them directly or out of the corner of their eye. Current eye trackers in VR headsets have limited accuracy, and accuracy decreases the further they deviate from the centre of their sight. Additionally, we considered participants were looking at VH when their gaze direction intersected with an invisible box roughly the size of VH's head. The resolution of the eye tracker was insufficient for a more detailed analysis to determine whether participants were looking at specific facial features, such as the eyes or mouth. Prior research has shown that people put their attention to different parts of virtual faces depending on the displayed emotions (Geraets et al., 2021), so in the future, eye trackers with higher precision will likely provide more information about participants' behaviour. Finally, future research could explore other scenarios besides the VR lift, which limited participants to close proximity (under 2 meters) from the VHs. Comparing this with scenarios involving greater social distances (e.g., entering a room) would be valuable, as interpersonal distance in VR can affect people's emotional and behavioural responses (Kroczek et al., 2020).

Our findings provide further evidence that character animations alter people's perceptions and experiences in VR. This may, for example, affect VR experiences focused on the understanding and treatment of paranoia. Therefore, careful consideration of character design and animation is likely to be important in developing future VR mental health applications.

References

Aburumman, N., Gillies, M., Ward, J. A., & Hamilton, A. F. de C. (2022). Nonverbal communication in virtual reality: Nodding as a social signal in virtual interactions. *International Journal of Human-Computer Studies*, *164*, 102819. https://doi.org/10.1016/j.ijhcs.2022.102819

Amaral, C., Mouga, S., Simões, M., Pereira, H. C., Bernardino, I., Quental, H., Playle, R., McNamara, R., Oliveira, G., & Castelo-Branco, M. (2018). A Feasibility Clinical Trial to Improve Social Attention in Autistic Spectrum Disorder (ASD) Using a Brain Computer Interface. *Frontiers in Neuroscience*, *12*, 477. https://doi.org/10.3389/fnins.2018.00477

Amestoy Alonso, B., Donegan, T., Calvis, I., Swidrak, J., Rodriguez, E., Vargas-Reverón, C. L., Combalia, A., Oliva Martinez, R., & Sanchez-Vives, M. V. (2024). Focus groups in the metaverse: Shared virtual spaces for patients, clinicians, and researchers. *Frontiers in Virtual Reality*, *5*. https://doi.org/10.3389/frvir.2024.1432282

APA. (2020). *Publication manual of the American Psychological Association, 7th ed* (pp. xxii, 427). American Psychological Association. https://doi.org/10.1037/0000165-000

Ardito, R. B., & Rabellino, D. (2011). Therapeutic Alliance and Outcome of Psychotherapy: Historical Excursus, Measurements, and Prospects for Research. *Frontiers in Psychology*, *2*, 270. https://doi.org/10.3389/fpsyg.2011.00270

Artiran, S., Bedmutha, P. S., & Cosman, P. (2024). Analysis of Gaze, Head Orientation, and Joint Attention in Autism With Triadic VR Interviews. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *32*, 759–769. https://doi.org/10.1109/TNSRE.2024.3363728

Arza, A., Garzón-Rey, J. M., Lázaro, J., Gil, E., Lopez-Anton, R., de la Camara, C., Laguna, P., Bailon, R., & Aguiló, J. (2019). Measuring acute stress response through physiological signals: Towards a quantitative assessment of stress. *Medical & Biological Engineering & Computing*, *57*(1), 271–287. https://doi.org/10.1007/s11517-018-1879-z

Ascione, M., Carulla-Roig, M., Miquel-Nabau, H., Porras-Garcia, B., Meschberger-Annweiler, F.-A., Serrano-Troncoso, E., Ferrer-Garcia, M., Moreno-Sánchez, M., & Gutierrez-Maldonado, J. (2023). Attentional Bias Modification Training Based on Virtual Reality and Eye Tracking in Anorexia Nervosa Patients. *Journal of Clinical Medicine*, *12*(18), 5932. https://doi.org/10.3390/jcm12185932

Atherton, S., Antley, A., Evans, N., Cernis, E., Lister, R., Dunn, G., Slater, M., & Freeman, D. (2016). Self-Confidence and Paranoia: An Experimental Study Using an Immersive Virtual Reality Social Situation. *Behavioural and Cognitive Psychotherapy*, *44*(1), 56–64. https://doi.org/10.1017/S1352465814000496

Aymerich-Franch, L., Kizilcec, R. F., & Bailenson, J. N. (2014). The Relationship between Virtual Self Similarity and Social Anxiety. *Frontiers in Human Neuroscience*, *8*, 944. https://doi.org/10.3389/fnhum.2014.00944

Aziz, S., & Komogortsev, O. (2022). An Assessment of the Eye Tracking Signal Quality Captured in the HoloLens 2. *2022 Symposium on Eye Tracking Research and Applications*, 1–6. https://doi.org/10.1145/3517031.3529626

Banakou, D., Beacco, A., Neyret, S., Blasco-Oliver, M., Seinfeld, S., & Slater, M. (2020). Virtual body ownership and its consequences for implicit racial bias are dependent on social context. *Royal Society Open Science*, *7*(12), 201848. https://doi.org/10.1098/rsos.201848

Banakou, D., Johnston, T., Beacco, A., Senel, G., & Slater, M. (2024). Desensitizing Anxiety Through Imperceptible Change: Feasibility Study on a Paradigm for Single-Session Exposure Therapy for Fear of Public Speaking. *JMIR Formative Research*, *8*(1), e52212. https://doi.org/10.2196/52212

Barisnikov, K., Thomasson, M., Stutzmann, J., & Lejeune, F. (2021). Sensitivity to Emotion Intensity and Recognition of Emotion Expression in Neurotypical Children. *Children*, *8*(12), 1108. https://doi.org/10.3390/children8121108

Barrett, L. F., & Russell, J. A. (1999). The Structure of Current Affect: Controversies and Emerging Consensus. *Current Directions in Psychological Science*, *8*(1), 10–14. https://doi.org/10.1111/1467-8721.00003

Bektas, S., Natali, L., Rowlands, K., Valmaggia, L., Di Pietro, J., Mutwalli, H., Himmerich, H., Treasure, J., & Cardi, V. (2023). Exploring Correlations of Food-Specific Disgust with Eating Disorder Psychopathology and Food Interaction: A Preliminary Study Using Virtual Reality. *Nutrients*, *15*(20), Article 20. https://doi.org/10.3390/nu15204443

Bell, I. H., Nicholas, J., Alvarez-Jimenez, M., Thompson, A., & Valmaggia, L. (2020). Virtual reality as a clinical tool in mental health research and practice. *Dialogues in Clinical Neuroscience*, *22*(2), 169–177. https://doi.org/10.31887/DCNS.2020.22.2/lvalmaggia

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, *190*(1), 80–91. https://doi.org/10.1016/j.jneumeth.2010.04.028

Bickmore, T. W., & Picard, R. W. (2004). Towards caring machines. *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, 1489–1492. https://doi.org/10.1145/985921.986097

Bolinski, F., Etzelmüller, A., De Witte, N. A. J., van Beurden, C., Debard, G., Bonroy, B., Cuijpers, P., Riper, H., & Kleiboer, A. (2021). Physiological and self-reported arousal in virtual reality versus face-to-face emotional activation and cognitive restructuring in university students: A crossover experimental study using wearable monitoring. *Behaviour Research and Therapy*, *142*, 103877. https://doi.org/10.1016/j.brat.2021.103877

Bombari, D., Schmid Mast, M., Canadas, E., & Bachmann, M. (2015). Studying social interactions through immersive virtual environment technology: Virtues, pitfalls, and future challenges. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00869

Bönsch, A., Radke, S., Ehret, J., Habel, U., & Kuhlen, T. W. (2020). The Impact of a Virtual Agent's Non-Verbal Emotional Expression on a User's Personal Space Preferences. *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*, 1–8. https://doi.org/10.1145/3383652.3423888

Brander, M., Egger, S. T., Hürlimann, N., Seifritz, E., Sumner, R. W., Vetter, S., & Magnenat, S. (2021). Virtual Reality Human–Human Interface to Deliver Psychotherapy to People Experiencing Auditory Verbal Hallucinations: Development and Usability Study. *JMIR Serious Games*, *9*(2). https://doi.org/10.2196/26820

Braun, H. I., Jackson, D. N., & Wiley, D. E. (Eds.). (2001). Socially Desirable Responding: The Evolution of a Construct. In *The Role of Constructs in Psychological and Educational Measurement* (0 ed., pp. 61–84). Routledge. https://doi.org/10.4324/9781410607454-10

Brenner, P. S., & DeLamater, J. (2016). Lies, Damned Lies, and Survey Self-Reports? Identity as a Cause of Measurement Bias. *Social Psychology Quarterly*, *79*(4), 333–354. https://doi.org/10.1177/0190272516628298

Brett, J., Staniszewska, S., Mockford, C., Herron-Marx, S., Hughes, J., Tysall, C., & Suleman, R. (2014). A Systematic Review of the Impact of Patient and Public Involvement on Service Users, Researchers and Communities. *The Patient - Patient-Centered Outcomes Research*, *7*(4), 387–395. https://doi.org/10.1007/s40271-014-0065-0

Brinkman, W.-P., Veling, W., Dorrestijn, E., Sandino, G., Vakili, V., & Gaag, M. (2011). Virtual Reality to study responses to social environmental stressors in individuals with and without psychosis. *Annual Review of CyberTherapy and Telemedicine*, 86–91. https://doi.org/10.3233/978-1-6-750-766-6-86

Brown, P., Waite, F., Lambe, S., Jones, J., Jenner, L., Diamond, R., & Freeman, D. (2022). Automated Virtual Reality Cognitive Therapy (gameChange) in Inpatient Psychiatric Wards: Qualitative Study of Staff and Patient Views Using an Implementation Framework. *JMIR Formative Research*, *6*(4), e34225. https://doi.org/10.2196/34225

Brown, P., Waite, F., Rovira, A., Nickless, A., & Freeman, D. (2020). Virtual reality clinical-experimental tests of compassion treatment techniques to reduce paranoia. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-64957-7

Bruder, L. R., Scharer, L., & Peters, J. (2020). *A Virtual Reality set-up for the combined assessment of behavioral and psychophysiological cue-reactivity in gambling disorder* (p. 2020.08.07.237826). bioRxiv. https://doi.org/10.1101/2020.08.07.237826

Burin, D., Cavanna, G., Rabellino, D., Kotozaki, Y., & Kawashima, R. (2022). Neuroendocrine Response and State Anxiety Due to Psychosocial Stress Decrease after a Training with Subject's Own (but Not Another) Virtual Body: An RCT Study. *International Journal of Environmental Research and Public Health*, *19*(10), 6340. https://doi.org/10.3390/ijerph19106340

Can, Y., Mahesh, B., & Andre, E. (2023). Approaches, Applications, and Challenges in Physiological Emotion Recognition—A Tutorial Overview. *Proceedings of the IEEE*, *PP*, 1–27. https://doi.org/10.1109/JPROC.2023.3286445

Cassell, J. (2001). Embodied Conversational Agents: Representation and Intelligence in User Interfaces. *AI Magazine*, *22*(4), Article 4. https://doi.org/10.1609/aimag.v22i4.1593

Cerritelli, F., Chiera, M., Abbro, M., Megale, V., Esteves, J., Gallace, A., & Manzotti, A. (2021). The Challenges and Perspectives of the Integration Between Virtual and Augmented Reality and Manual Therapies. *Frontiers in Neurology*, *12*, 700211. https://doi.org/10.3389/fneur.2021.700211

Chang, Z., Bai, H., Zhang, L., Gupta, K., He, W., & Billinghurst, M. (2022). The impact of virtual agents' multimodal communication on brain activity and cognitive load in Virtual Reality. *Frontiers in Virtual Reality*, *3*. https://doi.org/10.3389/frvir.2022.995090

Chattopadhyay, D., Ma, T., Sharifi, H., & Martyn-Nemeth, P. (2020). Computer-Controlled Virtual Humans in Patient-Facing Systems: Systematic Review and Meta-Analysis. *Journal of Medical Internet Research*, *22*(7), e18839. https://doi.org/10.2196/18839

Cho, S., Ku, J., Park, J., Han, K., Lee, H., Choi, Y. K., Jung, Y.-C., Namkoong, K., Kim, J.-J., Kim, I. Y., Kim, S. I., & Shen, D. F. (2008a). Development and Verification of an Alcohol Craving–Induction Tool Using Virtual Reality: Craving Characteristics in Social Pressure Situation. *CyberPsychology & Behavior*, *11*(3), Article 3. https://doi.org/10.1089/cpb.2007.0149

Cho, S., Ku, J., Park, J., Han, K., Lee, H., Choi, Y. K., Jung, Y.-C., Namkoong, K., Kim, J.-J., Kim, I. Y., Kim, S. I., & Shen, D. F. (2008b). Development and Verification of an Alcohol Craving–Induction Tool Using Virtual Reality: Craving Characteristics in Social Pressure Situation. *CyberPsychology & Behavior*, *11*(3), 302–309. https://doi.org/10.1089/cpb.2007.0149

Choudhary, Z., Norouzi, N., Erickson, A., Schubert, R., Bruder, G., & Welch, G. F. (2023). Exploring the Social Influence of Virtual Humans Unintentionally Conveying Conflicting Emotions. *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, 571–580. https://doi.org/10.1109/VR55154.2023.00072

Christinaki, E., Vidakis, N., & Triantafyllidis, G. (2013). Facial expression recognition teaching to preschoolers with autism: A natural user interface approach. *Proceedings of the 6th Balkan Conference in Informatics*, 141–148. https://doi.org/10.1145/2490257.2490262

Citron, F. M. M., Gray, M. A., Critchley, H. D., Weekes, B. S., & Ferstl, E. C. (2014). Emotional valence and arousal affect reading in an interactive way: Neuroimaging evidence for an approach-withdrawal framework. *Neuropsychologia*, *56*(100), 79–89. https://doi.org/10.1016/j.neuropsychologia.2014.01.002

Ciuk, D., Troy, A. K., & Jones, M. C. (2015). Measuring Emotion: Self-Reports vs. Physiological Indicators. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2595359

Clavel, C., Plessier, J., Martin, J.-C., Ach, L., & Morel, B. (2009). Combining Facial and Postural Expressions of Emotions in a Virtual Character. In Z. Ruttkay, M. Kipp, A. Nijholt, & H. H. Vilhjálmsson (Eds.), *Intelligent Virtual Agents* (pp. 287–300). Springer. https://doi.org/10.1007/978-3-642-04380-2_31

Cornwell, B. R., Johnson, L., Berardi, L., & Grillon, C. (2006). Anticipation of Public Speaking in Virtual Reality Reveals a Relationship Between Trait Social Anxiety and Startle Reactivity. *Biological Psychiatry*, *59*(7), 664–666. https://doi.org/10.1016/j.biopsych.2005.09.015

Del Piccolo, L., & Finset, A. (2018). Patients' autonomic activation during clinical interaction: A review of empirical studies. *Patient Education and Counseling*, *101*(2), 195–208. https://doi.org/10.1016/j.pec.2017.08.007

DeVault, D., Artstein, R., Benn, G., Dey, T., Fast, E., Gainer, A., Georgila, K., Gratch, J., Hartholt, A., Lhommet, M., Lucas, G., Marsella, S., Morbini, F., Nazarian, A., Scherer, S., Stratou, G., Suri, A., Traum, D., Wood, R., … Morency, L.-P. (2014). SimSensei kiosk: A virtual human interviewer for healthcare decision support. *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-Agent Systems*, 1061–1068.

DeVault, D., Georgila, K., Artstein, R., Morbini, F., Traum, D., Scherer, S., Rizzo, A. S., & Morency, L.-P. (2013). Verbal indicators of psychological distress in interactive dialogue with a virtual human. In M. Eskenazi, M. Strube, B. Di Eugenio, & J. D. Williams (Eds.), *Proceedings of the SIGDIAL 2013 Conference* (pp. 193–202). Association for Computational Linguistics. https://aclanthology.org/W13-4032

Devilly, G. J., & Borkovec, T. D. (2000). Psychometric properties of the credibility/expectancy questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*, *31*(2), 73–86. https://doi.org/10.1016/S0005-7916(00)00012-4

Dickinson, P., Gerling, K., Hicks, K., Murray, J., Shearer, J., & Greenwood, J. (2019). Virtual reality crowd simulation: Effects of agent density on user experience and behaviour. *Virtual Reality*, *23*(1), 19–32. https://doi.org/10.1007/s10055-018-0365-0

Diemer, J., Lohkamp, N., Mühlberger, A., & Zwanzger, P. (2016). Fear and physiological arousal during a virtual height challenge—Effects in patients with acrophobia and healthy controls. *Journal of Anxiety Disorders*, *37*, 30–39. https://doi.org/10.1016/j.janxdis.2015.10.007

Dimascio, A., Boyd, R. W., & Greenblatt, M. (1957). Physiological Correlates of Tension and Antagonism During Psychotherapy: A Study of 'Interpersonal Physiology'. *Psychosomatic Medicine*, *19*(2), 99.

Dobre, G. C., Gillies, M., & Pan, X. (2022). Immersive machine learning for social attitude detection in virtual reality narrative games. *Virtual Reality*, *26*(4), 1519–1538. https://doi.org/10.1007/s10055-022-00644-4

Dobs, K., Bülthoff, I., & Schultz, J. (2018). Use and Usefulness of Dynamic Face Stimuli for Face Perception Studies—A Review of Behavioral Findings and Methodology. *Frontiers in Psychology*, *9*, 1355. https://doi.org/10.3389/fpsyg.2018.01355

Dong, Y., & Peng, C. (2019). Real-Time Large Crowd Rendering with Efficient Character and Instance Management on GPU. *International Journal of Computer Games Technology*, *2019*, e1792304. https://doi.org/10.1155/2019/1792304

du Sert, O. P., Potvin, S., Lipp, O., Dellazizzo, L., Laurelli, M., Breton, R., Lalonde, P., Phraxayavong, K., O'Connor, K., Pelletier, J.-F., Boukhalfi, T., Renaud, P., & Dumais, A. (2018). Virtual reality therapy for refractory auditory verbal hallucinations in schizophrenia: A pilot clinical trial. *Schizophrenia Research*, *197*, 176–181. https://doi.org/10.1016/j.schres.2018.02.031

Dyer, M. L., Attwood, A. S., Penton-Voak, I. S., & Munafò, M. R. (2022). The role of state and trait anxiety in the processing of facial expressions of emotion. *Royal Society Open Science*, *9*(1), 210056. https://doi.org/10.1098/rsos.210056

Ekman, P. (2004). *Emotions Revealed: Understanding Faces and Feelings*. https://www.amazon.co.uk/Emotions-Revealed-Understanding-Faces-Feelings/dp/0753817659

El Ali, A., Ney, R., Van Berlo, Z. M. C., & Cesar, P. (2023). Is that My Heartbeat? Measuring and Understanding Modality-Dependent Cardiac Interoception in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, *29*(11), 4805–4815. https://doi.org/10.1109/TVCG.2023.3320228

Elkin, L. A., Kay, M., Higgins, J. J., & Wobbrock, J. O. (2021). An Aligned Rank Transform Procedure for Multifactor Contrast Tests. *The 34th Annual ACM Symposium on User Interface Software and Technology*, 754–768. https://doi.org/10.1145/3472749.3474784

Elshan, E., Zierau, N., Engel, C., Janson, A., & Leimeister, J. M. (2022). Understanding the Design Elements Affecting User Acceptance of Intelligent Agents: Past, Present and Future. *Information Systems Frontiers*, *24*(3), 699–730. https://doi.org/10.1007/s10796-021-10230-9

Empatica. (2024). *E4 data—EDA potential issues*. Empatica Support. https://support.empatica.com/hc/en-us/articles/204676479-E4-data-EDA-potential-issues

Faita, C., Vanni, F., Lorenzini, C., Carrozzino, M., Tanca, C., & Bergamasco, M. (2015). Perception of Basic Emotions from Facial Expressions of Dynamic Virtual Avatars. In L. T. De Paolis & A. Mongelli (Eds.), *Augmented and Virtual Reality* (pp. 409–419). Springer International Publishing. https://doi.org/10.1007/978-3-319-22888-4_30

Falconer, C. J., Rovira, A., King, J. A., Gilbert, P., Antley, A., Fearon, P., Ralph, N., Slater, M., & Brewin, C. R. (2016). Embodying self-compassion within virtual reality and its effects on patients with depression. *BJPsych Open*, *2*(1), 74–80. https://doi.org/10.1192/bjpo.bp.115.002147

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Felnhofer, A., Knaust, T., Weiss, L., Goinska, K., Mayer, A., & Kothgassner, O. D. (2024). A Virtual Character's Agency Affects Social Responses in Immersive Virtual Reality: A Systematic Review and Meta-Analysis. *International Journal of Human–Computer Interaction*, *40*(16), 4167–4182. https://doi.org/10.1080/10447318.2023.2209979

Ferrer-Garcia M., Porras-Garcia B., Gonzalez-Ibanez C., Gracia-Blanes M., Vilalta-Abella F., Pla-Sanjuanelo J., Riva G., Dakanalis A., Achotegui-Loizate J., Talarn-Caparros A., Ribas-Sabate J., Andreu-Gracia A., Diaz-Marsa M., Monras-Arnau M., Serrano-Troncoso E., Treasure J., & Gutierrez-Maldonado J. (2017). Does owning a 'fatter' virtual body increase body anxiety in college students? *Annual Review of CyberTherapy and Telemedicine*, *15*, 147–153.

Ferrer-Garcia, M., Porras-Garcia, B., Moreno, M., Bertomeu, P., & Maldonado, J. G. (2018). Embodiment in different size virtual bodies produces changes in women's body image distortion and dissatisfaction. *Annual Review of CyberTherapy and Telemedicine*, *16*(Human Factors Engineering [4010]), 111–117.

Fett, A.-K. J., Shergill, S. S., Joyce, D. W., Riedl, A., Strobel, M., Gromann, P. M., & Krabbendam, L. (2012). To trust or not to trust: The dynamics of social interaction in psychosis. *Brain*, *135*(3), 976–984. https://doi.org/10.1093/brain/awr359

Finset, A., Stensrud, T. L., Holt, E., Verheul, W., & Bensing, J. (2011). Electrodermal activity in response to empathic statements in clinical interviews with fibromyalgia patients. *Patient Education and Counseling*, *82*(3), 355–360. Scopus. https://doi.org/10.1016/j.pec.2010.12.029

Fisher, S., Abdullah, A., Charvin, I., Da Fonseca, D., & Bat-Pitault, F. (2020). Comparison of body image evaluation by virtual reality and paper-based figure rating scales in adolescents with anorexia nervosa: Retrospective study. *Eating and Weight Disorders: EWD*, *25*(3), 735–743. https://doi.org/10.1007/s40519-019-00680-1

Flavián, C., Ibáñez-Sánchez, S., & Orús, C. (2021). The influence of scent on virtual reality experiences: The role of aroma-content congruence. *Journal of Business Research*, *123*, 289–301. https://doi.org/10.1016/j.jbusres.2020.09.036

Fornells-Ambrojo, M., Elenbaas, M., Barker, C., Swapp, D., Navarro, X., Rovira, A., Sanahuja, J. M. T., & Slater, M. (2016). Hypersensitivity to Contingent Behavior in Paranoia: A New Virtual Reality Paradigm. *The Journal of Nervous and Mental Disease*, *204*(2), 148–152. https://doi.org/10.1097/NMD.0000000000000414

Fornells-Ambrojo, M., Freeman, D., Slater, M., Swapp, D., Antley, A., & Barker, C. (2015). How Do People with Persecutory Delusions Evaluate Threat in a Controlled Social Environment? A Qualitative Study Using Virtual Reality. *Behavioural and Cognitive Psychotherapy*, *43*(1), 89–107. https://doi.org/10.1017/S1352465813000830

Freeman, D., Bradley, J., Antley, A., Bourke, E., DeWeever, N., Evans, N., Černis, E., Sheaves, B., Waite, F., Dunn, G., Slater, M., & Clark, D. M. (2016). Virtual reality in the treatment of persecutory delusions: Randomised controlled experimental study testing how to reduce delusional conviction. *British Journal of Psychiatry*, *209*(1), 62–67. https://doi.org/10.1192/bjp.bp.115.176438

Freeman, D., Dunn, G., Murray, R. M., Evans, N., Lister, R., Antley, A., Slater, M., Godlewska, B., Cornish, R., Williams, J., Di Simplicio, M., Igoumenou, A., Brenneisen, R., Tunbridge, E. M., Harrison, P. J., Harmer, C. J., Cowen, P., & Morrison, P. D. (2015). How Cannabis Causes Paranoia: Using the Intravenous Administration of Δ 9 -Tetrahydrocannabinol (THC) to Identify Key Cognitive Mechanisms Leading to Paranoia. *Schizophrenia Bulletin*, *41*(2), 391–399. https://doi.org/10.1093/schbul/sbu098

Freeman, D., Freeman, J., Ahmed, M., Haynes, P., Beckwith, H., Rovira, A., Miguel, A. L., Ward, R., Bousfield, M., Riffiod, L., Kabir, T., Waite, F., & Rosebrock, L. (2024). Automated VR therapy for improving positive self-beliefs and psychological well-being in young patients with psychosis: A proof of concept evaluation of Phoenix VR self-confidence therapy. *Behavioural and Cognitive Psychotherapy*, *52*(3), 277–287. https://doi.org/10.1017/S1352465823000553

Freeman, D., Garety, P. A., Bebbington, P. E., Smith, B., Rollinson, R., Fowler, D., Kuipers, E., Ray, K., & Dunn, G. (2005). Psychological investigation of the structure of paranoia in a non-clinical population. *The British Journal of Psychiatry*, *186*(5), 427–435. https://doi.org/10.1192/bjp.186.5.427

Freeman, D., Garety, P. A., Bebbington, P., Slater, M., Kuipers, E., Fowler, D., Green, C., Jordan, J., Ray, K., & Dunn, G. (2005). The psychology of persecutory ideation II: A virtual reality experimental study. *The Journal of Nervous and Mental Disease*, *193*(5), 309–315. https://doi.org/10.1097/01.nmd.0000161686.53245.70

Freeman, D., Garety, P. A., Kuipers, E., Fowler, D., Bebbington, P. E., & Dunn, G. (2007). Acting on persecutory delusions: The importance of safety seeking. *Behaviour Research and Therapy*, *45*(1), 89–99. https://doi.org/10.1016/j.brat.2006.01.014

Freeman, D., Haselton, P., Freeman, J., Spanlang, B., Kishore, S., Albery, E., Denne, M., Brown, P., Slater, M., & Nickless, A. (2018a). Automated psychological therapy using immersive virtual reality for treatment of fear of heights: A single-blind, parallel-group, randomised controlled trial. *The Lancet Psychiatry*, *5*(8), 625–632. https://doi.org/10.1016/S2215-0366(18)30226-8

Freeman, D., Haselton, P., Freeman, J., Spanlang, B., Kishore, S., Albery, E., Denne, M., Brown, P., Slater, M., & Nickless, A. (2018b). Automated psychological therapy using immersive virtual reality for treatment of fear of heights: A single-blind, parallel-group, randomised controlled trial. *The Lancet. Psychiatry*, *5*(8), 625–632. https://doi.org/10.1016/S2215-0366(18)30226-8

Freeman, D., Haselton, P., Freeman, J., Spanlang, B., Kishore, S., Albery, E., Denne, M., Brown, P., Slater, M., & Nickless, A. (2018c). Automated psychological therapy using immersive virtual reality for treatment of fear of heights: A single-blind, parallel-group, randomised controlled trial. *The Lancet Psychiatry*, *5*(8), 625–632. https://doi.org/10.1016/S2215-0366(18)30226-8

Freeman, D., Lambe, S., Galal, U., Yu, L.-M., Kabir, T., Petit, A., Rosebrock, L., Dudley, R., Chapman, K., Morrison, A., O'Regan, E., Murphy, E., Aynsworth, C., Jones, J., Powling, R., Grabey, J., Rovira, A., Freeman, J., Clark, D. M., & Waite, F. (2022). Agoraphobic avoidance in patients with psychosis: Severity and response to automated VR therapy in a secondary analysis of a randomised controlled clinical trial. *Schizophrenia Research*, *250*, 50–59. https://doi.org/10.1016/j.schres.2022.10.008

Freeman, D., Lambe, S., Kabir, T., Petit, A., Rosebrock, L., Yu, L.-M., Dudley, R., Chapman, K., Morrison, A., O'Regan, E., Aynsworth, C., Jones, J., Murphy, E., Powling, R., Galal, U., Grabey, J., Rovira, A., Martin, J., Hollis, C., … West, J. (2022a). Automated virtual reality therapy to treat agoraphobic avoidance and distress in patients with psychosis (gameChange): A multicentre, parallel-group, single-blind, randomised, controlled trial in England with mediation and moderation analyses. *The Lancet Psychiatry*, *9*(5), 375–388. https://doi.org/10.1016/S2215-0366(22)00060-8

Freeman, D., Lambe, S., Kabir, T., Petit, A., Rosebrock, L., Yu, L.-M., Dudley, R., Chapman, K., Morrison, A., O'Regan, E., Aynsworth, C., Jones, J., Murphy, E., Powling, R., Galal, U., Grabey, J., Rovira, A., Martin, J., Hollis, C., … West, J. (2022b). Automated virtual reality therapy to treat agoraphobic avoidance and distress in patients with psychosis (gameChange): A multicentre, parallel-group, single-blind, randomised, controlled trial in England with mediation and moderation analyses. *The Lancet Psychiatry*, *9*(5), 375–388. https://doi.org/10.1016/S2215-0366(22)00060-8

Freeman, D., Lister, R., Waite, F., Galal, U., Yu, L.-M., Lambe, S., Beckley, A., Bold, E., Jenner, L., Diamond, R., Kirkham, M., Twivy, E., Causier, C., Carr, L., Saidel, S., Day, R., Beacco, A., Rovira, A., Ivins, A., … Rosebrock, L. (2023). *Automated virtual reality cognitive therapy versus virtual reality mental relaxation therapy for the treatment of persistent persecutory delusions in patients with psychosis (THRIVE): A parallel-group, single-blind, randomised controlled trial in England with mediation analyses. 10*.

Freeman, D., & Loe, B. S. (2023). Explaining paranoia: Cognitive and social processes in the occurrence of extreme mistrust. *BMJ Ment Health*, *26*(1). https://doi.org/10.1136/bmjment-2023-300880

Freeman, D., Loe, B. S., Kingdon, D., Startup, H., Molodynski, A., Rosebrock, L., Brown, P., Sheaves, B., Waite, F., & Bird, J. C. (2021). The revised Green *et al.* , Paranoid Thoughts Scale (R-GPTS): Psychometric properties, severity ranges, and clinical cut-offs. *Psychological Medicine*, *51*(2), 244–253. https://doi.org/10.1017/S0033291719003155

Freeman, D., Pugh, K., Antley, A., Slater, M., Bebbington, P., Gittins, M., Dunn, G., Kuipers, E., Fowler, D., & Garety, P. (2008). Virtual reality study of paranoid thinking in the general population. *The British Journal of Psychiatry*, *192*(4), 258–263. https://doi.org/10.1192/bjp.bp.107.044677

Freeman, D., Pugh, K., Green, C., Valmaggia, L., Dunn, G., & Garety, P. (2007). A Measure of State Persecutory Ideation for Experimental Studies. *Journal of Nervous & Mental Disease*, *195*(9), 781–784. https://doi.org/10.1097/NMD.0b013e318145a0a9

Freeman, D., Pugh, K., Vorontsova, N., Antley, A., & Slater, M. (2010). Testing the continuum of delusional beliefs: An experimental study using virtual reality. *Journal of Abnormal Psychology*, *119*(1), 83–92. https://doi.org/10.1037/a0017514

Freeman, D., Rosebrock, L., Waite, F., Loe, B. S., Kabir, T., Petit, A., Dudley, R., Chapman, K., Morrison, A., O'Regan, E., Aynsworth, C., Jones, J., Murphy, E., Powling, R., Peel, H., Walker, H., Byrne, R., Freeman, J., Rovira, A., … Lambe, S. (2023). Virtual reality (VR) therapy for patients with psychosis: Satisfaction and side effects. *Psychological Medicine*, *53*(10), 4373–4384. https://doi.org/10.1017/S0033291722001167

Freeman, D., Slater, M., Bebbington, P. E., Garety, P. A., Kuipers, E., Fowler, D., Met, A., Read, C. M., Jordan, J., & Vinayagamoorthy, V. (2003a). Can Virtual Reality be Used to Investigate Persecutory Ideation? *The Journal of Nervous and Mental Disease*, *191*(8), 509–514. https://doi.org/10.1097/01.nmd.0000082212.83842.fe

Freeman, D., Slater, M., Bebbington, P. E., Garety, P. A., Kuipers, E., Fowler, D., Met, A., Read, C. M., Jordan, J., & Vinayagamoorthy, V. (2003b). Can Virtual Reality be Used to Investigate Persecutory Ideation?: *The Journal of Nervous and Mental Disease*, *191*(8), 509–514. https://doi.org/10.1097/01.nmd.0000082212.83842.fe

Frith, C. (2009). Role of facial expressions in social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3453–3458. https://doi.org/10.1098/rstb.2009.0142

Fujihara, Y., Guo, K., & Liu, C. H. (2023). Relationship between types of anxiety and the ability to recognize facial expressions. *Acta Psychologica*, *241*, 104100. https://doi.org/10.1016/j.actpsy.2023.104100

Fusaro, M., Fanti, V., & Chakrabarti, B. (2023). Greater interpersonal distance in adults with autism. *Autism Research*, *16*(10), 2002–2007. https://doi.org/10.1002/aur.3013

Gamberini, L., Chittaro, L., Spagnolli, A., & Carlesso, C. (2015). Psychological response to an emergency in virtual reality: Effects of victim ethnicity and emergency type on helping behavior and navigation. *Computers in Human Behavior*, *48*, 104–113. https://doi.org/10.1016/j.chb.2015.01.040

Garety, P. A., Kuipers, E., Fowler, D., Freeman, D., & Bebbington, P. E. (2001). A cognitive model of the positive symptoms of psychosis. *Psychological Medicine, 31(2)*, 189–195. https://doi.org/10.1017/S0033291701003312

Geraets, C. N. W., Klein Tuente, S., Lestestuiver, B. P., van Beilen, M., Nijman, S. A., Marsman, J. B. C., & Veling, W. (2021). Virtual reality facial emotion recognition in social environments: An eye-tracking study. *Internet Interventions*, *25*, 100432. https://doi.org/10.1016/j.invent.2021.100432

Giguère, S., Potvin, S., Beaudoin, M., Dellazizzo, L., Giguère, C.-É., Furtos, A., Gilbert, K., Phraxayavong, K., & Dumais, A. (2023). Avatar Intervention for Cannabis Use Disorder in Individuals with Severe Mental Disorders: A Pilot Study. *Journal of Personalized Medicine*, *13*(5), Article 5. https://doi.org/10.3390/jpm13050766

Gillies, M. (2018). Creating Virtual Characters. *Proceedings of the 5th International Conference on Movement and Computing*, 1–8. https://doi.org/10.1145/3212721.3212835

Giron, J., Mladenovic, J., Hanrieder, M., Golland, Y., & Friedman, D. (2023). Physiological correlates of stress induced by virtual humans in a naturalistic virtual reality scenario. *2023 11th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, 1–8. https://doi.org/10.1109/ACIIW59127.2023.10388148

Gonzalez-Franco, M., Egan, Z., Peachey, M., Antley, A., Randhavane, T., Panda, P., Zhang, Y., Wang, C. Y., Reilly, D. F., Peck, T. C., Won, A. S., Steed, A., & Ofek, E. (2020). MoveBox: Democratizing MoCap for the Microsoft Rocketbox Avatar Library. *2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, 91–98. https://doi.org/10.1109/AIVR50618.2020.00026

Gonzalez-Franco, M., Ofek, E., Pan, Y., Antley, A., Steed, A., Spanlang, B., Maselli, A., Banakou, D., Pelechano, N., Orts-Escolano, S., Orvalho, V., Trutoiu, L., Wojcik, M., Sanchez-Vives, M. V., Bailenson, J., Slater, M., & Lanier, J. (2020). The Rocketbox Library and the Utility of Freely Available Rigged Avatars. *Frontiers in Virtual Reality*, *1*. https://doi.org/10.3389/frvir.2020.561558

Gorisse, G., Senel, G., Banakou, D., Beacco, A., Oliva, R., Freeman, D., & Slater, M. (2021). Self-observation of a virtual body-double engaged in social interaction reduces persecutory thoughts. *Scientific Reports*, *11*(1), 23923. https://doi.org/10.1038/s41598-021-03373-x

Green, C. E. L., Freeman, D., Kuipers, E., Bebbington, P., Fowler, D., Dunn, G., & Garety, P. A. (2011). Paranoid Explanations of Experience: A Novel Experimental Study. *Behavioural and Cognitive Psychotherapy*, *39*(1), 21–34. https://doi.org/10.1017/S1352465810000457

Guillén, V., Baños, R. M., & Botella, C. (2018). Users' Opinion About a Virtual Reality System as an Adjunct to Psychological Treatment for Stress-Related Disorders: A Quantitative and Qualitative Mixed-Methods Study. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01038

Guldager, J. D., Kjær, S. L., Grittner, U., & Stock, C. (2022). Efficacy of the Virtual Reality Intervention VR FestLab on Alcohol Refusal Self-Efficacy: A Cluster-Randomized Controlled Trial. *International Journal of Environmental Research and Public Health*, *19*(6), 3293. https://doi.org/10.3390/ijerph19063293

Gupta, K., Zhang, Y., Gunasekaran, T. S., Krishna, N., Pai, Y. S., & Billinghurst, M. (2024). CAEVR: Biosignals-Driven Context-Aware Empathy in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, 1–11. https://doi.org/10.1109/TVCG.2024.3372130

Gupta, T., Osborne, K. J., Nadig, A., Haase, C. M., & Mittal, V. A. (2023). Alterations in facial expressions in individuals at risk for psychosis: A facial electromyography approach using emotionally evocative film clips. *Psychological Medicine*, *53*(12), 5829–5838. https://doi.org/10.1017/S0033291722003087

Gürerk, Ö., Bönsch, A., Kittsteiner, T., & Staffeldt, A. (2019). Virtual humans as co-workers: A novel methodology to study peer effects. Journal of Behavioral and Experimental Economics, 78, 17–29. https://doi.org/10.1016/j.socec.2018.11.003

Günther, V., Kropidlowski, A., Schmidt, F. M., Koelkebeck, K., Kersting, A., & Suslow, T. (2021). Attentional processes during emotional face perception in social anxiety disorder: A systematic review and meta-analysis of eye-tracking findings. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 111, 110353. https://doi.org/10.1016/j.pnpbp.2021.110353

Halbig, A., & Latoschik, M. E. (2021). A Systematic Review of Physiological Measurements, Factors, Methods, and Applications in Virtual Reality. *Frontiers in Virtual Reality*, *2*. https://doi.org/10.3389/frvir.2021.694567

Hale, J., & Hamilton, A. F. D. C. (2016). Testing the relationship between mimicry, trust and rapport in virtual reality conversations. *Scientific Reports*, *6*(1), Article 1. https://doi.org/10.1038/srep35295

Halim, I., Stemmet, L., Hach, S., Porter, R., Liang, H.-N., Vaezipour, A., Henry, J. D., & Baghaei, N. (2023). Individualized Virtual Reality for Increasing Self-Compassion: Evaluation Study. *JMIR Mental Health*, *10*(1), e47617. https://doi.org/10.2196/47617

Han, K., Shin, J., Yoon, S. Y., Jang, D.-P., & Kim, J.-J. (2014). Deficient gaze pattern during virtual multiparty conversation in patients with schizophrenia. *Computers in Biology and Medicine*, *49*, 60–66. https://doi.org/10.1016/j.compbiomed.2014.03.012

Hariton, E., & Locascio, J. J. (2018). Randomised controlled trials—The gold standard for effectiveness research. *BJOG : An International Journal of Obstetrics and Gynaecology*, *125*(13), 1716. https://doi.org/10.1111/1471-0528.15199

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), Article 7825. https://doi.org/10.1038/s41586-020-2649-2

Hartanto, D., Kampmann, I. L., Morina, N., Emmelkamp, P. G. M., Neerincx, M. A., & Brinkman, W.-P. (2014). Controlling Social Stress in Virtual Reality Environments. *PLOS ONE*, *9*(3), e92804. https://doi.org/10.1371/journal.pone.0092804

Hidding, M., Veling, W., Pijnenborg, G. H. M., & van der Stouwe, E. C. D. (2024). A single-session VR intervention addressing self-compassion and self-criticism with and without perspective change: Results of a randomized controlled experiment. *Behaviour Research and Therapy*, *173*, 104466. https://doi.org/10.1016/j.brat.2023.104466

Holmqvist, K., Nyström, M., & Mulvey, F. (2012). Eye tracker data quality: What it is and how to measure it. *Proceedings of the Symposium on Eye Tracking Research and Applications*, 45–52. https://doi.org/10.1145/2168556.2168563

Horigome, T., Yoshida, S., Tanikawa, T., Mimura, M., & Kishimoto, T. (2023). Modification of the therapist's facial expressions using virtual reality technology during the treatment of social anxiety disorder: A case series. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1030050

Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology*, *38*(2), 139–149. https://doi.org/10.1037/0022-0167.38.2.139

Inoue, M., Irino, T., Furuyama, N., & Hanada, R. (2021). Observational and Accelerometer Analysis of Head Movement Patterns in Psychotherapeutic Dialogue. *Sensors*, *21*, 3162. https://doi.org/10.3390/s21093162

Jack, R. E., & Schyns, P. G. (2015). The Human Face as a Dynamic Tool for Social Communication. *Current Biology*, *25*(14), R621–R634. https://doi.org/10.1016/j.cub.2015.05.052

Jain, D., Jain, U., Rao, K., Dhingra, D., Nagrath, P., Varshney, A. R., & Nagrath, P. (2024). A Real-Time Crowd Tracking and Control System Using Deep Learning. In A. E. Hassanien, S. Anand, A. Jaiswal, & P. Kumar (Eds.), *Innovative Computing and Communications* (pp. 587–601). Springer Nature. https://doi.org/10.1007/978-981-97-3817-5_42

Jebb, A. T., Tay, L., Wang, W., & Huang, Q. (2015). Time series analysis for psychological research: Examining and forecasting change. *Frontiers in Psychology*, *6*, 727. https://doi.org/10.3389/fpsyg.2015.00727

Jeong, H. S., Lee, J. H., Kim, H. E., & Kim, J.-J. (2021). Appropriate Number of Treatment Sessions in Virtual Reality-Based Individual Cognitive Behavioral Therapy for Social Anxiety Disorder. *Journal of Clinical Medicine*, *10*(5), 915. https://doi.org/10.3390/jcm10050915

Jerome, L., Feduccia, A. A., Wang, J. B., Hamilton, S., Yazar-Klosinski, B., Emerson, A., Mithoefer, M. C., & Doblin, R. (2020). Long-term follow-up outcomes of MDMA-assisted psychotherapy for treatment of PTSD: A longitudinal pooled analysis of six phase 2 trials. *Psychopharmacology*, *237*(8), 2485–2497. https://doi.org/10.1007/s00213-020-05548-2

Jorjafki, E. M., Sagarin, B. J., & Butail, S. (2018). Drawing power of virtual crowds. *Journal of The Royal Society Interface*, *15*(145), 20180335. https://doi.org/10.1098/rsif.2018.0335

Kaminska, O. K., Magnuski, M., Olszanowski, M., Gola, M., Brzezicka, A., & Winkielman, P. (2020). Ambiguous at the second sight: Mixed facial expressions trigger late electrophysiological responses linked to lower social impressions. *Cognitive, Affective, & Behavioral Neuroscience*, *20*(2), 441–454. https://doi.org/10.3758/s13415-020-00778-5

Kane, H. S., McCall, C., Collins, N. L., & Blascovich, J. (2012). Mere presence is not enough: Responsive support in a virtual world. *Journal of Experimental Social Psychology*, *48*(1), 37–44. https://doi.org/10.1016/j.jesp.2011.07.001

Kang, W., Kim, G., Kim, H., & Lee, S.-H. (2019). The Influence of Anxiety on the Recognition of Facial Emotion Depends on the Emotion Category and Race of the Target Faces. *Experimental Neurobiology*, *28*(2), 261–269. https://doi.org/10.5607/en.2019.28.2.261

Kao, D., Ratan, R., Mousas, C., & Magana, A. J. (2021). The Effects of a Self-Similar Avatar Voice in Educational Games. *Proc. ACM Hum.-Comput. Interact.*, *5*(CHI PLAY), 238:1-238:28. https://doi.org/10.1145/3474665

Kappas, A., Krumhuber, E., & Küster, D. (2013). 6 Facial behavior. In J. A. Hall & M. L. Knapp (Eds.), *Nonverbal Communication* (pp. 131–166). De Gruyter Mouton. https://doi.org/10.1515/9783110238150.131

Kilteni, K., Groten, R., & Slater, M. (2012). The Sense of Embodiment in Virtual Reality. *Presence: Teleoperators and Virtual Environments*, *21*(4), 373–387. https://doi.org/10.1162/PRES_a_00124

Kim, J., Jung, Y. H., Shin, Y.-B., Kim, M.-K., Eom, H., Kim, E., Kim, J., & Kim, J.-J. (2020). Development and Validation of a Virtual Reality-Based Training Program for Promoting Subjective Well-Being. *Psychiatry Investigation*, *17*(12), 1207–1215. https://doi.org/10.30773/pi.2020.0311

Kimmel, S., Jung, F., Matviienko, A., Heuten, W., & Boll, S. (2023). Let's Face It: Influence of Facial Expressions on Social Presence in Collaborative Virtual Reality. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–16. https://doi.org/10.1145/3544548.3580707

Kleinsmith, A., & Bianchi-Berthouze, N. (2013). Affective Body Expression Perception and Recognition: A Survey. *IEEE Transactions on Affective Computing*, *4*(1), 15–33. IEEE Transactions on Affective Computing. https://doi.org/10.1109/T-AFFC.2012.16

Knight, I., West, J., Matthews, E., Kabir, T., Lambe, S., Waite, F., & Freeman, D. (2021). Participatory design to create a VR therapy for psychosis. *Design for Health*, *5*(1), 98–119. https://doi.org/10.1080/24735132.2021.1885889

Kocur, M., Habler, F., Schwind, V., Woźniak, P. W., Wolff, C., & Henze, N. (2021). Physiological and Perceptual Responses to Athletic Avatars while Cycling in Virtual Reality. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–18. https://doi.org/10.1145/3411764.3445160

Kothgassner, O. D., Goreis, A., Glenk, L. M., Kafka, J. X., Beutl, L., Kryspin-Exner, I., Hlavacs, H., Palme, R., & Felnhofer, A. (2021). Virtual and real-life ostracism and its impact on a subsequent acute stressor. *Physiology & Behavior*, *228*, 113205. https://doi.org/10.1016/j.physbeh.2020.113205

Kothgassner, O. D., Goreis, A., Kafka, J. X., Kaufmann, M., Atteneder, K., Beutl, L., Hennig-Fast, K., Hlavacs, H., & Felnhofer, A. (2019). Virtual social support buffers stress response: An experimental comparison of real-life and virtual support prior to a social stressor. *Journal of Behavior Therapy and Experimental Psychiatry*, *63*, 57–65. https://doi.org/10.1016/j.jbtep.2018.11.003

Kroczek, L. O. H., Pfaller, M., Lange, B., Müller, M., & Mühlberger, A. (2020). Interpersonal Distance During Real-Time Social Interaction: Insights From Subjective Experience, Behavior, and Physiology. *Frontiers in Psychiatry*, *11*. https://www.frontiersin.org/articles/10.3389/fpsyt.2020.00561

Krumhuber, E. G., Skora, L. I., Hill, H. C. H., & Lander, K. (2023). The role of facial movements in emotion recognition. *Nature Reviews Psychology*, *2*(5), Article 5. https://doi.org/10.1038/s44159-023-00172-1

Kullmann, P., Menzel, T., Botsch, M., & Latoschik, M. E. (2023). An Evaluation of Other-Avatar Facial Animation Methods for Social VR. *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–7. https://doi.org/10.1145/3544549.3585617

Kwon, J. H., Powell, J., & Chalmers, A. (2013). How level of realism influences anxiety in virtual reality environments for a job interview. *International Journal of Human-Computer Studies*, *71*(10), 978–987. https://doi.org/10.1016/j.ijhcs.2013.07.003

Lambe, S., Knight, I., Kabir, T., West, J., Patel, R., Lister, R., Rosebrock, L., Rovira, A., Garnish, B., Freeman, J., M. Clark, D., Waite, F., & Freeman, D. (2020).

Developing an automated VR cognitive treatment for psychosis: gameChange VR therapy. *Journal of Behavioral and Cognitive Therapy*, *30*(1), 33–40. https://doi.org/10.1016/j.jbct.2019.12.001

Låver, J., McAleavey, A., Valaker, I., Castonguay, L. G., & Moltu, C. (2024). Therapists' and patients' experiences of using patients' self-reported data in ongoing psychotherapy processes—A systematic review and meta-analysis of qualitative studies. *Psychotherapy Research*, *34*(3), 293–310. https://doi.org/10.1080/10503307.2023.2222896

Lawson, A. P., & Mayer, R. E. (2022). Does the emotional stance of human and virtual instructors in instructional videos affect learning processes and outcomes? *Contemporary Educational Psychology*, *70*, 102080. https://doi.org/10.1016/j.cedpsych.2022.102080

Lee, H.-S., Hong, S.-J. J., Baxter, T., Scott, J., Shenoy, S., Buck, L., Bodenheimer, B., & Park, S. (2021). Altered Peripersonal Space and the Bodily Self in Schizophrenia: A Virtual Reality Study. *Schizophrenia Bulletin*, *47*(4), 927–937. https://doi.org/10.1093/schbul/sbab024

Lee, J., Lim, Y., Graham, S. J., Kim, G., Wiederhold, B. K., Wiederhold, M. D., Kim, I. Y., & Kim, S. I. (2004). Nicotine craving and cue exposure therapy by using virtual environments. *Cyberpsychology & Behavior : The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society*, *7*(6), 705–713. https://doi.org/10.1089/cpb.2004.7.705

Leonardis, D., Frisoli, A., Barsotti, M., Carrozzino, M., & Bergamasco, M. (2014). Multisensory Feedback Can Enhance Embodiment Within an Enriched Virtual Walking Scenario. *Presence: Teleoperators and Virtual Environments*, *23*(3), 253–266. https://doi.org/10.1162/PRES_a_00190

Liew, T. W., Mat Zin, N. A., Ashaari, N., & Tan, S.-M. (2016). The Effects of a Pedagogical Agent's Smiling Expression on the Learner's Emotions and Motivation in a Virtual Learning Environment. *The International Review of Research in Open and Distributed Learning*, *17*, 248–266. https://doi.org/10.19173/irrodl.v17i5.2350

Lindner, P., Miloff, A., Fagernäs, S., Andersen, J., Sigeman, M., Andersson, G., Furmark, T., & Carlbring, P. (2019). Therapist-led and self-led one-session virtual reality exposure therapy for public speaking anxiety with consumer hardware and software: A randomized controlled trial. *Journal of Anxiety Disorders*, *61*, 45–54. https://doi.org/10.1016/j.janxdis.2018.07.003

Lisetti, C., Amini, R., Yasavur, U., & Rishe, N. (2013). I Can Help You Change! An Empathic Virtual Agent Delivers Behavior Change Health Interventions. *ACM Transactions on Management Information Systems*, *4*(4), 1–28. https://doi.org/10.1145/2544103

Liu, F., Lijn, F. van der, Schurmann, C., Zhu, G., Chakravarty, M. M., Hysi, P. G., Wollstein, A., Lao, O., Bruijne, M. de, Ikram, M. A., Lugt, A. van der, Rivadeneira, F., Uitterlinden, A. G., Hofman, A., Niessen, W. J., Homuth, G., Zubicaray, G. de, McMahon, K. L., Thompson, P. M., … Kayser, M. (2012). A Genome-Wide Association Study Identifies Five Loci Influencing Facial Morphology in Europeans. *PLOS Genetics*, *8*(9), e1002932. https://doi.org/10.1371/journal.pgen.1002932

Llanes-Jurado, J., Gómez-Zaragozá, L., Minissi, M. E., Alcañiz, M., & Marín-Morales, J. (2024). Developing conversational Virtual Humans for social emotion elicitation based on large language models. *Expert Systems with Applications*, *246*, 123261. https://doi.org/10.1016/j.eswa.2024.123261

Llanes-Jurado, J., Marín-Morales, J., Guixeres, J., & Alcañiz, M. (2020). Development and Calibration of an Eye-Tracking Fixation Identification Algorithm for Immersive Virtual Reality. *Sensors*, *20*(17), Article 17. https://doi.org/10.3390/s20174956

Loth, E., Garrido, L., Ahmad, J., Watson, E., Duff, A., & Duchaine, B. (2018). Facial expression recognition as a candidate marker for autism spectrum disorder: How frequent and severe are deficits? *Molecular Autism*, *9*(1), 7. https://doi.org/10.1186/s13229-018-0187-7

Lundin, R. M., Yeap, Y., & Menkes, D. B. (2023). Adverse Effects of Virtual and Augmented Reality Interventions in Psychiatry: Systematic Review. *JMIR Mental Health*, *10*, e43240. https://doi.org/10.2196/43240

Luo, L., Weng, D., Ding, N., Hao, J., & Tu, Z. (2022). The effect of avatar facial expressions on trust building in social virtual reality. *The Visual Computer*. https://doi.org/10.1007/s00371-022-02700-1

Luong, T., & Holz, C. (2022). Characterizing Physiological Responses to Fear, Frustration, and Insight in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, *28*(11), 3917–3927. IEEE Transactions on Visualization and Computer Graphics. https://doi.org/10.1109/TVCG.2022.3203113

Ma, F., & Pan, X. (2022). Visual Fidelity Effects on Expressive Self-avatar in Virtual Reality: First Impressions Matter. *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 57–65. https://doi.org/10.1109/VR51125.2022.00023

Maestri, G. (2006). Facial and Dialogue Animation. In *Digital Character Animation 3* (pp. 196–219). New Riders.

Magnenat-Thalmann, N., & Thalmann, D. (2005). Virtual humans: Thirty years of research, what next? *The Visual Computer*, *21*(12), 997–1015. https://doi.org/10.1007/s00371-005-0363-6

Makowski, D., Pham, T., Lau, Z. J., Brammer, J. C., Lespinasse, F., Pham, H., Schölzel, C., & Chen, S. H. A. (2021). NeuroKit2: A Python toolbox for neurophysiological signal processing. *Behavior Research Methods*, *53*(4), 1689–1696. https://doi.org/10.3758/s13428-020-01516-y

Maples-Keller, J. L., Rauch, S. A. M., Jovanovic, T., Yasinski, C. W., Goodnight, J. M., Sherrill, A., Black, K., Michopoulos, V., Dunlop, B. W., Rothbaum, B. O., & Norrholm, S. D. (2019). Changes in trauma-potentiated startle, skin conductance, and heart rate within prolonged exposure therapy for PTSD in high and low treatment responders. *Journal of Anxiety Disorders*, *68*, 102147. https://doi.org/10.1016/j.janxdis.2019.102147

Marci, C. D., Moran, E. K., & Orr, S. P. (2004). Physiologic Evidence for the Interpersonal Role of Laughter During Psychotherapy. *The Journal of Nervous and Mental Disease*, *192*(10), 689. https://doi.org/10.1097/01.nmd.0000142032.04196.63

Marci, C., & Riess, H. (2005). The clinical relevance of psychophysiology: Support for the psychobiology of empathy and psychodynamic process. *American Journal of Psychotherapy*, *59*(3), 213–226. https://doi.org/10.1176/appi.psychotherapy.2005.59.3.213

Martens, M. A., Antley, A., Freeman, D., Slater, M., Harrison, P. J., & Tunbridge, E. M. (2019). It feels real: Physiological responses to a stressful virtual reality environment and its impact on working memory. *Journal of Psychopharmacology*, *33*(10), 1264–1273. https://doi.org/10.1177/0269881119860156

McCorry, L. K. (2007). Physiology of the Autonomic Nervous System. *American Journal of Pharmaceutical Education*, *71*(4), 78.

McKinney, W. (2010). *Data Structures for Statistical Computing in Python*. 56–61. https://doi.org/10.25080/Majora-92bf1922-00a

Mendoza-Medialdea, M. T., Meschberger-Annweiler, F.-A., Ascione, M., Rueda-Pina, A., Rabarbari, E., Porras-Garcia, B., Ferrer-Garcia, M., & Gutiérrez-Maldonado, J. (2023). Body Dissatisfaction and Body-Related Attentional Bias: Is There a Causal Relationship? *Journal of Clinical Medicine*, *12*(17), Article 17. https://doi.org/10.3390/jcm12175659

Mertens, G., Wagensveld, P., & Engelhard, I. M. (2019). Cue conditioning using a virtual spider discriminates between high and low spider fearful individuals. *Computers in Human Behavior*, *91*, 192–200. https://doi.org/10.1016/j.chb.2018.10.006

Miloff, A., Carlbring, P., Hamilton, W., Andersson, G., Reuterskiöld, L., & Lindner, P. (2020a). Measuring Alliance Toward Embodied Virtual Therapists in the Era of Automated Treatments With the Virtual Therapist Alliance Scale (VTAS): Development and Psychometric Evaluation. *Journal of Medical Internet Research*, *22*(3), e16660. https://doi.org/10.2196/16660

Miloff, A., Carlbring, P., Hamilton, W., Andersson, G., Reuterskiöld, L., & Lindner, P. (2020b). Measuring Alliance Toward Embodied Virtual Therapists in the Era of Automated Treatments With the Virtual Therapist Alliance Scale (VTAS): Development and Psychometric Evaluation. *Journal of Medical Internet Research*, *22*(3), e16660. https://doi.org/10.2196/16660

Mitruț, O., Moldoveanu, A., Petrescu, L., Petrescu, C., & Moldoveanu, F. (2021). A Review of Virtual Therapists in Anxiety and Phobias Alleviating Applications. In J. Y. C. Chen & G. Fragomeni (Eds.), *Virtual, Augmented and Mixed Reality* (pp. 71–79). Springer International Publishing. https://doi.org/10.1007/978-3-030-77599-5_6

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine*, *19*(2), 98–100. IEEE Robotics & Automation Magazine. https://doi.org/10.1109/MRA.2012.2192811

Mountford, V. A., Tchanturia, K., & Valmaggia, L. (2016). 'What Are You Thinking When You Look at Me?' A Pilot Study of the Use of Virtual Reality in Body Image. *Cyberpsychology, Behavior and Social Networking*, *19*(2), 93–99. https://doi.org/10.1089/cyber.2015.0169

Mousas, C., Anastasiou, D., & Spantidi, O. (2018). The effects of appearance and motion of virtual characters on emotional reactivity. *Computers in Human Behavior*, *86*, 99–108. https://doi.org/10.1016/j.chb.2018.04.036

Mühlberger, A., Herrmann, M. J., Wiedemann, G., Ellgring, H., & Pauli, P. (2001). Repeated exposure of flight phobics to flights in virtual reality. *Behaviour Research and Therapy*, *39*(9), 1033–1050. https://doi.org/10.1016/S0005-7967(00)00076-0

Nason, G. P., & Silverman, B. W. (1995). The Stationary Wavelet Transform and some Statistical Applications. In A. Antoniadis & G. Oppenheim (Eds.), *Wavelets and Statistics* (pp. 281–299). Springer. https://doi.org/10.1007/978-1-4612-2544-7_17

Natali, L., Meregalli, V., Rowlands, K., Di Pietro, J., Treasure, J., Collantoni, E., Meneguzzo, P., Tenconi, E., Favaro, A., Fontana, F., Ceccato, E., Sala, A., Valmaggia, L., & Cardi, V. (2024). Virtual food exposure with positive mood induction or social support to reduce food anxiety in anorexia nervosa: A feasibility study. *International Journal of Eating Disorders*, *57*(3), 703–715. https://doi.org/10.1002/eat.24155

Nijman, S. A., Veling, W., Greaves-Lord, K., Vos, M., Zandee, C. E. R., Aan het Rot, M., Geraets, C. N. W., & Pijnenborg, G. H. M. (2020). Dynamic Interactive Social Cognition Training in Virtual Reality (DiSCoVR) for People With a Psychotic Disorder: Single-Group Feasibility and Acceptability Study. *JMIR Mental Health*, *7*(8). https://doi.org/10.2196/17808

Nowak, K., & Biocca, F. (2003). The Effect of the Agency and Anthropomorphism on Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments. *Presence Teleoperators & Virtual Environments*, *12*, 481–494. https://doi.org/10.1162/105474603322761289

Oh Kruzic, C., Kruzic, D., Herrera, F., & Bailenson, J. (2020). Facial expressions contribute more than body movements to conversational outcomes in avatar-mediated virtual environments. *Scientific Reports*, *10*(1), 20626. https://doi.org/10.1038/s41598-020-76672-4

Oh, S. Y., Bailenson, J., Krämer, N., & Li, B. (2016). Let the Avatar Brighten Your Smile: Effects of Enhancing Facial Expressions in Virtual Environments. *PLOS ONE*, *11*(9), e0161794. https://doi.org/10.1371/journal.pone.0161794

Osugi, T., & Kawahara, J. I. (2018). Effects of Head Nodding and Shaking Motions on Perceptions of Likeability and Approachability. *Perception*, *47*(1), 16–29. https://doi.org/10.1177/0301006617733209

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ*, *372*, n160. https://doi.org/10.1136/bmj.n160

Pan, X., Gillies, M., & Slater, M. (2015). Virtual Character Personality Influences Participant Attitudes and Behavior – An Interview with a Virtual Human Character about Her Social Anxiety. *Frontiers in Robotics and AI*, *2*. https://doi.org/10.3389/frobt.2015.00001

Pan, X., & Hamilton, A. F. de C. (2018a). Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology*, *109*(3), 395–417. https://doi.org/10.1111/bjop.12290

Pan, X., & Hamilton, A. F. de C. (2018b). Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology*, *109*(3), Article 3. https://doi.org/10.1111/bjop.12290

Park, I. H., Kim, J.-J., Ku, J., Jang, H. J., Park, S.-H., Kim, C.-H., Kim, I. Y., & Kim, S. I. (2009). Characteristics of social anxiety from virtual interpersonal interactions in patients with schizophrenia. *Psychiatry*, *72*(1), 79–93. https://doi.org/10.1521/psyc.2009.72.1.79

Parmar, D., Olafsson, S., Utami, D., Murali, P., & Bickmore, T. (2022). Designing empathic virtual agents: Manipulating animation, voice, rendering, and empathy to create persuasive agents. *Autonomous Agents and Multi-Agent Systems*, *36*(1). https://doi.org/10.1007/s10458-021-09539-1

Perpiñá, C., Botella, C., Baños, R., Marco, H., Alcañiz, M., & Quero, S. (1999). Body Image and Virtual Reality in Eating Disorders: Is Exposure to Virtual Reality More Effective than the Classical Body Image Treatment? *CyberPsychology & Behavior*, *2*(2), 149–155. https://doi.org/10.1089/cpb.1999.2.149

Pertaub, D.-P., Slater, M., & Barker, C. (2002). An Experiment on Public Speaking Anxiety in Response to Three Different Types of Virtual Audience. *Presence: Teleoperators and Virtual Environments*, *11*(1), 68–78. https://doi.org/10.1162/105474602317343668

Peterson, S. M., Furuichi, E., & Ferris, D. P. (2018). Effects of virtual reality high heights exposure during beam-walking on physiological stress and cognitive loading. *PLoS ONE*, *13*(7), e0200306. https://doi.org/10.1371/journal.pone.0200306

Ponton, J. L., Yun, H., Andujar, C., & Pelechano, N. (2022). Combining Motion Matching and Orientation Prediction to Animate Avatars for Consumer-Grade VR

Devices. *Computer Graphics Forum*, *41*(8), 107–118. https://doi.org/10.1111/cgf.14628

Porras Garcia, B., Ferrer Garcia, M., Olszewska, A., Yilmaz, L., González Ibañez, C., Gracia Blanes, M., Gültekin, G., Serrano Troncoso, E., & Gutiérrez Maldonado, J. (2019). Is This My Own Body? Changing the Perceptual and Affective Body Image Experience among College Students Using a New Virtual Reality Embodiment-Based Technique. *Journal of Clinical Medicine*, *8*(7). https://doi.org/10.3390/jcm8070925

Porras-Garcia, B., Ferrer-Garcia, M., Ghita, A., Moreno, M., López-Jiménez, L., Vallvé-Romeu, A., Serrano-Troncoso, E., & Gutiérrez-Maldonado, J. (2019). The influence of gender and body dissatisfaction on body-related attentional bias: An eye-tracking and virtual reality study. *International Journal of Eating Disorders*, *52*(10), 1181–1190. https://doi.org/10.1002/eat.23136

Porras-Garcia, B., Ferrer-Garcia, M., Yilmaz, L., Sen, Y. O., Olszewska, A., Ghita, A., Serrano-Troncoso, E., Treasure, J., & Gutiérrez-Maldonado, J. (2020). Body-related attentional bias as mediator of the relationship between body mass index and body dissatisfaction. *European Eating Disorders Review : The Journal of the Eating Disorders Association*, *28*(4), 454–464. https://doi.org/10.1002/erv.2730

Porras-Garcia, B., Ghita, A., Moreno, M., Ferrer-Garcia, M., Bertomeu-Panisello, P., Serrano-Troncoso, E., Riva, G., Dakanalis, A., Achotegui-Loizate, J., Talarn-Caparros, A., Andreu-Gracia, A., Treasure, J., Martinez-Mallen, E., Moreno-Perez, E., & Gutierrez-Maldonado, J. (2018). Gender differences in attentional bias after owning a virtual avatar with increased weight. *Annual Review of CyberTherapy and Telemedicine*, *16*(Human Factors Engineering [4010]), 73–79.

Post, R., Haberman, J., Iwaki, L., & Whitney, D. (2012). The Frozen Face Effect: Why Static Photographs May Not Do You Justice. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00022

Pot-Kolder, R. M. C. A., Geraets, C. N. W., Veling, W., Beilen, M. van, Staring, A. B. P., Gijsman, H. J., Delespaul, P. A. E. G., & Gaag, M. van der. (2018a). Virtual-reality-based cognitive behavioural therapy versus waiting list control for paranoid ideation and social avoidance in patients with psychotic disorders: A single-blind randomised controlled trial. *The Lancet Psychiatry*, *5*(3), 217–226. https://doi.org/10.1016/S2215-0366(18)30053-1

Pot-Kolder, R. M. C. A., Geraets, C. N. W., Veling, W., Beilen, M. van, Staring, A. B. P., Gijsman, H. J., Delespaul, P. A. E. G., & Gaag, M. van der. (2018b). Virtual-reality-based cognitive behavioural therapy versus waiting list control for paranoid ideation and social avoidance in patients with psychotic disorders: A single-blind randomised controlled trial. *The Lancet Psychiatry*, *5*(3), 217–226. https://doi.org/10.1016/S2215-0366(18)30053-1

Powers, M. B., Briceno, N. F., Gresham, R., Jouriles, E. N., Emmelkamp, P. M. G., & Smits, J. A. J. (2013). Do conversations with virtual avatars increase feelings of social anxiety? *Journal of Anxiety Disorders*, *27*(4), 398–403. https://doi.org/10.1016/j.janxdis.2013.03.003

Poyatos, F. (2010). Forms and Functions of Nonverbal Communication in the Novel: A New Perspective of the Author-Character-Reader Relationship. In A. Kendon (Ed.), *Nonverbal Communication, Interaction, and Gesture: Selections from SEMIOTICA* (pp. 107–150). De Gruyter Mouton. https://doi.org/10.1515/9783110880021.107

Provenzano, L., Porciello, G., Ciccarone, S., Lenggenhager, B., Tieri, G., Marucci, M., Dazzi, F., Loriedo, C., & Bufalari, I. (2020). Characterizing Body Image Distortion and Bodily Self-Plasticity in Anorexia Nervosa via Visuo-Tactile Stimulation in Virtual Reality. *Journal of Clinical Medicine*, *9*(1), Article 1. https://doi.org/10.3390/jcm9010098

Quintana, P., Nolet, K., Baus, O., & Bouchard, S. (2019). The Effect of Exposure to Fear-Related Body Odorants on Anxiety and Interpersonal Trust Toward a Virtual Character. *Chemical Senses*, *44*(9), 683–692. https://doi.org/10.1093/chemse/bjz063

Raffegeau, T. E., Fawver, B., Young, W. R., Williams, A. M., Lohse, K. R., & Fino, P. C. (2020). The direction of postural threat alters balance control when standing at virtual elevation. *Experimental Brain Research*, *238*(11), 2653–2663. https://doi.org/10.1007/s00221-020-05917-5

Rahman, M. A., Brown, D. J., Mahmud, M., Harris, M., Shopland, N., Heym, N., Sumich, A., Turabee, Z. B., Standen, B., Downes, D., Xing, Y., Thomas, C., Haddick, S., Premkumar, P., Nastase, S., Burton, A., & Lewis, J. (2023). Enhancing biofeedback-driven self-guided virtual reality exposure therapy through arousal detection from multimodal data using machine learning. *Brain Informatics*, *10*(1), 14. https://doi.org/10.1186/s40708-023-00193-9

Raimbaud, P., Jovane, A., Zibrek, K., Pacchierotti, C., Christie, M., Hoyet, L., Pettré, J., & Olivier, A.-H. (2022). The Stare-in-the-Crowd Effect in Virtual Reality. *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 281–290. https://doi.org/10.1109/VR51125.2022.00047

Ranjbartabar, H., Richards, D., Bilgin, A. A., & Kutay, C. (2021). First Impressions Count! The Role of the Human's Emotional State on Rapport Established with an Empathic versus Neutral Virtual Therapist. *IEEE Transactions on Affective Computing*, *12*(3), 788–800. IEEE Transactions on Affective Computing. https://doi.org/10.1109/TAFFC.2019.2899305

Reblin, M., Ellington, L., Latimer, S., Uchino, B. N., Roter, D., & Maxwell, A. (2012). Communication Style Affects Physiological Response in Simulated Cancer Risk Reduction Interactions. *Journal of Applied Biobehavioral Research*, *17*(2), 129–156. https://doi.org/10.1111/j.1751-9861.2012.00080.x

Recio, G., Sommer, W., & Schacht, A. (2011). Electrophysiological correlates of perceiving and evaluating static and dynamic facial emotional expressions. *Brain Research*, *1376*, 66–75. https://doi.org/10.1016/j.brainres.2010.12.041

Reichenberger, J., Pfaller, M., Forster, D., Gerczuk, J., Youssef Shiban, & Mühlberger, A. (2019). Men Scare Me More: Gender Differences in Social Fear

Conditioning in Virtual Reality. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.01617

Reichenberger, J., Pfaller, M., & Mühlberger, A. (2020). Gaze Behavior in Social Fear Conditioning: An Eye-Tracking Study in Virtual Reality. *Frontiers in Psychology*, *11*. https://www.frontiersin.org/articles/10.3389/fpsyg.2020.00035

Reichenberger, J., Wechsler, T. F., Diemer, J., Mühlberger, A., & Notzon, S. (2022). Fear, psychophysiological arousal, and cognitions during a virtual social skills training in social anxiety disorder while manipulating gaze duration. *Biological Psychology*, *175*, 108432. https://doi.org/10.1016/j.biopsycho.2022.108432

Ridout, S. J., Spofford, C. M., van't Wout-Frank, M., Philip, N. S., Unger, W. S., Carpenter, L. L., Tyrka, A. R., & Shea, M. T. (2017). Heart Rate Variability Responses to a Standardized Virtual Reality Exposure in Veterans with PTSD. *Current Treatment Options in Psychiatry*, *4*(3), 271–280. https://doi.org/10.1007/s40501-017-0118-9

Riva, G., & Melis, L. (1997). Virtual Reality for the Treatment of Body Image Disturbances. *Studies in Health Technology and Informatics*, 95–112.

Rizzo, A. S., Difede, J., Rothbaum, B. O., Reger, G., Spitalnick, J., Cukor, J., & McLay, R. (2010). Development and early evaluation of the Virtual Iraq/Afghanistan exposure therapy system for combat-related PTSD. *Annals of the New York Academy of Sciences*, *1208*, 114–125. https://doi.org/10.1111/j.1749-6632.2010.05755.x

Rizzo, A. 'Skip', & Shilling, R. (2017). Clinical Virtual Reality tools to advance the prevention, assessment, and treatment of PTSD. *European Journal of Psychotraumatology*, *8*(sup5). https://doi.org/10.1080/20008198.2017.1414560

Robitaille, N., Jackson, P. L., Hébert, L. J., Mercier, C., Bouyer, L. J., Fecteau, S., Richards, C. L., & McFadyen, B. J. (2017). A Virtual Reality avatar interaction (VRai) platform to assess residual executive dysfunction in active military personnel with previous mild traumatic brain injury: Proof of concept. *Disability and Rehabilitation. Assistive Technology*, *12*(7), 758–764. https://doi.org/10.1080/17483107.2016.1229048

Rotenberg, K. J., Sharp, C., & Venta, A. (2023). Trust Beliefs in Significant Others, Interpersonal Stress, and Internalizing Psychopathology of Adolescents with Psychiatric Disorders. *Child Psychiatry & Human Development*, *54*(2), 450–459. https://doi.org/10.1007/s10578-021-01255-x

Rothbaum, B. O., Price, M., Jovanovic, T., Norrholm, S. D., Gerardi, M., Dunlop, B., Davis, M., Bradley, B., Duncan, E. J., Rizzo, A., & Ressler, K. J. (2014). A Randomized, Double-Blind Evaluation of d-Cycloserine or Alprazolam Combined With Virtual Reality Exposure Therapy for Posttraumatic Stress Disorder in Iraq and Afghanistan War Veterans. *American Journal of Psychiatry*, *171*(6), 640–648. https://doi.org/10.1176/appi.ajp.2014.13121625

Rovira, A., & Slater, M. (2022a). Encouraging bystander helping behaviour in a violent incident: A virtual reality study using reinforcement learning. *Scientific Reports*, *12*(1), 3843. https://doi.org/10.1038/s41598-022-07872-3

Rovira, A., & Slater, M. (2022b). Encouraging bystander helping behaviour in a violent incident: A virtual reality study using reinforcement learning. *Scientific Reports*, *12*(1), Article 1. https://doi.org/10.1038/s41598-022-07872-3

Rovira, A., Southern, R., Swapp, D., Campbell, C., Zhang, J. J., Levine, M., & Slater, M. (2021). Bystander Affiliation Influences Intervention Behavior: A Virtual Reality Study. *SAGE Open*, *11*(3), 21582440211040076. https://doi.org/10.1177/21582440211040076

Rovira, A., Swapp, D., Spanlang, B., & Slater, M. (2009). The use of virtual reality in the study of people's responses to violent incidents. *Frontiers in Behavioral Neuroscience*, *3*. https://doi.org/10.3389/neuro.08.059.2009

RStudio Team. (2021). *RStudio: Integrated Development Environment for R* [Computer software]. RStudio, PBC. http://www.rstudio.com/

Rubin, M., Minns, S., Muller, K., Tong, M. H., Hayhoe, M. M., & Telch, M. J. (2020). Avoidance of social threat: Evidence from eye movements during a public speaking challenge using 360°- video. *Behaviour Research and Therapy*, *134*, 103706. https://doi.org/10.1016/j.brat.2020.103706

Rubin, M., Muller, K., Hayhoe, M. M., & Telch, M. J. (2022). Attention guidance augmentation of virtual reality exposure therapy for social anxiety disorder: A pilot randomized controlled trial. *Cognitive Behaviour Therapy*, *51*(5), 371–387. https://doi.org/10.1080/16506073.2022.2053882

Ryan, V., & Griffin, R. (2016). An Investigation Into Anxiety In Virtual Reality Following A Self-Compassion Induction. *Annual Review of CyberTherapy and Telemedicine*, *14*, pp 109-114.

Saberi, M., Bernardet, U., & DiPaola, S. (2014). *An Architecture for Personality-based, Nonverbal Behavior in Affective Virtual Humanoid Character | Elsevier Enhanced Reader*. https://doi.org/10.1016/j.procs.2014.11.104

Sagui-Henson, S. J., Welcome Chamberlain, C. E., Smith, B. J., Li, E. J., Castro Sweet, C., & Altman, M. (2022). Understanding Components of Therapeutic Alliance and Well-Being from Use of a Global Digital Mental Health Benefit During the COVID-19 Pandemic: Longitudinal Observational Study. *Journal of Technology in Behavioral Science*, 1–12. https://doi.org/10.1007/s41347-022-00263-5

Salber, D., & Coutaz, J. (1993). Applying the Wizard of Oz technique to the study of multimodal systems. In L. J. Bass, J. Gornostaev, & C. Unger (Eds.), *Human-Computer Interaction* (pp. 219–230). Springer. https://doi.org/10.1007/3-540-57433-6_51

Sanchez-Vives, M. V., & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience*, *6*(4), 332–339. https://doi.org/10.1038/nrn1651

Savulich, G., Freeman, D., Shergill, S., & Yiend, J. (2015). Interpretation Biases in Paranoia. *Behavior Therapy*, *46*(1), 110–124. https://doi.org/10.1016/j.beth.2014.08.002

Sawyer, S. F. (2009). Analysis of Variance: The Fundamental Concepts. *Journal of Manual & Manipulative Therapy*, *17*(2), 27E-38E. https://doi.org/10.1179/jmt.2009.17.2.27E

Scarpina, F., Serino, S., Keizer, A., Chirico, A., Scacchi, M., Castelnuovo, G., Mauro, A., & Riva, G. (2019). The Effect of a Virtual-Reality Full-Body Illusion on Body Representation in Obesity. *Journal of Clinical Medicine*, *8*(9). https://doi.org/10.3390/jcm8091330

Schmuckler, M. (2001). What Is Ecological Validity? A Dimensional Analysis. *Infancy*, *2*. https://doi.org/10.1207/S15327078IN0204_02

Schroeder, P. A., Gehrer, N. A., Reents, M., Reimer, N., Vagedes, J., & Svaldi, J. (2023). Body Dissatisfaction Directs Avatar Perception: Embodiment and Selective Visual Attention in Body Mass-Modified Self-Avatars. *Cyberpsychology, Behavior, and Social Networking*, *26*(11), 850–860. https://doi.org/10.1089/cyber.2022.0385

Schroeder, R. (2002). Social Interaction in Virtual Environments: Key Issues, Common Themes, and a Framework for Research. In R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments* (pp. 1–18). Springer. https://doi.org/10.1007/978-1-4471-0277-9_1

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments*, *10*(3), 266–281.

Schuetz, I., & Fiehler, K. (2022). Eye tracking in virtual reality: Vive pro eye spatial accuracy, precision, and calibration reliability. *Journal of Eye Movement Research*, *15*(3), Article 3. https://doi.org/10.16910/jemr.15.3.3

Schulte, D. (2008). Patients' outcome expectancies and their impression of suitability as predictors of treatment outcome. *Psychotherapy Research*, *18*(4), 481–494. https://doi.org/10.1080/10503300801932505

Seiferth, N. Y., Pauly, K., Habel, U., Kellermann, T., Jon Shah, N., Ruhrmann, S., Klosterkötter, J., Schneider, F., & Kircher, T. (2008). Increased neural response related to neutral faces in individuals at risk for psychosis. *NeuroImage*, *40*(1), 289–297. https://doi.org/10.1016/j.neuroimage.2007.11.020

Selaskowski, B., Wiebe, A., Kannen, K., Asché, L., Pakos, J., Philipsen, A., & Braun, N. (2024). Clinical adoption of virtual reality in mental health is challenged by lack of high-quality research. *Npj Mental Health Research*, *3*(1), 1–5. https://doi.org/10.1038/s44184-024-00069-8

Seo, S., Kim, E., Mundy, P., Heo, J., & Kim, K. (Kenny). (2019). Joint Attention Virtual Classroom: A Preliminary Study. *Psychiatry Investigation*, *16*(4), 292–299. https://doi.org/10.30773/pi.2019.02.08

Sep, M. S. C., van Osch, M., van Vliet, L. M., Smets, E. M. A., & Bensing, J. M. (2014). The power of clinicians' affective communication: How reassurance about non-abandonment can reduce patients' physiological arousal and increase information recall in bad news consultations. An experimental study using analogue patients. *Patient Education and Counseling*, *95*(1), 45–52. https://doi.org/10.1016/j.pec.2013.12.022

Shadiev, R., & Li, D. (2023). A review study on eye-tracking technology usage in immersive virtual reality learning environments. *Computers & Education*, *196*, 104681. https://doi.org/10.1016/j.compedu.2022.104681

Shaffer, F., & Ginsberg, J. P. (2017). An Overview of Heart Rate Variability Metrics and Norms. *Frontiers in Public Health*, *5*. https://www.frontiersin.org/articles/10.3389/fpubh.2017.00258

Shiban, Y., Reichenberger, J., Neumann, I. D., & Mühlberger, A. (2015). Social conditioning and extinction paradigm: A translational study in virtual reality. *Frontiers in Psychology*, *6*, 400. https://doi.org/10.3389/fpsyg.2015.00400

Shin, Y.-B., Kim, J.-J., Kim, M.-K., Kyeong, S., Jung, Y. H., Eom, H., & Kim, E. (2018a). Development of an effective virtual environment in eliciting craving in adolescents and young adults with internet gaming disorder. *PLOS ONE*, *13*(4), Article 4. https://doi.org/10.1371/journal.pone.0195677

Shin, Y.-B., Kim, J.-J., Kim, M.-K., Kyeong, S., Jung, Y. H., Eom, H., & Kim, E. (2018b). Development of an effective virtual environment in eliciting craving in adolescents and young adults with internet gaming disorder. *PLOS ONE*, *13*(4), e0195677. https://doi.org/10.1371/journal.pone.0195677

Silva, N., Schreck, T., Veas, E., Sabol, V., Eggeling, E., & Fellner, D. W. (2018). Leveraging eye-gaze and time-series features to predict user interests and build a recommendation model for visual analysis. *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*, 1–9. https://doi.org/10.1145/3204493.3204546

Slater, M., Neyret, S., Johnston, T., Iruretagoyena, G., Crespo, M. Á. de la C., Alabèrnia-Segura, M., Spanlang, B., & Feixas, G. (2019). An experimental study of a virtual reality counselling paradigm using embodied self-dialogue. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-46877-3

Slater, M., Pertaub, D.-P., & Steed, A. (1999). Public Speaking in Virtual Reality: Facing an Audience of Avatars. *Computer Graphics and Applications, IEEE*, *19*, 6–9. https://doi.org/10.1109/38.749116

Slater, M., & Wilbur, S. (1997). A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence:*

Teleoperators and Virtual Environments, *6*(6), 603–616. https://doi.org/10.1162/pres.1997.6.6.603

Souto, T., Baptista, A., Tavares, D., Queirós, C., & António, M. (2013). Facial emotional recognition in schizophrenia: Preliminary results of the virtual reality program for facial emotional recognition. *Archives of Clinical Psychiatry (São Paulo)*, *40*(4), 129–134. https://doi.org/10.1590/S0101-60832013000400001

Spanlang, B., Nierula, B., Haffar, M., & Debruille, J. B. (2018). Mimicking Schizophrenia: Reducing P300b by Minimally Fragmenting Healthy Participants' Selves Using Immersive Virtual Reality Embodiment. *Frontiers in Human Neuroscience*, *12*, 504. https://doi.org/10.3389/fnhum.2018.00504

Spanlang, B., Normand, J.-M., Borland, D., Kilteni, K., Giannopoulos, E., Pomés, A., González-Franco, M., Perez-Marcos, D., Arroyo-Palacios, J., Muncunill, X. N., & Slater, M. (2014). How to Build an Embodiment Lab: Achieving Body Representation Illusions in Virtual Reality. *Frontiers in Robotics and AI*, *1*, 9. https://doi.org/10.3389/frobt.2014.00009

Spyridonis, F., Daylamani-Zad, D., & Nightingale, J. (2024). PublicVR: A virtual reality exposure therapy intervention for adults with speech anxiety. *Virtual Reality*, *28*. https://doi.org/10.1007/s10055-024-00998-x

Stallmann, L., Tran, M., Rudrauf, D., Dukes, D., & Samson, A. C. (2023). Simulating Social Emotion Regulation in Virtual Reality: Effect of Virtual Social Support Following Ostracism in a Cyberball Game. *International Journal of Emerging Technologies in Learning (iJET)*, *18*(07), 4–27. https://doi.org/10.3991/ijet.v18i07.29419

Stein, J.-P., & Ohler, P. (2017). Venturing into the uncanny valley of mind—The influence of mind attribution on the acceptance of human-like characters in a virtual reality setting. *Cognition*, *160*, 43–50. https://doi.org/10.1016/j.cognition.2016.12.010

Steinman, S. A., & Teachman, B. A. (2011). Cognitive Processing and Acrophobia: Validating the Heights Interpretation Questionnaire. *Journal of Anxiety Disorders*, *25*(7), 896–902. https://doi.org/10.1016/j.janxdis.2011.05.001

Szekely, R., Mason, O., Frohlich, D., & Barley, E. (2024). 'It's not everybody's snapshot. It's just an insight into that world': A qualitative study of multiple perspectives towards understanding the mental health experience and addressing stigma in healthcare students through virtual reality. *DIGITAL HEALTH*, *10*, 20552076231223801. https://doi.org/10.1177/20552076231223801

Takac, M., Collett, J., Blom, K. J., Conduit, R., Rehm, I., & De Foe, A. (2019). Public speaking anxiety decreases within repeated virtual reality training sessions. *PloS One*, *14*(5), e0216288. https://doi.org/10.1371/journal.pone.0216288

Takemoto, A., Aispuriete, I., Niedra, L., & Dreimane, L. F. (2023). Differentiating depression using facial expressions in a virtual avatar communication system. *Frontiers in Digital Health*, *5*. https://doi.org/10.3389/fdgth.2023.1080023

Talaat, F. M. (2023). Real-time facial emotion recognition system among children with autism based on deep learning and IoT. *Neural Computing and Applications*, *35*(17), 12717–12728. https://doi.org/10.1007/s00521-023-08372-9

Tarnanas, Ioannis, Tsoukalas, Ioannis, & Stogiannidou, A. (2003). Virtual reality as a psychosocial coping environment. *Annual Review of CyberTherapy and Telemedicine*, *1*(Bowers, D., Meichenbaum, D. (1994). The unconscious reconsidered. New York: John Willey.), 77–81.

ter Stal, S., Jongbloed, G., & Tabak, M. (2021). Embodied Conversational Agents in eHealth: How Facial and Textual Expressions of Positive and Neutral Emotions Influence Perceptions of Mutual Understanding. *Interacting with Computers*, *33*(2), 167–176. https://doi.org/10.1093/iwc/iwab019

Thalmann, D. (2000). *Challenges for the Research in Virtual Humans*. 9.

Thompson-Hollands, J., Bentley, K. H., Gallagher, M. W., Boswell, J. F., & Barlow, D. H. (2014a). Credibility and Outcome Expectancy in the Unified Protocol: Relationship to Outcomes. *Journal of Experimental Psychopathology*, *5*(1), 72–82. https://doi.org/10.5127/jep.033712

Thompson-Hollands, J., Bentley, K. H., Gallagher, M. W., Boswell, J. F., & Barlow, D. H. (2014b). Credibility and Outcome Expectancy in the Unified Protocol: Relationship to Outcomes. *Journal of Experimental Psychopathology*, *5*(1), 72–82. https://doi.org/10.5127/jep.033712

Thompson-Lake, D. G. Y., Cooper, K. N., Mahoney, J. J., Bordnick, P. S., Salas, R., Kosten, T. R., Dani, J. A., & De La Garza, R. (2015). Withdrawal Symptoms and Nicotine Dependence Severity Predict Virtual Reality Craving in Cigarette-Deprived Smokers. *Nicotine & Tobacco Research*, *17*(7), 796–802. https://doi.org/10.1093/ntr/ntu245

Tinwell, A., Grimshaw, M., Nabi, D. A., & Williams, A. (2011). Facial expression of emotion and perception of the Uncanny Valley in virtual characters. *Computers in Human Behavior*, *27*(2), 741–749. https://doi.org/10.1016/j.chb.2010.10.018

Treal, T., Jackson, P. L., Jeuvrey, J., Vignais, N., & Meugnot, A. (2021). Natural human postural oscillations enhance the empathic response to a facial pain expression in a virtual character. *Scientific Reports*, *11*(1), 12493. https://doi.org/10.1038/s41598-021-91710-5

Trivedi, H., & Mousas, C. (2023). Human-virtual crowd interaction: Towards understanding the effects of crowd avoidance proximity in an immersive virtual environment. *Computer Animation and Virtual Worlds*, *34*(3–4), e2169. https://doi.org/10.1002/cav.2169

Trotta, A., Kang, J., Stahl, D., & Yiend, J. (2021). Interpretation Bias in Paranoia: A Systematic Review and Meta-Analysis. *Clinical Psychological Science*, *9*(1), 3–23. https://doi.org/10.1177/2167702620951552

Tsiourti, C., Weiss, A., Wac, K., & Vincze, M. (2019). Multimodal Integration of Emotional Signals from Voice, Body, and Context: Effects of (In)Congruence on Emotion Recognition and Attitudes Towards Robots. *International Journal of Social Robotics*, *11*(4), 555–573. https://doi.org/10.1007/s12369-019-00524-z

Tsoli, A., Mahmood, N., & Black, M. J. (2014). Breathing life into shape: Capturing, modeling and animating 3D human breathing. *ACM Trans. Graph.*, *33*(4), 52:1-52:11. https://doi.org/10.1145/2601097.2601225

Ulicny, B., & Thalmann, D. (2001). Crowd simulation for interactive virtual environments and VR training systems. In N. Magnenat-Thalmann & D. Thalmann (Eds.), *Computer Animation and Simulation 2001* (pp. 163–170). Springer Vienna. https://doi.org/10.1007/978-3-7091-6240-8_15

Urbaniak, G., & Plous, S. (2013). *Research randomizer (version 4.0)[computer software]. 2013* [Computer software].

Vahle, N., & Tomasik, M. J. (2022). The Embodiment of an Older Avatar in a Virtual Reality Setting Impacts the Social Motivation of Young Adults. *Experimental Aging Research*, *48*(2), 164–176. https://doi.org/10.1080/0361073X.2021.1943793

Valmaggia, L. R., Day, F., & Rus-Calafell, M. (2016). Using virtual reality to investigate psychological processes and mechanisms associated with the onset and maintenance of psychosis: A systematic review. *Social Psychiatry and Psychiatric Epidemiology*, *51*(7), 921–936. https://doi.org/10.1007/s00127-016-1245-0

Valmaggia, L. R., Freeman, D., Green, C., Garety, P., Swapp, D., Antley, A., Prescott, C., Fowler, D., Kuipers, E., Bebbington, P., Slater, M., Broome, M., & McGuire, P. K. (2007). Virtual reality and paranoid ideations in people with an 'at-risk mental state' for psychosis. *The British Journal of Psychiatry*, *191*(S51), s63–s68. https://doi.org/10.1192/bjp.191.51.s63

Van Gelder, J.-L., Cornet, L. J. M., Zwalua, N. P., Mertens, E. C. A., & van der Schalk, J. (2022). Interaction with the future self in virtual reality reduces self-defeating behavior in a sample of convicted offenders. *Scientific Reports*, *12*(1), Article 1. https://doi.org/10.1038/s41598-022-06305-5

Veldmeijer, L., Terlouw, G., Van Os, J., Van Dijk, O., Van 't Veer, J., & Boonstra, N. (2023). The Involvement of Service Users and People With Lived Experience in Mental Health Care Innovation Through Design: Systematic Review. *JMIR Mental Health*, *10*, e46590. https://doi.org/10.2196/46590

Venning, A., Herd, M. C., Oswald, T. K., Razmi, S., Glover, F., Hawke, T., Quartermain, V., & Redpath, P. (2021). Exploring the acceptability of a digital mental health platform incorporating a virtual coach: The good, the bad, and the opportunities. *Health Informatics Journal*, *27*(1), 1460458221994873. https://doi.org/10.1177/1460458221994873

Vinayagamoorthy, V., Gillies, M., Steed, A., Tanguy, E., Pan, X., Loscos, C., & Slater, M. (2006). *Building Expression into Virtual Characters*. 41.

Volonte, M., Hsu, Y.-C., Liu, K.-Y., Mazer, J. P., Wong, S.-K., & Babu, S. V. (2020). Effects of Interacting with a Crowd of Emotional Virtual Humans on Users' Affective and Non-Verbal Behaviors. *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 293–302. https://doi.org/10.1109/VR46266.2020.00049

Volonte, M., Ofek, E., Jakubzak, K., Bruner, S., & Gonzalez-Franco, M. (2022). HeadBox: A Facial Blendshape Animation Toolkit for the Microsoft Rocketbox Library. *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 39–42. https://doi.org/10.1109/VRW55335.2022.00015

Voutilainen, L., Henttonen, P., Kahri, M., Ravaja, N., Sams, M., & Peräkylä, A. (2018). Empathy, Challenge, and Psychophysiological Activation in Therapist–Client Interaction. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00530

Wang, Z., Giannopoulos, E., Slater, M., Peer, A., & Buss, M. (2011). Handshake: Realistic Human-Robot Interaction in Haptic Enhanced Virtual Reality. *Presence: Teleoperators and Virtual Environments*, *20*(4), 371–392. https://doi.org/10.1162/PRES_a_00061

Wei, S., Freeman, D., & Rovira, A. (2023). A randomised controlled test of emotional attributes of a virtual coach within a virtual reality (VR) mental health treatment. *Scientific Reports*, *13*(1), Article 1. https://doi.org/10.1038/s41598-023-38499-7

Wienrich, C., & Latoschik, M. E. (2021). eXtended Artificial Intelligence: New Prospects of Human-AI Interaction Research. *Frontiers in Virtual Reality*, *2*. https://doi.org/10.3389/frvir.2021.686783

Wieser, M., Pauli, P., Grosseibl, M., Molzow, I., & Mühlberger, A. (2010a). Virtual Social Interactions in Social Anxiety-The Impact of Sex, Gaze, and Interpersonal Distance. *Cyberpsychology, Behavior and Social Networking*, *13*, 547–554. https://doi.org/10.1089/cyber.2009.0432

Wieser, M., Pauli, P., Grosseibl, M., Molzow, I., & Mühlberger, A. (2010b). Virtual Social Interactions in Social Anxiety-The Impact of Sex, Gaze, and Interpersonal Distance. *Cyberpsychology, Behavior and Social Networking*, *13*, 547–554. https://doi.org/10.1089/cyber.2009.0432

Williams, J. G. L., Jones, J. R., Workhoven, M. N., & Williams, B. (1975). The Psychological Control of Preoperative Anxiety. *Psychophysiology*, *12*(1), 50–54. https://doi.org/10.1111/j.1469-8986.1975.tb03059.x

Witmer, B., & Singer, M. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, *7*(3), 225–240.

Wobbrock, J. O., Findlater, L., Gergle, D., & Higgins, J. J. (2011). The aligned rank transform for nonparametric factorial analyses using only anova procedures. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 143–146. https://doi.org/10.1145/1978942.1978963

Wolf, E., Merdan, N., Dölinger, N., Mal, D., Wienrich, C., Botsch, M., & Latoschik, M. E. (2021). The Embodiment of Photorealistic Avatars Influences Female Body Weight Perception in Virtual Reality. *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, 65–74. https://doi.org/10.1109/VR50410.2021.00027

Wonghempoom, A., Chuangchai, W., & Selanon, P. (2023). Physiological Stress Responses to Fear and Anxiety in a Height Change Experiment among Non-Labor Teenagers. *Safety*, *9*(2), Article 2. https://doi.org/10.3390/safety9020019

Wu, J., Sun, Y., Zhang, G., Zhou, Z., & Ren, Z. (2021). Virtual Reality-Assisted Cognitive Behavioral Therapy for Anxiety Disorders: A Systematic Review and Meta-Analysis. *Frontiers in Psychiatry*, *12*. https://doi.org/10.3389/fpsyt.2021.575094

Wu, Y., Wang, Y., Jung, S., Hoermann, S., & Lindeman, R. W. (2021). Using a Fully Expressive Avatar to Collaborate in Virtual Reality: Evaluation of Task Performance, Presence, and Attraction. *Frontiers in Virtual Reality*, *2*. https://www.frontiersin.org/articles/10.3389/frvir.2021.641296

Wuehr, M., Breitkopf, K., Decker, J., Ibarra, G., Huppert, D., & Brandt, T. (2019). Fear of heights in virtual reality saturates 20 to 40 m above ground. *Journal of Neurology*, *266*(Suppl 1), 80–87. https://doi.org/10.1007/s00415-019-09370-5

Yamamoto, K., & Nakao, T. (2022). Top-Down Influence Leads to a Reduced Sense of Body Ownership in Individuals With Depersonalization Tendencies: A Focus on Full Body Illusion. *Frontiers in Psychology*, *13*. https://www.frontiersin.org/articles/10.3389/fpsyg.2022.790960

Zell, E., Zibrek, K., & McDonnell, R. (2019). Perception of virtual characters. *ACM SIGGRAPH 2019 Courses*, 1–17. https://doi.org/10.1145/3305366.3328101

Zibrek, K., Kokkinara, E., & Mcdonnell, R. (2018). The Effect of Realistic Appearance of Virtual Characters in Immersive Environments—Does the Character's Personality Play a Role? *IEEE Transactions on Visualization and Computer Graphics*, *24*(4), 1681–1690. IEEE Transactions on Visualization and Computer Graphics. https://doi.org/10.1109/TVCG.2018.2794638

6 Conclusions

Virtual humans (VHs) play a significant role in VR mental health applications, yet their impact as a distinct focus in mental health research has been underexplored. The aim of this thesis was to investigate how these characteristics are designed, especially focusing on how their facial expressions can enhance VR mental health applications. This concluding chapter recaps the findings of each chapter, positioning them within the context of existing research. It also discusses the overall limitations of the thesis and suggests directions for future research.

6.1 THESIS SUMMARY

Chapter 1 introduced the concepts related to VR mental health research, VHs in mental health applications, and their facial expressions.

The systematic review presented in Chapter 2 delved into how VHs have been used in VR mental health research, from the early studies to the current state of the art. We highlighted how VHs are pivotal in several research areas such as in social anxiety disorders and body dysmorphic disorders, and the versatility of their roles in mental health applications. Nevertheless, very few studies provided detailed information on how the VHs were designed. Chapter 2 concluded that there is a lack of comprehensive studies on the effects of important VH characteristics (e.g. facial expressions). This led to the exploration of these effects in two randomized controlled tests, focusing on the two types of VHs: a VR coach that guides users throughout the programme (Chapters 3, Chapter 4) and the virtual characters used as a stimulus to trigger their fears (Chapter 5).

Chapter 3 examined the effects of a virtual coach's emotional attributes (warm facial expressions and affirmative head nods) on people's therapeutic alliance and expectancy of therapy outcome, which are reliable predictors of treatment success (Ardito & Rabellino, 2011; Thompson-Hollands et al., 2014), during the introductory stage of an automated VR therapy. The findings showed that both warm facial expressions and affirmative head nods enhanced therapy engagement and confidence in the treatment. Chapter 4 explored participants' physiological responses by looking at their cardiac responses and EDA using the collected data from Chapter 3. Exposure to virtual heights elicited stronger cardiac and EDA responses compared to the indoor consultation. During the indoor consultation, virtual coach's warm facial expressions increased the SCRs and that affirmative head nods showed a trend to increase the SCL. However, no significant effects were observed on cardiac measures across experimental conditions. These results provide evidence that the VR environment and the coach's non-verbal behaviours can modulate physiological responses.

Chapter 5 focused on the design and implementation of VH faces when the characters were used as part of the stimuli in a VR social situation to assess people's paranoid interpretation and visual attention. Both the addition of facial animation and positive expressions separately reduced the likelihood of having paranoid thoughts towards the VHs. Positive expressions resulted in less visual attention to VHs when their faces were static. The findings indicated that the detailed design of VH facial expressions significantly influenced the occurrence of paranoid thoughts and visual attention.

Overall Contribution: This thesis addressed the gap in the current literature and provided initial evidence of the psychological, behavioural, and physiological impacts of VH design, particularly related to facial expressions. Two randomised controlled tests with an overall of around 250 participants examined the detailed design of VHs for two conditions (fear of heights and paranoia). Facial characteristics showed medium to large effect sizes in both studies, demonstrating the importance of incorporating emotional attributes and carefully crafted VH faces in enhancing mental health outcomes and shaping interpretations of virtual characters' intentions. This highlights the crucial need for thoughtful attention to VH design and implementation, tailored to their specific purposes. For instance, in VR-based phobia interventions, positive emotional attributes in a virtual coach can strengthen therapy engagement.

Conversely, in VR applications for paranoia, neutral emotional attributes may be more beneficial, as they challenge unfounded beliefs and help individuals reinterpret ambiguous social cues from a more balanced perspective.

6.2 LIMITATIONS

This thesis has several limitations beyond those specific to the individual studies. First, our experimental studies focused on two specific VR exemplars related to mental health applications — phobia (fear of heights) and paranoia. Other mental health conditions frequently studied in VR, such as social anxiety disorders (Shiban et al., 2015; Takac et al., 2019) and eating disorders (Bektas et al., 2023; Natali et al., 2024), were not covered. Thus, our findings may not be generalisable to these other conditions.

Second, there were constraints related to the experimental design and recruitment, particularly regarding the use of randomised controlled tests and our participant sample. Randomised controlled tests are resource-intensive and time-consuming; while they provide a robust design to minimise bias and establish causality (Hariton & Locascio, 2018), they limited our ability to test a broader range of VH characteristics or conduct multi-session experiments. Additionally, our sample was from the non-clinical population based around Oxfordshire. Although the gender and age distribution of our participants was well-represented (with both experimental studies covering the age range from 18 to over 70), the recruited sample may not be representative of a wider or clinical population. Furthermore, as we targeted a non-clinical group and conducted single-session lab studies, the studies lacked direct outcome measures for assessing and treating mental health conditions. Existing VR clinical trials with outcome evaluation normally employed multiple sessions of training and therapy (Freeman et al., 2018, 2022; Pot-Kolder et al., 2018), and it remains unclear whether the VH design improvements explored in this thesis will translate into clinical outcomes.

Third, the thesis focused on specific aspects of facial expressions, specifically examining animation (static vs animated) and positive expressions (neutral vs warm). We did not explore other relevant facial characteristics such as emotion intensity (Barisnikov et al., 2021), negative and mixed emotions (Kaminska et al., 2020; Treal et al., 2021; Volonte et al., 2020), or the influence of adaptive facial expressions that changed in real time based on participants' responses (Dobs et al., 2018; Horigome et al., 2023). Additionally, the relationship between participants' facial features and those of VHs was not investigated. Previous studies have shown that VHs with facial features similar to the participant can intensify emotional responses (Aymerich-Franch et al., 2014), little is known about whether facial expression similarity or facial mimicry from VHs also affects therapeutic outcomes in VR. Therefore, although this thesis provided initial insights into VH faces, a broader range of facial features needs to be tested for a deeper understanding.

Fourth, regarding the use cases of VHs, we tested virtual coaches (Chapters 3 and Chapter 4) and virtual crowds as part of stimuli (Chapter 5). Virtual body representations (e.g. participants' avatars), which are another key role of VHs discussed in our systematic review (Chapter 2), were not studied. This decision was made to prioritise the two types of VHs most commonly seen in social VR mental health settings, with explicit and implicit interaction with participants. Also, most existing mental health studies did not directly mapped participants' facial expressions onto their avatars (Falconer et al., 2016; Mendoza-Medialdea et al., 2023; Schroeder et al., 2023). However, the omission of avatar studies leaves a gap in understanding how the design of personal avatars might impact mental health experience, especially in cases where embodiment and identity are essential.

Finally, regarding data collection and analysis, we primarily analysed the data separately to evaluate the impacts of VH design on people's psychological, behavioural, physiological responses. We did not explore interaction factors between different data types, which could have provided a more holistic understanding of participant responses, such as predicting emotion from the physiological indicators in real time (El Ali et al., 2023; Gupta et al., 2024). Additionally, we collected only quantitative data and did not include any qualitative insights. This may have limited our insight into participants' thought processes and the reasoning behind their responses, thereby affecting our understanding of how they interact with and perceive VHs (Fornells-Ambrojo et al., 2016).

6.3 FUTURE DIRECTIONS

This thesis can serve as a starting point for understanding the impact of VH characteristics in VR mental health applications, from which research can take different paths to understand their role to design more effective mental health assessment and treatments in VR.

6.3.1 Clinical Applications Extension

Building on the positive results obtained from the in non-clinical populations, future research should test the effects of VH design in clinical populations. Our VR studies were designed as single-session tests; multi-session clinical trials would be sensitive next step to understand the impact of VH design on mental health outcomes (Freeman et al., 2022; Jeong et al., 2021). Recent studies suggest that single-session VR exposure therapy with gradual increase of the stimuli intensity can be as effective as multi-session exposure (Banakou et al., 2024). However, longitudinal studies are needed to understand the long-term effects of the VH interactions. Evaluating the lasting effects of VH interactions across well-represented demographics through multi-session and multi-site clinical trials is vital. Furthermore, longitudinal studies could help determine parameters such as the optimal frequency and duration of VH-based therapies, ultimately enhancing their long-term effectiveness and therapeutic value.

6.3.2 Additional VH Features and Use Cases

The experimental tests showed the impact of the facial animations and expressions on VHs, future research can expand the investigation to other facial parameters. A better understanding of these complex interactions can guide the optimisation of VH design for therapeutic purposes. For instance, examining the intensity, timing of facial expressions between two or more people, or the interaction effects of facial features and other VH behaviours, could lead to a better understanding of the VH faces. The findings in Chapter 3 regarding the virtual coach showed a trend where positive facial expressions and affirmative head nods together resulted in the greatest improvement in therapeutic alliance. This highlights the potential for combining different non-verbal cues in VH design to amplify the empathetic and supportive interactions within VR. In contrast, combining positive facial expressions with incongruent postures could weaken the intended emotional impact (Clavel et al., 2009).

Beyond facial expressions, future research should also explore other VH characteristics, such as visual appearance (e.g., realistic vs. stylised), body animation, and voice. These features have been shown to influence user perception and emotional responses in virtual interactions (Ma & Pan, 2022; Mousas et al., 2018; Tsiourti et al., 2019; Zibrek et al., 2018), but their effects on the context of mental health are not yet understood. Furthermore, generative AI offer new opportunities to customise interactions with VHs. Large language models (LLMs), such as ChatGPT [15] and

[15] https://chat.openai.com/

Claude [16], can generate real-time, personalised responses during consultations, addressing users' specific fears and concerns. Also, AI-driven voice synthesis (e.g. ElevenLabs [17], resemble.ai [18]) could enable VHs to speak in participants' native languages or adjust vocal tones to suit emotional contexts. These enhancements could be used in virtual coaching for personalised guidance or to make stimuli more effective in triggering anxiety, thereby improving therapeutic outcomes.

Given the acknowledged limitations around the scope of VH roles and mental health conditions, future work would benefit from systematically examining different VH use cases across a wider range of mental health disorders. One potential direction is to explore how the detailed design of avatars influences people's perceptions and behaviours. For instance, in Slater et al.'s (2019) embodied VR counselling paradigm, participants alternated between embodying themselves and Sigmund Freud, describing their personal problems to Freud before switching roles to offer advice to their own avatar from Freud's perspective. Future studies could utilise this unique body-swapping technique to test whether adding facial animation capture to the avatars further enhances participants' emotional connection and perceptions of self-change. This counselling paradigm could also be adapted to enhance self-awareness and decrease negative thoughts in psychosis, depression, and anxiety, further diversifying the scope of VR interventions.

6.3.3 Holistic Multimodal Data Analysis

To gain a more holistic understanding of the participant experience, it is worth incorporating a wider range of data collection and analysis techniques. Qualitative methods, such as interviews or focus groups, could offer deeper insights into the program's usability, tolerance, and limitations (Amestoy Alonso et al., 2024; Brown et

[16] https://www.anthropic.com/
[17] https://elevenlabs.io/
[18] https://www.resemble.ai/

al., 2022; Szekely et al., 2024). Current VR headsets like the Meta Quest Pro or Apple Vision Pro, which support integrated facial tracking, enable researchers to collect additional data about participants' facial expressions during interactions with VHs. This would expand our understanding of emotional responses and mimicry effects in facial expressions. For data analysis, time series analysis of eye gaze behaviour and physiological data (e.g., heart rate and skin conductance) could also track changes over time, identifying key patterns in participant responses (Jebb et al., 2015; Silva et al., 2018). Furthermore, employing machine learning algorithms could integrate these multimodal datasets, including subjective measures, behavioural and physiological responses, to predict emotional and cognitive states in real time (Chang et al., 2022; Rahman et al., 2023). Such systems could adjust virtual environments and VH behaviour over the time during the same VR session, maximising therapeutic efficacy and user engagement.

6.3.4 User-centred Experience Design

Finally, involving individuals with lived experience in the development of VR applications can benefit both patients and researchers (Brett et al., 2014), and should also be considered for VH design. Patients and public involvement (PPI) groups play a crucial role in fostering a sense of empowerment and ownership among patients, making them feel valued and respected. By allowing them to directly contribute to the VH design from the early stages, patients can actively shape the research and VR content, ensuring their lived experiences are acknowledged and utilised. This not only enhances the relevance and usability of the VR programme but also increases patient engagement and satisfaction with the overall experience. For researchers, it provides critical feedback from individuals with first-hand experience, helping to refine scenarios and focus on key character features. This participatory approach is likely to improve the therapeutic outcomes and effectiveness of VR assessments and interventions (Knight et al., 2021; Veldmeijer et al., 2023; Selaskowski et al., 2024).

6.4 PERSONAL REFLECTION

My DPhil represents interdisciplinary work encompassing computer science (VR) and experimental psychology, which has been both fulfilling and challenging. One challenge was navigating the technical complexities of 3D animation and VR development, while ensuring the psychological validity of the VR scenarios and VH behaviours. Implementation often takes multiple iterations; for example, the development of the VR coach (in Chapter 3) involved a complete software reiteration after testing. This necessitated restarting the user study from scratch, even after collecting 60 datasets, due to necessary adjustments.

Additionally, both experimental studies required recruiting a large number of participants, and I was solely responsible for managing all testing end-to-end. How to efficiently, ethically, and economically recruiting target participants within a short timeframe became an interesting applied research in itself. Another challenge was handling the uncertainties of behavioural and physiological data in VR. When using commercial physiological sensors and devices, what may appear as noise for one participant could serve as a signal for another. Understanding the capabilities and limitations of the chosen technology and devices was essential for interpreting the data. Although this process was time-consuming and sometimes felt out of scope, it often led to unplanned but valuable contributions to the broader research community. For instance, the eye-tracking accuracy test discussed in Chapter 5 arose from the attempt to bridge the information gap of the limited information about Quest Pro's eye-tracking specifications during my development of the VR stimuli.

For someone starting in this field, my advice would be to embrace both technical and psychological perspectives and seek collaboration across disciplines early on. Programming skills and computer literacy are important, but equally critical is the ability to translate psychological or clinical theories into VR experience creation. Be prepared for rapid learning, trial and error, and experimenting with different implementation pipelines and constantly updated software. Many of these efforts may seem to lead nowhere on the project Gantt chart, but there are always insights to be gained during detours. Research isn't just about writing manuscripts; hands-on prototyping or participating in hackathons can keep your problem-solving skills flexible. I highly recommend events like the *MIT Reality Hack* for XR prototyping.

And last, learning from the users. Spending substantial time observing how participants interact with the system provides crucial insights. In-person user testing can be tedious and sometimes feels like it lacks technical glamour. However, after conducting over 300 hours of in-person VR sessions, I began to recognise patterns in user behaviour and interactions with the virtual characters. Researchers might spend hours optimising the polycount of a model or enhancing rendering quality, only for the immersive experience to be undermined because a user finds the fabric of the headset strap uncomfortable.

REFERENCES

Amestoy Alonso, B., Donegan, T., Calvis, I., Swidrak, J., Rodriguez, E., Vargas-Reverón, C. L., Combalia, A., Oliva Martinez, R., & Sanchez-Vives, M. V. (2024). Focus groups in the metaverse: Shared virtual spaces for patients, clinicians, and researchers. *Frontiers in Virtual Reality*, *5*. https://doi.org/10.3389/frvir.2024.1432282

Ardito, R. B., & Rabellino, D. (2011). Therapeutic Alliance and Outcome of Psychotherapy: Historical Excursus, Measurements, and Prospects for Research. *Frontiers in Psychology*, *2*, 270. https://doi.org/10.3389/fpsyg.2011.00270

Aymerich-Franch, L., Kizilcec, R. F., & Bailenson, J. N. (2014). The Relationship between Virtual Self Similarity and Social Anxiety. *Frontiers in Human Neuroscience*, *8*, 944. https://doi.org/10.3389/fnhum.2014.00944

Banakou, D., Johnston, T., Beacco, A., Senel, G., & Slater, M. (2024). Desensitizing Anxiety Through Imperceptible Change: Feasibility Study on a Paradigm for Single-Session Exposure Therapy for Fear of Public Speaking. *JMIR Formative Research*, *8*(1), e52212. https://doi.org/10.2196/52212

Barisnikov, K., Thomasson, M., Stutzmann, J., & Lejeune, F. (2021). Sensitivity to Emotion Intensity and Recognition of Emotion Expression in Neurotypical Children. *Children*, *8*(12), 1108. https://doi.org/10.3390/children8121108

Bektas, S., Natali, L., Rowlands, K., Valmaggia, L., Di Pietro, J., Mutwalli, H., Himmerich, H., Treasure, J., & Cardi, V. (2023). Exploring Correlations of Food-Specific Disgust with Eating Disorder Psychopathology and Food Interaction: A Preliminary Study Using Virtual Reality. *Nutrients*, *15*(20), Article 20. https://doi.org/10.3390/nu15204443

Brett, J., Staniszewska, S., Mockford, C., Herron-Marx, S., Hughes, J., Tysall, C., & Suleman, R. (2014). A Systematic Review of the Impact of Patient and Public Involvement on Service Users, Researchers and Communities. *The Patient - Patient-Centered Outcomes Research*, *7*(4), 387–395. https://doi.org/10.1007/s40271-014-0065-0

Brown, P., Waite, F., Lambe, S., Jones, J., Jenner, L., Diamond, R., & Freeman, D. (2022). Automated Virtual Reality Cognitive Therapy (gameChange) in Inpatient Psychiatric Wards: Qualitative Study of Staff and Patient Views Using an Implementation Framework. *JMIR Formative Research*, *6*(4), e34225. https://doi.org/10.2196/34225

Chang, Z., Bai, H., Zhang, L., Gupta, K., He, W., & Billinghurst, M. (2022). The impact of virtual agents' multimodal communication on brain activity and cognitive load in Virtual Reality. *Frontiers in Virtual Reality*, *3*. https://doi.org/10.3389/frvir.2022.995090

Clavel, C., Plessier, J., Martin, J.-C., Ach, L., & Morel, B. (2009). Combining Facial and Postural Expressions of Emotions in a Virtual Character. In Z. Ruttkay, M. Kipp, A. Nijholt, & H. H. Vilhjálmsson (Eds.), *Intelligent Virtual Agents* (pp. 287–300). Springer. https://doi.org/10.1007/978-3-642-04380-2_31

Dobs, K., Bülthoff, I., & Schultz, J. (2018). Use and Usefulness of Dynamic Face Stimuli for Face Perception Studies—A Review of Behavioral Findings and Methodology. *Frontiers in Psychology*, *9*, 1355. https://doi.org/10.3389/fpsyg.2018.01355

El Ali, A., Ney, R., Van Berlo, Z. M. C., & Cesar, P. (2023). Is that My Heartbeat? Measuring and Understanding Modality-Dependent Cardiac Interoception in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, *29*(11), 4805–4815. https://doi.org/10.1109/TVCG.2023.3320228

Falconer, C. J., Rovira, A., King, J. A., Gilbert, P., Antley, A., Fearon, P., Ralph, N., Slater, M., & Brewin, C. R. (2016). Embodying self-compassion within virtual reality and its effects on patients with depression. *BJPsych Open*, *2*(1), 74–80. https://doi.org/10.1192/bjpo.bp.115.002147

Fornells-Ambrojo, M., Elenbaas, M., Barker, C., Swapp, D., Navarro, X., Rovira, A., Sanahuja, J. M. T., & Slater, M. (2016). Hypersensitivity to Contingent Behavior in Paranoia: A New Virtual Reality Paradigm. *The Journal of Nervous and Mental Disease*, *204*(2), 148–152. https://doi.org/10.1097/NMD.0000000000000414

Freeman, D., Haselton, P., Freeman, J., Spanlang, B., Kishore, S., Albery, E., Denne, M., Brown, P., Slater, M., & Nickless, A. (2018). Automated psychological therapy using immersive virtual reality for treatment of fear of heights: A single-blind, parallel-group, randomised controlled trial. *The Lancet. Psychiatry*, *5*(8), 625–632. https://doi.org/10.1016/S2215-0366(18)30226-8

Freeman, D., Lambe, S., Kabir, T., Petit, A., Rosebrock, L., Yu, L.-M., Dudley, R., Chapman, K., Morrison, A., O'Regan, E., Aynsworth, C., Jones, J., Murphy, E., Powling, R., Galal, U., Grabey, J., Rovira, A., Martin, J., Hollis, C., … West, J. (2022). Automated virtual reality therapy to treat agoraphobic avoidance and distress in patients with psychosis (gameChange): A multicentre, parallel-group, single-blind, randomised, controlled trial in England with mediation and moderation analyses. *The Lancet Psychiatry*, *9*(5), 375–388. https://doi.org/10.1016/S2215-0366(22)00060-8

Gupta, K., Zhang, Y., Gunasekaran, T. S., Krishna, N., Pai, Y. S., & Billinghurst, M. (2024). CAEVR: Biosignals-Driven Context-Aware Empathy in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, 1–11. https://doi.org/10.1109/TVCG.2024.3372130

Hariton, E., & Locascio, J. J. (2018). Randomised controlled trials—The gold standard for effectiveness research. *BJOG : An International Journal of Obstetrics and Gynaecology*, *125*(13), 1716. https://doi.org/10.1111/1471-0528.15199

Horigome, T., Yoshida, S., Tanikawa, T., Mimura, M., & Kishimoto, T. (2023). Modification of the therapist's facial expressions using virtual reality technology during the treatment of social anxiety disorder: A case series. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1030050

Jebb, A. T., Tay, L., Wang, W., & Huang, Q. (2015). Time series analysis for psychological research: Examining and forecasting change. *Frontiers in Psychology*, *6*, 727. https://doi.org/10.3389/fpsyg.2015.00727

Jeong, H. S., Lee, J. H., Kim, H. E., & Kim, J.-J. (2021). Appropriate Number of Treatment Sessions in Virtual Reality-Based Individual Cognitive Behavioral Therapy for Social Anxiety Disorder. *Journal of Clinical Medicine*, *10*(5), 915. https://doi.org/10.3390/jcm10050915

Kaminska, O. K., Magnuski, M., Olszanowski, M., Gola, M., Brzezicka, A., & Winkielman, P. (2020). Ambiguous at the second sight: Mixed facial expressions trigger late electrophysiological responses linked to lower social impressions. *Cognitive, Affective, & Behavioral Neuroscience*, *20*(2), 441–454. https://doi.org/10.3758/s13415-020-00778-5

Knight, I., West, J., Matthews, E., Kabir, T., Lambe, S., Waite, F., & Freeman, D. (2021). Participatory design to create a VR therapy for psychosis. *Design for Health*, *5*(1), 98–119. https://doi.org/10.1080/24735132.2021.1885889

Ma, F., & Pan, X. (2022). Visual Fidelity Effects on Expressive Self-avatar in Virtual Reality: First Impressions Matter. *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 57–65. https://doi.org/10.1109/VR51125.2022.00023

Mendoza-Medialdea, M. T., Meschberger-Annweiler, F.-A., Ascione, M., Rueda-Pina, A., Rabarbari, E., Porras-Garcia, B., Ferrer-Garcia, M., & Gutiérrez-Maldonado, J. (2023). Body Dissatisfaction and Body-Related Attentional Bias: Is There a Causal Relationship? *Journal of Clinical Medicine*, *12*(17), Article 17. https://doi.org/10.3390/jcm12175659

Mousas, C., Anastasiou, D., & Spantidi, O. (2018). The effects of appearance and motion of virtual characters on emotional reactivity. *Computers in Human Behavior*, *86*, 99–108. https://doi.org/10.1016/j.chb.2018.04.036

Natali, L., Meregalli, V., Rowlands, K., Di Pietro, J., Treasure, J., Collantoni, E., Meneguzzo, P., Tenconi, E., Favaro, A., Fontana, F., Ceccato, E., Sala, A., Valmaggia, L., & Cardi, V. (2024). Virtual food exposure with positive mood induction or social support to reduce food anxiety in anorexia nervosa: A feasibility study. *International Journal of Eating Disorders*, *57*(3), 703–715. https://doi.org/10.1002/eat.24155

Pot-Kolder, R. M. C. A., Geraets, C. N. W., Veling, W., Beilen, M. van, Staring, A. B. P., Gijsman, H. J., Delespaul, P. A. E. G., & Gaag, M. van der. (2018). Virtual-reality-based cognitive behavioural therapy versus waiting list control for paranoid ideation and social avoidance in patients with psychotic disorders: A single-blind randomised controlled trial. *The Lancet Psychiatry*, *5*(3), 217–226. https://doi.org/10.1016/S2215-0366(18)30053-1

Rahman, M. A., Brown, D. J., Mahmud, M., Harris, M., Shopland, N., Heym, N., Sumich, A., Turabee, Z. B., Standen, B., Downes, D., Xing, Y., Thomas, C., Haddick, S., Premkumar, P., Nastase, S., Burton, A., & Lewis, J. (2023). Enhancing biofeedback-driven self-guided virtual reality exposure therapy through arousal detection from multimodal data using machine learning. *Brain Informatics*, *10*(1), 14. https://doi.org/10.1186/s40708-023-00193-9

Schroeder, P. A., Gehrer, N. A., Reents, M., Reimer, N., Vagedes, J., & Svaldi, J. (2023). Body Dissatisfaction Directs Avatar Perception: Embodiment and Selective Visual Attention in Body Mass-Modified Self-Avatars. *Cyberpsychology, Behavior, and Social Networking*, *26*(11), 850–860. https://doi.org/10.1089/cyber.2022.0385

Selaskowski, B., Wiebe, A., Kannen, K., Asché, L., Pakos, J., Philipsen, A., & Braun, N. (2024). Clinical adoption of virtual reality in mental health is challenged by lack of high-quality research. *Npj Mental Health Research*, *3*(1), 1–5. https://doi.org/10.1038/s44184-024-00069-8

Shiban, Y., Reichenberger, J., Neumann, I. D., & Mühlberger, A. (2015). Social conditioning and extinction paradigm: A translational study in virtual reality. *Frontiers in Psychology*, *6*, 400. https://doi.org/10.3389/fpsyg.2015.00400

Silva, N., Schreck, T., Veas, E., Sabol, V., Eggeling, E., & Fellner, D. W. (2018). Leveraging eye-gaze and time-series features to predict user interests and build a recommendation model for visual analysis. *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*, 1–9. https://doi.org/10.1145/3204493.3204546

Slater, M., Neyret, S., Johnston, T., Iruretagoyena, G., Crespo, M. Á. de la C., Alabèrnia-Segura, M., Spanlang, B., & Feixas, G. (2019). An experimental study of a virtual reality counselling paradigm using embodied self-dialogue. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-46877-3

Szekely, R., Mason, O., Frohlich, D., & Barley, E. (2024). 'It's not everybody's snapshot. It's just an insight into that world': A qualitative study of multiple perspectives towards understanding the mental health experience and addressing stigma in healthcare students through virtual reality. *DIGITAL HEALTH*, *10*, 20552076231223801. https://doi.org/10.1177/20552076231223801

Takac, M., Collett, J., Blom, K. J., Conduit, R., Rehm, I., & De Foe, A. (2019). Public speaking anxiety decreases within repeated virtual reality training sessions. *PloS One*, *14*(5), e0216288. https://doi.org/10.1371/journal.pone.0216288

Thompson-Hollands, J., Bentley, K. H., Gallagher, M. W., Boswell, J. F., & Barlow, D. H. (2014). Credibility and Outcome Expectancy in the Unified Protocol: Relationship to Outcomes. *Journal of Experimental Psychopathology*, *5*(1), 72–82. https://doi.org/10.5127/jep.033712

Treal, T., Jackson, P. L., Jeuvrey, J., Vignais, N., & Meugnot, A. (2021). Natural human postural oscillations enhance the empathic response to a facial pain expression in a virtual character. *Scientific Reports*, *11*(1), 12493. https://doi.org/10.1038/s41598-021-91710-5

Tsiourti, C., Weiss, A., Wac, K., & Vincze, M. (2019). Multimodal Integration of Emotional Signals from Voice, Body, and Context: Effects of (In)Congruence on Emotion Recognition and Attitudes Towards Robots. *International Journal of Social Robotics*, *11*(4), 555–573. https://doi.org/10.1007/s12369-019-00524-z

Veldmeijer, L., Terlouw, G., Van Os, J., Van Dijk, O., Van 't Veer, J., & Boonstra, N. (2023). The Involvement of Service Users and People With Lived Experience in Mental Health Care Innovation Through Design: Systematic Review. *JMIR Mental Health*, *10*, e46590. https://doi.org/10.2196/46590

Volonte, M., Hsu, Y.-C., Liu, K.-Y., Mazer, J. P., Wong, S.-K., & Babu, S. V. (2020). Effects of Interacting with a Crowd of Emotional Virtual Humans on Users' Affective and Non-Verbal Behaviors. *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 293–302. https://doi.org/10.1109/VR46266.2020.00049

Zibrek, K., Kokkinara, E., & Mcdonnell, R. (2018). The Effect of Realistic Appearance of Virtual Characters in Immersive Environments—Does the Character's Personality Play a Role? *IEEE Transactions on Visualization and Computer Graphics*, *24*(4), 1681–1690. IEEE Transactions on Visualization and Computer Graphics. https://doi.org/10.1109/TVCG.2018.2794638

7 Appendices

7.1 APPENDICES – CHAPTER3

A3.1. 2-way ANOVA assumption testing results

	Homogeneity of Variance				Shapiro-Wilk Normality Test	
Measures	**df1**	**df2**	**F value**	**Pr(>F)**	**w**	**p**
Therapeutic Alliance	3	114	0.527	0.665	0.96	0.001
Treatment Credibility	3	113	0.613	0.608	0.973	0.02
Treatment Expectancy	3	114	0.955	0.417	0.989	0.44
Presence	3	113	2.6	0.056	0.958	0.001
Warmness of Voice	3	110	2.51	0.063	0.921	<0.001

A3.2. 2-way ANOVA results for alliance

Alliance	Df	Sum Sq	Mean Sq	F value	Pr(>F)	Mean Difference with 95% CI	Eta2 (partial) with 95% CI
WarmFace	1	1631	1631.1	12.389	<0.001 ***	7.437 [3.251,11.622]	0.1 [0.02, 0.21]
HeadNod	1	568	568.5	4.318	0.040 *	4.360 [0.205, 8.575]	0.04 [0.00, 0.13]
WarmFace:HeadNod	1	93	92.8	0.705	0.403		0.006 [0.00, 0.06]
Residuals	114	15009	131.7				

A3.3. 2-way ANOVA results for Credibility

Credibility	Df	Sum Sq	Mean Sq	F value	Pr(>F)	Mean Difference with 95% CI	Eta2 (partial) with 95% CI
WarmFace	1	11.9	11.88	0.833	0.363	0.637 [-0.746, 2.021]	0.007 [0.00, 0.07]
HeadNod	1	87.1	87.14	6.110	0.015 *	1.760 [0.343, 3.109]	0.05 [0.00, 0.15]
WarmFace:HeadNod	1	47	46.96	3.293	0.072		0.03 [0.00, 0.11]
Residuals	113	1611.7	14.26				

A3.4. 2-way ANOVA results for Expectancy

Expectancy	Df	Sum Sq	Mean Sq	F value	Pr(>F)	Mean Difference with 95% CI	Eta2 (partial) with 95% CI
WarmFace	1	3.8	3.81	0.150	0.700	0.359 [-1.479, 2.198]	0.001 [0.00, 0.04]
HeadNod	1	154	153.96	6.055	0.015 *	2.284 [0.445, 4.124]	0.05 [0.00, 0.15]
WarmFace:HeadNod	1	30.6	30.56	1.202	0.275		0.01 [0.00, 0.08]
Residuals	114	2898.7	25.43				

A3.5. 2-way ANOVA results for Warmness of Voice

Warmness of Voice	Df	Sum Sq	Mean Sq	F value	Pr(>F)	Mean Difference with 95% CI	Eta2 (partial) with 95% CI
WarmFace	1	5.93	5.93	9.442	0.003 **	0.454 [0.162, 0.750]	0.08 [0.01, 0.19]
HeadNod	1	4.1	4.104	6.535	0.012 *	0.389 [0.085, 0.674]	0.06 [0.00, 0.16]
WarmFace:HeadNod	1	0.99	0.991	1.579	0.212		0.01 [0.00, 0.09]
Residuals	110	69.08	0.628				

A3.6. 2-way ANOVA results for Presence

Presence	Df	Sum Sq	Mean Sq	F value	Pr(>F)	Mean Difference with 95% CI	Eta2 (partial) with 95% CI
WarmFace	1	14.29	14.287	8.119	0.005 **	0.699 [0.213, 1.185]	0.07 [0.01, 0.17]
HeadNod	1	4.66	4.661	2.649	0.106	0.399 [-0.087, 0.885]	0.02 [0.00, 0.10]
WarmFace:HeadNod	1	0.31	0.313	0.178	0.674		0.001 [0.00, 0.05]
Residuals	113	198.86	1.76				

A3.7. Multiple pairwise-comparison between the means of groups

Pairwise Condition	Alliance p-adj	Credibility p-adj	Expectancy p-adj	WarmnessOfVoice p-adj	Presence p-adj
WarmFace-NeutralFace	0.013542 *	0.229	0.743	0.016984 *	0.105
NeutralFacewithNod-NeutralFace	0.169	0.0156883 *	0.142	0.0399989 *	0.465
WarmFacewithNod-NeutralFace	0.0007801 ***	0.090	0.330	0.00073 ***	0.0106219 *
NeutralFacewithNod-WarmFace	0.732	0.683	0.662	0.992	0.825
WarmFacewithNod-WarmFace	0.826	0.969	0.903	0.794	0.838
WarmFacewithNod-NeutralFacewithNod	0.236	0.912	0.967	0.632	0.320

A3.8. Ethics Approval

MEDICAL SCIENCES INTERDIVISIONAL RESEARCH ETHICS COMMITTEE
Research Services, Boundary Brook House, Churchill Drive, Headington, Oxford, OX3 7GB
Tel: +44(0)1865 616575
ethics@medsci.ox.ac.uk

UNIVERSITY OF OXFORD

CONFIDENTIAL

Shu Wei
Department of Psychiatry
University of Oxford
Warneford Hospital
Oxford

21 October 2021

Dear Shu,

Research Ethics Approval - CUREC 1

Ethics Approval Reference: R77367/RE001

Study title: The effect of a virtual coach's facial expression and head nod in virtual reality (VR): experiences of the introductory section of a VR therapy for fear of heights

The above application has been considered on behalf of the Medical Sciences Interdivisional Research Ethics Committee (MS IDREC) in accordance with the University's procedures for ethical approval of all research involving human participants.

I am pleased to inform you that, on the basis of the information provided to the IDREC, the proposed research has been judged as meeting appropriate ethical standards, and approval has been granted for a period of **18 months**, commencing on **21st October 2021.**

Amendments

Should there be any subsequent changes to the study, you should submit details to the MS IDREC for consideration and approval. Details of changes must be listed on an amendment form.

Yours Sincerely

H. Barnby-Porritt

Dr Helen Barnby-Porritt
Research Ethics Manager

A3.9. Screening questionnaires

Heights Interpretation Questionnaire (Steinman and Teachman, 2011)

Situation #1

Imagine that you are climbing a ladder that is leaning against the side of a two-story house. As you move from one rung to the next, you feel the cold metal beneath your hands. You pass a window on the first floor of the house. You continue to climb, feeling the wind on your face. You pass a window on the second floor of the house. You look down and the ground looks very far away.

How likely is it that…

	Not Likely		**Some-what Likely**		**Very Likely**
1. You will hurt yourself.	1	2	3	4	5
2. You will fall.	1	2	3	4	5
3. You will not be able to tolerate your anxiety.	1	2	3	4	5
4. You will panic and lose control.	1	2	3	4	5
5. You are not safe.	1	2	3	4	5
6. You will faint.	1	2	3	4	5
7. You will freeze and not be able to climb back down the ladder.	1	2	3	4	5
8. Being on the ladder is dangerous.	1	2	3	4	5

Situation #2

Imagine that you are on a balcony on the 15th floor of a building. As you hold onto the warm metal railing that comes up to your waist, you feel the heat of the sun on your face. You listen to the sounds of cars and people down below. You look down and the people and cars on the ground seem small and very distant. Even the treetops down below seem far away.

How likely is it that…

	Not Likely		**Some-what Likely**		**Very Likely**
1. You will hurt yourself.	1	2	3	4	5
2. You will fall.	1	2	3	4	5
3. You will not be able to tolerate your anxiety.	1	2	3	4	5
4. You will panic and lose control.	1	2	3	4	5
5. You are not safe.	1	2	3	4	5
6. You will faint.	1	2	3	4	5
7. You will freeze and not be able to get off the balcony.	1	2	3	4	5
8. Being on the balcony is dangerous.	1	2	3	4	5

A3.10. Questionnaires (for measures after VR experience)

Virtual Therapist Alliance Scale (VTAS) (Miloff *et al.*, 2020)

Item	**Do not agree at all**				**Very Likely**
1. I experienced the virtual coach as friendly	0	1	2	3	4
2. I experienced the virtual coach as warming	0	1	2	3	4
3. I felt that the virtual coach gave clear instructions	0	1	2	3	4
4. I experienced the virtual coach as supportive	0	1	2	3	4
5. The presence of the virtual coach made the experience more enjoyable	0	1	2	3	4
6. It felt like the virtual coach shared the virtual environment with me	0	1	2	3	4
7. The virtual coach appeared alive to me	0	1	2	3	4
8. I felt that the virtual coach and I interacted	0	1	2	3	4
9. The way that the virtual coach communicated was captivating	0	1	2	3	4
10. I felt that the virtual coach was trustworthy	0	1	2	3	4
11. It felt comforting to have the virtual coach there with me	0	1	2	3	4
12. The presence of the virtual coach helped me achieve my goals	0	1	2	3	4
13. The virtual coach and I shared common goals	0	1	2	3	4
14. I felt that the virtual coach understood my fears	0	1	2	3	4
15. I felt that the virtual coach tailored the treatment according to my needs and progress	0	1	2	3	4
16. The encouragement of the virtual coach helped me	0	1	2	3	4
17. The virtual coach gave me new perspectives on my troubles	0	1	2	3	4
17. The voice of the virtual coach was warm and friendly.	0	1	2	3	4

Factor 1: consists of item 5-17 from task, goal, and copresence categories.
Factor 2: consists of item 1-5, 7 and 10 from bond and empathy.

Credibility/expectancy questionnaire (Devilly and Borkovec, 2000)

Please answer the questions below. In the first set, answer in terms of what you think. In the second set answer in terms of what you really and truly feel.

Set I

1. At this point, how logical does the VR therapy offered to you seem?

1 2 3 4 5 6 7 8 9

not at all logical somewhat logical very logical

2. At this point, how successfully do you think this VR treatment will be in reducing your fear of height symptoms?

1 2 3 4 5 6 7 8 9

not at all useful somewhat useful very useful

3. How confident would you be in recommending this VR treatment to a friend who experiences similar problems?

1 2 3 4 5 6 7 8 9

not at all confident somewhat confident very confident

4. By the end of the VR therapy period, how much improvement in your fear of height symptoms do you think would occur (if you had it)?

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Set II

For this set, close your eyes for a few moments, and try to identify what you really *feel* about the therapy and its likely success. Then answer the following questions.

1. At this point, how much do you really *feel* that VR therapy will help you to reduce your fear of height symptoms?

1 2 3 4 5 6 7 8 9

not at all somewhat Very much

2. By the end of the VR therapy period (if you had it), how much improvement in your fear of height symptoms do you really *feel* will occur?

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

A.3.11. Questionnaires for paranoid thoughts in real-life

The Revised Green et al., Paranoid Thoughts Scale (R-GPTS) (Freeman *et al.*, 2021)

Please read each of the statements carefully.
They refer to thoughts and feelings you may have had about others over the last month.

Part A Item	**Do not agree at all**		**Somewhat**		**Very Likely**
1. I spent time thinking about friends gossiping about me.	0	1	2	3	4
2. I often heard people referring to me.	0	1	2	3	4
3. I have been upset by friends and colleagues judging me critically.	0	1	2	3	4
4. People definitely laughed at me behind my back.	0	1	2	3	4
5. I have been thinking a lot about people avoiding me.	0	1	2	3	4
6. People have been dropping hints for me.	0	1	2	3	4
7. I believed that certain people were not what they seemed.	0	1	2	3	4
8. People talking about me behind my back upset me.	0	1	2	3	4

Part B Item	**Do not agree at all**		**Somewhat**		**Very Likely**
1. Certain individuals have had it in for me.	0	1	2	3	4
2. People wanted me to feel threatened, so they stared at me.	0	1	2	3	4
3. I was certain people did things in order to annoy me.	0	1	2	3	4
4. I was convinced there was a conspiracy against me.	0	1	2	3	4
5. I was sure someone wanted to hurt me.	0	1	2	3	4
6. I couldn't stop thinking about people wanting to confuse me.	0	1	2	3	4
7. I was distressed by being persecuted.	0	1	2	3	4

8. It was difficult to stop thinking about people wanting to make me feel bad.	0	1	2	3	4
9. People have been hostile towards me on purpose.	0	1	2	3	4
10. I was angry that someone wanted to hurt me.	0	1	2	3	4

A.3.12 Recruitment Material

Poster

Facebook Ads

We are looking for volunteers aged 18 years and over to come in to the University of Oxford for a 45-min virtual reality (VR) study, looking at the introductory experience of VR consultation for fear of heights. Strict measures will be followed to minimise the risk from COVID-19, and you will be reimbursed for your time!

7.2 Appendices - Chapter5

A5.1. 2-way ANCOVA results for VAS Paranoia

VASParanoia	Df	Sum Sq	Mean Sq	F value	Pr(>F)		Eta2 (partial)
BaselinePartB	1	50579	50579	2.678	0.104		0.022
Animation	1	322406	322406	17.071	0.000	***	0.125
Expression	1	74364	74364	3.938	0.049	*	0.033
Animation:Expression	1	47571	47571	2.519	0.115		0.021
Residuals	117	2209669	18886				

A5.2. 2-way ANCOVA results for SSPSPersecutory

	Homogeneity of Variance				Shapiro-Wilk Normality Test	
Measures	**df1**	**df2**	**F value**	**Pr(>F)**	**w**	**p**
Therapeutic Alliance	3	114	0.527	0.665	0.96	0.001
Treatment Credibility	3	113	0.613	0.608	0.973	0.02
Treatment Expectancy	3	114	0.955	0.417	0.989	0.44
Presence	3	113	2.6	0.056	0.958	0.001
Warmness of Voice	3	110	2.51	0.063	0.921	<0.001

SSPS-Persecutiory	Df	Sum Sq	Mean Sq	F value	Pr(>F)		Eta2 (partial)
BaselinePartB	1	394	394.1	3.624	0.059		0.030
Animation	1	1138	1138	10.464	0.002	**	0.081
Expression	1	276	275.6	2.534	0.114		0.021
Animation:Expression	1	206	206.3	1.897	0.171		0.016
Residuals	117	12724	108.8				

A5.3. 2-way ANCOVA results for SSPSNeutral

	Homogeneity of Variance				Shapiro-Wilk Normality Test	
Measures	**df1**	**df2**	**F value**	**Pr(>F)**	**w**	**p**
Therapeutic Alliance	3	114	0.527	0.665	0.96	0.001
Treatment Credibility	3	113	0.613	0.608	0.973	0.02
Treatment Expectancy	3	114	0.955	0.417	0.989	0.44
Presence	3	113	2.6	0.056	0.958	0.001
Warmness of Voice	3	110	2.51	0.063	0.921	<0.001

SSPS-Neutral	Df	Sum Sq	Mean Sq	F value	Pr(>F)		Eta2 (partial)
BaselinePartB	1	10.2	10.2	0.442	0.508		0.002
Animation	1	181	180.99	7.843	0.006	**	0.063
Expression	1	3.9	3.91	0.170	0.681		0.001
Animation:Expression	1	44.2	44.17	1.914	0.169		0.016
Residuals	117	2699.9	23.08				

A5.4. 2-way ANCOVA results for SSPSPositive

	Homogeneity of Variance				Shapiro-Wilk Normality Test	
Measures	**df1**	**df2**	**F value**	**Pr(>F)**	**w**	**p**
Therapeutic Alliance	3	114	0.527	0.665	0.96	0.001
Treatment Credibility	3	113	0.613	0.608	0.973	0.02
Treatment Expectancy	3	114	0.955	0.417	0.989	0.44
Presence	3	113	2.6	0.056	0.958	0.001
Warmness of Voice	3	110	2.51	0.063	0.921	<0.001

SSPS-Positive	**Df**	**Sum Sq**	**Mean Sq**	**F value**	**Pr(>F)**		**Eta2 (partial)**
BaselinePartB	1	1.7	1.73	0.100	0.752		0.000
Animation	1	101	100.96	5.867	0.017	*	0.047
Expression	1	58.5	58.49	3.399	0.068		0.028
Animation:Expression	1	74	73.95	4.297	0.040	*	0.036
Residuals	117	2013.5	17.21				

A5.5. Multiple Pairwise Comparisons Results

	Homogeneity of Variance				Shapiro-Wilk Normality Test	
Measures	**df1**	**df2**	**F value**	**Pr(>F)**	**w**	**p**
Therapeutic Alliance	3	114	0.527	0.665	0.96	0.001
Treatment Credibility	3	113	0.613	0.608	0.973	0.02
Treatment Expectancy	3	114	0.955	0.417	0.989	0.44
Presence	3	113	2.6	0.056	0.958	0.001
Warmness of Voice	3	110	2.51	0.063	0.921	<0.001

Pairwise Condition	**VASParanoia p-adj**	**SPSSPersecutory p-adj**	**SPSSNeutral p-adj**	**SPSSPositive p-adj**
Neutral:Static-Positive:Animated	*0.000*	*0.005*	0.324	*0.016*
Neutral:Static-Neutral:Animated	*0.001*	*0.008*	*0.019*	0.996
Positive:Static-Neutral:Animated	0.451	0.670	0.115	0.981
Positive:Static-Positive:Animated	0.290	0.569	0.745	*0.011*
Positive:Static-Neutral:Static	0.061	0.159	0.902	0.999
Positive:Animated-Neutral:Animated	0.992	0.999	0.586	*0.032*

A5.6. Ethics Approval

MEDICAL SCIENCES INTERDIVISIONAL RESEARCH ETHICS COMMITTEE
Research Services, Boundary Brook House, Churchill Drive, Headington, Oxford, OX3 7GB
Tel: +44(0)1865 616575
ethics@medsci.ox.ac.uk

UNIVERSITY OF OXFORD

CONFIDENTIAL

Dr Aitor Rovira & Shu Wei
Department of Psychiatry
University of Oxford
Warneford Hospital
Oxford

22 March 2023

Dear Dr Rovira and Shu,

Research Ethics Approval - CUREC 1

Ethics Approval Reference: R85111/RE001

Study title: Experience of a virtual reality (VR) social situation: Effects of animated faces and facial expression on appraisal of virtual humans' intentions

The above application has been considered on behalf of the Medical Sciences Interdivisional Research Ethics Committee (MS IDREC) in accordance with the University's procedures for ethical approval of all research involving human participants.

I am pleased to inform you that, on the basis of the information provided to the IDREC, the proposed research has been judged as meeting appropriate ethical standards, and approval has been granted for a period of **18 months**, commencing on **22nd March 2023.**

Amendments

Should there be any subsequent changes to the study, you should submit details to the MS IDREC for consideration and approval. Details of changes must be listed on an amendment form.

Yours Sincerely

DocuSigned by:
Leah Butts
9F14889D2BC549A...

Mrs Leah Butts
Research Ethics Administrator

for
Dr Helen Barnby-Porritt
Research Ethics Manager

A.5.7. Screening questionnaire for paranoid thoughts

The Revised Green et al., Paranoid Thoughts Scale (R-GPTS) (Freeman *et al.*, 2021)

Please read each of the statements carefully.
They refer to thoughts and feelings you may have had about others over the last month.

Part A Item	**Do not agree at all**		**Somewhat**		**Very Likely**
1. I spent time thinking about friends gossiping about me.	0	1	2	3	4
2. I often heard people referring to me.	0	1	2	3	4
3. I have been upset by friends and colleagues judging me critically.	0	1	2	3	4
4. People definitely laughed at me behind my back.	0	1	2	3	4
5. I have been thinking a lot about people avoiding me.	0	1	2	3	4
6. People have been dropping hints for me.	0	1	2	3	4
7. I believed that certain people were not what they seemed.	0	1	2	3	4
8. People talking about me behind my back upset me.	0	1	2	3	4

Part B Item	**Do not agree at all**		**Somewhat**		**Very Likely**
1. Certain individuals have had it in for me.	0	1	2	3	4
2. People wanted me to feel threatened, so they stared at me.	0	1	2	3	4
3. I was certain people did things in order to annoy me.	0	1	2	3	4
4. I was convinced there was a conspiracy against me.	0	1	2	3	4
5. I was sure someone wanted to hurt me.	0	1	2	3	4
6. I couldn't stop thinking about people wanting to confuse me.	0	1	2	3	4
7. I was distressed by being persecuted.	0	1	2	3	4

8. It was difficult to stop thinking about people wanting to make me feel bad.	0	1	2	3	4
9. People have been hostile towards me on purpose.	0	1	2	3	4
10. I was angry that someone wanted to hurt me.	0	1	2	3	4

A.5.8. Questionnaire as visual analogue scales for virtual human appraisals

Thoughts about Virtual Humans

For the first part, please draw a **X** on any part of the blue line to indicate your thoughts.

1. Right now I feel suspicious of the people in the lift.

0 50 100
Not at all Extremely

2. Right now I feel that people in the lift wanted to harm me.

0 50 100
Not at all Extremely

3. Right now I feel like the people in the lift wanted to upset me.

0 50 100
Not at all Extremely

4. Right now I feel like the people in the lift were against me.

0 50 100

Not at all Extremely

5. Right now I am thinking that the people in the lift were trying to persecute me.

0 50 100

Not at all Extremely

6. Right now I feel like the people in the lift were hostile towards me.

0 50 100

Not at all Extremely

A.5.9. Questionnaire for co-presence

Co Presence of Virtual Humans

Co-presence items	**Do not agree at all**		**Somewhat**		**Very Likely**
1. I noticed the virtual humans.	0	1	2	3	4
2. The virtual humans noticed me.	0	1	2	3	4
3. The presence of the virtual humans was obvious to me.	0	1	2	3	4
4. My presence was obvious to the virtual humans.	0	1	2	3	4
5. The virtual humans caught my attention.	0	1	2	3	4
6. I caught the virtual humans' attention.	0	1	2	3	4

Co-presence: the degree to which the observer believes he/she is not alone and secluded, their level of peripheral or focal awareness of the other, and their sense of the degree to which the other is peripherally or focally aware of them.

A.5.10. Questionnaire for state social thoughts

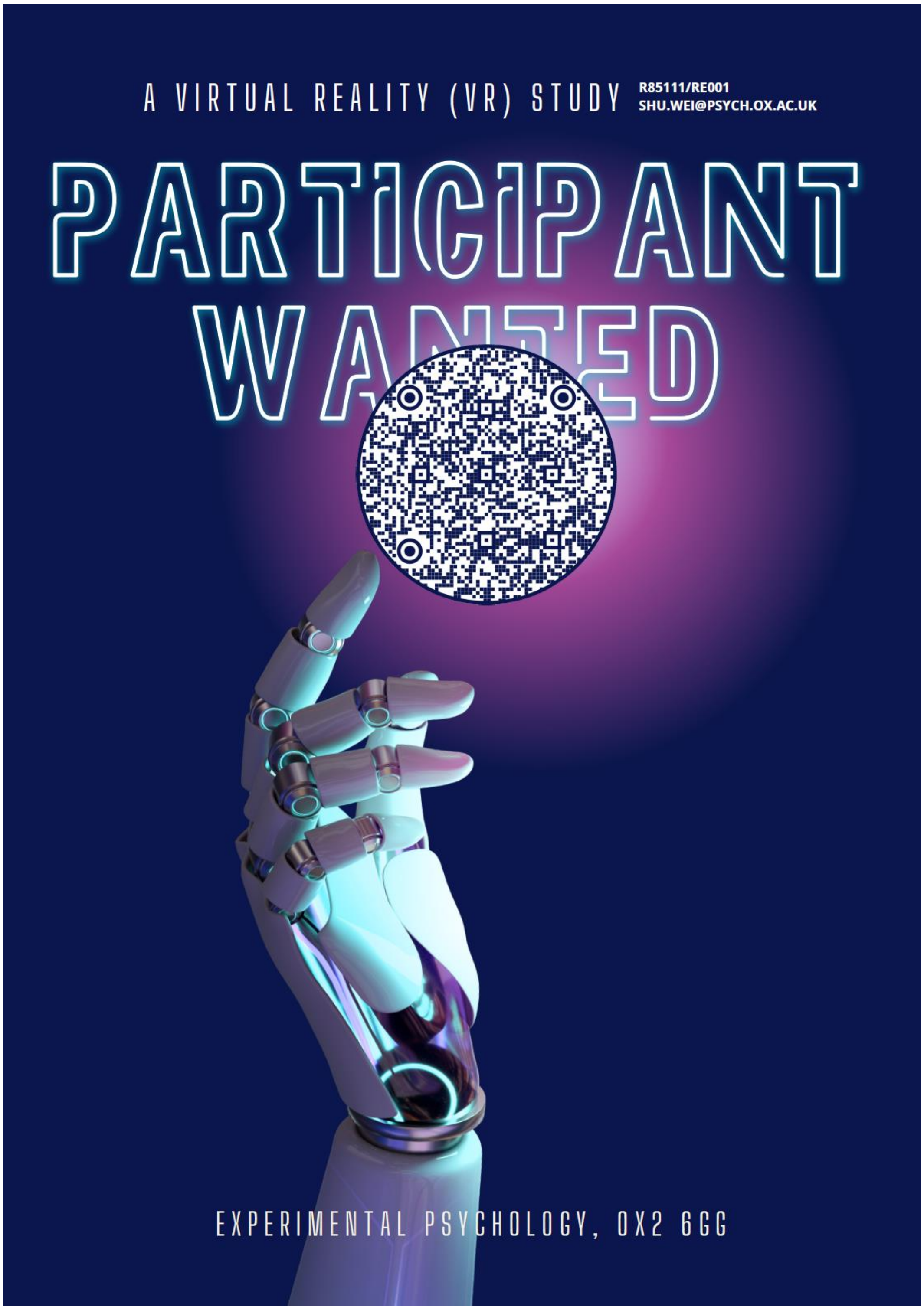
State Social Thoughts

	Not At All				**Totally**
1. Someone was hostile towards me.	0	1	2	3	4
N 2. No-one had any particular feelings about me.	0	1	2	3	4
3. Someone had bad intentions towards me.	0	1	2	3	4
P 4. Someone was friendly towards me.	0	1	2	3	4
5. Someone was trying to make me distressed.	0	1	2	3	4
P 6. I felt very safe in their company.	0	1	2	3	4
7. Someone stared at me in order to upset me.	0	1	2	3	4
P 8. Everyone was trustworthy.	0	1	2	3	4
9. Someone wanted me to feel threatened.	0	1	2	3	4
N 10. I wasn't really noticed by anybody.	0	1	2	3	4
P 11. Someone had kind intentions toward me.	0	1	2	3	4
12. Someone would have harmed me in some way if they could.	0	1	2	3	4
13. Someone had it in for me.	0	1	2	3	4
N 14. Everyone was neutral towards me.	0	1	2	3	4
15. Someone was trying to intimidate me.	0	1	2	3	4
P 16. Everyone was pleasant.	0	1	2	3	4
17. Someone was trying to isolate me.	0	1	2	3	4
N18. No-one had any intentions towards me.	0	1	2	3	4
N 19. Everyone seemed unconcerned by my presence.	0	1	2	3	4
20. Someone was trying to irritate me.	0	1	2	3	4

* P indicates positive thoughts; N indicates neutral thoughts; the rest indicates negative thoughts

A.5.11. Recruitment Material

<u>Poster</u>



Facebook/Meta Ads

Experiences of a virtual reality (VR) social situation: We are looking for participants aged 18 years and over to come to the University of Oxford for a 45-min virtual reality (VR) study. Eligible participants will experience a social situation in VR. You will be reimbursed for your time!

7.3 PUBLISHED MANUSCRIPTS

A. Wei, S., Freeman, D. & Rovira, A. A randomised controlled test of emotional attributes of a virtual coach within a virtual reality (VR) mental health treatment. Sci Rep 13, 11517 (2023). https://doi.org/10.1038/s41598-023-38499-7

B. Wei, S., Freeman, D., Harris, V., & Rovira, A. (2024). A randomised controlled test in virtual reality of the effects on paranoid thoughts of virtual humans' facial animation and expression. Sci Rep 14, 17102. https://doi.org/10.1038/s41598-024-67534-4

C. Wei, S., Bloemers, D., & Rovira, A. (2023). A Preliminary Study of the Eye Tracker in the Meta Quest Pro. Proceedings of the 2023 ACM International Conference on Interactive Media Experiences, 216–221. https://doi.org/10.1145/3573381.3596467

www.nature.com/scientificreports

scientific reports

Check for updates

OPEN

A randomised controlled test of emotional attributes of a virtual coach within a virtual reality (VR) mental health treatment

Shu Wei[1]✉, Daniel Freeman[2,3] & Aitor Rovira[2,3]

We set out to test whether positive non-verbal behaviours of a virtual coach can enhance people's engagement in automated virtual reality therapy. 120 individuals scoring highly for fear of heights participated. In a two-by-two factor, between-groups, randomised design, participants met a virtual coach that varied in warmth of facial expression (with/without) and affirmative nods (with/without). The virtual coach provided a consultation about treating fear of heights. Participants rated the therapeutic alliance, treatment credibility, and treatment expectancy. Both warm facial expressions (group difference = 7.44 [3.25, 11.62], p = 0.001, eta_p^2=0.10) and affirmative nods (group difference = 4.36 [0.21, 8.58], p = 0.040, eta_p^2 = 0.04) by the virtual coach independently increased therapeutic alliance. Affirmative nods increased the treatment credibility (group difference = 1.76 [0.34, 3.11], p = 0.015, eta_p^2 = 0.05) and expectancy (group difference = 2.28 [0.45, 4.12], p = 0.015, eta_p^2 = 0.05) but warm facial expressions did not increase treatment credibility (group difference = 0.64 [− 0.75, 2.02], p = 0.363, eta_p^2 = 0.01) or expectancy (group difference = 0.36 [− 1.48, 2.20], p = 0.700, eta_p^2 = 0.001). There were no significant interactions between head nods and facial expressions in the occurrence of therapeutic alliance (p = 0.403, eta_p^2 = 0.01), credibility (p = 0.072, eta_p^2 = 0.03), or expectancy (p = 0.275, eta_p^2 = 0.01). Our results demonstrate that in the development of automated VR therapies there is likely to be therapeutic value in detailed consideration of the animations of virtual coaches.

Automated virtual reality (VR) therapy is likely to prove a key approach to scale up the delivery of efficacious psychological treatment for mental health difficulties[1,2]. Without reliance on the relatively scarce resource of trained therapists, but with the opportunity for patients to access help in their own homes via the latest standalone consumer headsets, automated VR therapies offer a route to much greater mental health treatment provision. Virtual coaches—who provide instruction, education, encouragement, and feedback to patients—will thus form a crucial element of VR therapy design. In this paper we set out to test two specific characteristics of the virtual coach's non-verbal behaviour that could enhance the VR treatment experience. If characteristics of the virtual reality therapist do affect the patient experience—including markers of better treatment outcomes—then there could be a programme of work testing a range of potentially important factors in their realisation.

Therapeutic alliance, a positive relationship between patient and therapist, is a reliable predictor of better mental health treatment outcomes[3,4], and even affects the efficacy of psychological treatments delivered in digital forms[5–7]. Similarly, patient belief in the credibility of a therapy offered, and expectations of successful outcomes, predict better treatment outcomes[8,9]. Therefore, creating VR coaches that enhance therapeutic alliance and treatment credibility and expectancy could help maximise outcomes from automated VR therapies. Conducting randomized controlled clinical trials to compare treatment outcomes for slight modifications of a virtual coach is not practical, since clinical trials are typically labour and resource intensive studies. Instead, the use of proxy measures for good outcomes, such as therapeutic alliance and treatment credibility and expectancy, provides a pragmatic solution for examining potential treatment effects of variation in a virtual coach.

[1]Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford OX3 7JX, UK. [2]Department of Experimental Psychology, University of Oxford, Oxford, UK. [3]Oxford Health NHS Foundation Trust, Oxford, UK. ✉email: shu.wei@psych.ox.ac.uk

A growing body of research has focused on the experience of virtual humans in coaching and therapies within non-immersive modalities. For example, an early test from Bickmore and Picard[10] compared empathic and neutral versions of a virtual exercise advisor presented on a desktop computer. The empathic advisor displayed caring behaviours, such as direct gaze towards the participant and a concerned facial expression when participants felt unwell. Participants perceived more care from the empathic advisor and were more willing to continue the consultation. Likewise, Lawson and Mayer[11] found that people reported a favourable social connection with a virtual instructor that had a positive voice and body gestures in video coaching. Furthermore, Ter Stal et al.[12] tested the effects of positive facial expressions and response texts of an online virtual coach, who provided tips on physical activity and healthy nutrition. Results showed that positive text responses from the coach, programmed as responses with a greater number of positive words and longer word count, significantly increased participants' perceived rapport with the coach. However, positive facial expressions did not have a significant effect.

Other studies have looked at virtual humans in mental health digital interventions. DeVault et al.[13] created a virtual interview program on a desktop computer, where a virtual interviewer assessed people's distress indicators. They compared two versions of the interviewer—an automated version and a Wizard-of-Oz version in which the human operators triggered the virtual interviewer's spoken and gestural responses. The results showed that people who experienced the Wizard-of-Oz version reported greater rapport, high system usability, and a strong sense that the virtual human was a good listener. Lisetti et al.[14] evaluated an intervention for alcohol dependence delivered with an empathic or non-empathic virtual counsellor presented on a computer screen. Adding empathic qualities (e.g. nodding, smiling, head posture mimicry, and eyebrow movement) led to a higher level of trust in the counsellor and a more significant social influence. On the other hand, Ranjbartabar et al.[15] reported in a study of virtual therapists presented on a computer screen that empathic virtual therapists might not necessarily deliver better emotional outcomes than neutral therapists. Overall, reviews of the use of virtual humans have highlighted the potential benefits of realisation of emotional behaviours in facilitating participant engagement[7,16].

In studies of virtual humans in VR, research has suggested that characters' behavioural realism and positive non-verbal communication can enhance their social impact[17–19]. Wu et al.[18] reported that people perceived stronger social presence and interpersonal attractions when collaborating with a highly expressive virtual human, featuring detailed facial movements and body tracking, compared to a low expressive version. More specifically, non-verbal behaviours such as positive facial expressions with smiles[19] and responsive nodding[17] by characters increases perceived friendliness, trust, and bonding in VR social situations. However, relationships with virtual coaches in automated VR therapies for mental health difficulties have not been experimentally examined. Furthermore, the potential influence of participant factors on the experience of a VR coach is unknown. For instance, individuals who are especially mistrustful in everyday life may find it harder to form a therapeutic alliance with a virtual coach[20], but this has not been tested.

The current study tested the impact of a VR coach's positive non-verbal behaviours (warmth of facial expression, head nodding) on therapeutic alliance and treatment credibility and expectancy for an acrophobia treatment. Additionally, we tested whether a participant's level of mistrust may moderate the relationship with a virtual coach. Our primary hypotheses were that the addition of warm facial expressions and affirmative nods would independently enhance the therapeutic alliance and treatment credibility and expectancy. Further, we hypothesised that the combined use of warm facial expressions and affirmative nods would have the strongest positive effect (i.e. there would be a significant interaction).

Methods

Experimental design. A balanced two-by-two factorial between-groups experimental design was used. The two factors were warm facial expression (with/without, i.e. neutral face) and affirmative head nods (with/without). Therefore, participants were randomised to one of four virtual coach conditions: (1) neutral face (2) neutral face and affirmative nods (3) warm facial expressions and (4) warm facial expressions and affirmative nods. In all experimental conditions the virtual coach's facial expression included basic behaviours such as eye blinking and lip syncing. The study was single-blind. Participants were unaware of the study hypotheses or that they were being randomised to interact with one of the different versions of the virtual coach.

We calculated a target sample size for a between-factors ANOVA using *G*power 3.1*[21]. We specified a medium effect size of partial eta-squared = 0.06 and conventional values of power = 0.80 and $\alpha = 0.05$. A total of 120 participants (30 per condition) would be needed. A randomization list was created using *Research Randomizer*[22].

Participants and recruitment. Participants were primarily recruited via social media advertisements in Oxfordshire. We screened for fear of heights using the *Heights Interpretation Questionnaire (HIQ)*[23](HIQ score > 29, as used in our trial of automated VR therapy for acrophobia[1]) among the general population. Exclusion criteria were individuals who were (a) under 18 years of age, or who reported (b) having photosensitive epilepsy or a significant visual, hearing or mobility impairment that meant that they would not be able to use VR or (c) taking medication which can cause motion sicknesses.

Ethical approval was received from the University of Oxford Medical Sciences Interdivisional Research Ethics Committee. The study was performed in accordance with relevant guidelines and regulations and written informed consent was obtained from all participants. 120 participants (female = 66, male = 50, non-binary = 4) with a mean age of 44.4 (SD = 16.4) took part in the in-person VR study. Participants had a mean fear of heights score of 43.8 (SD = 10.8). Table 1 presents a summary of participant characteristics.

Apparatus and VR scenario. We used a Windows 10 computer (Intel i7-8700K, Nvidia GeForce GTX 1080Ti, 32 GB RAM) to run the VR scenario and render it on a Meta Quest 2 (Meta, formerly Facebook, 2022)

	Neutral face (n = 30)	Neutral face with nod (n = 30)	Warm face (n = 30)	Warm face with nod (n = 30)
Age in years, mean (SD), range	40.7 (16.6), 18–70 (range)	48.2 (16.4), 18–72 (range)	41.1 (17), 19–77 (range)	47.8 (14.9), 24–74 (range)
Gender Female (F), male (M), non-binary (NB)	15 F/13 M/2 NB	15 F/15 M	18 F/12 M	18 F/10 M/2 NB
Fear of heights scores (HIQ scores)	43.8 (10.5)	43.7 (10.8)	43.9 (11.2)	43.8 (11.0)
VR experience	1.97 (0.85)	1.63 (0.72)	1.80 (0.96)	1.80 (0.85)

Table 1. Participant characteristics by randomisation group.

through a wireless connection (Air Link). This VR headset resolution is 1832 × 1920 pixels per eye and was set up at a 90 Hz refresh rate.

We developed the VR experience in Unity game engine, version 2020.3.22. The experience consisted of an indoor scene where participants met the virtual coach for the first time (Fig. 1a) and then they were taken to an outdoor area for a walking task (Fig. 1b). A video of the VR experience is provided as supplementary data.

Indoor scene. The indoor scene was a standing experience. Participants faced the virtual coach for an introductory consultation. The consultation script was from our previous VR fear of heights trial[1]. The virtual coach first introduced herself and explained the cognitive approach to understanding fear of heights (e.g. "The reason we're afraid of heights is because we think something bad is going to happen. And that makes us feel anxious. Then

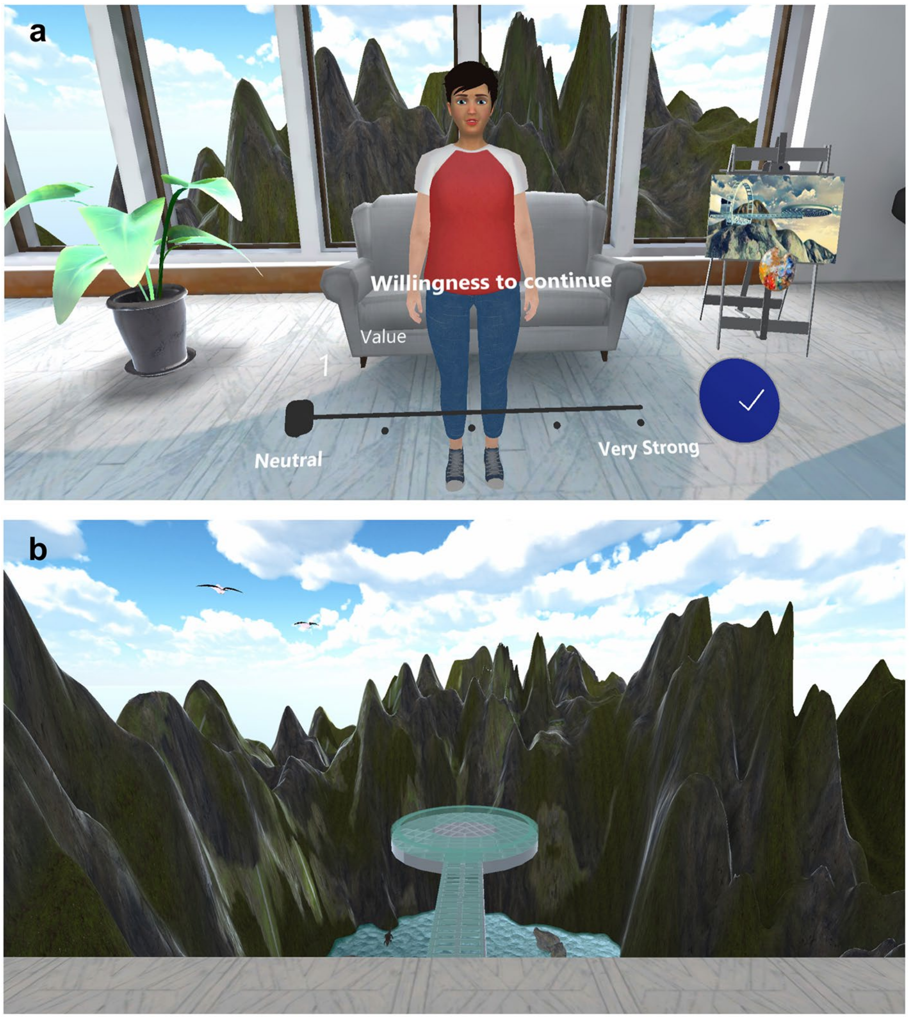


Figure 1. Screenshots of the VR experience. (**a**) Indoor scene: the virtual coach provided an introductory consultation about fear of heights and its treatment. The scene ended with a question about willingness to continue the VR therapy. (**b**) Outdoor scene: participants were instructed to step out on a glass-floor walkway.

we end up avoiding heights because they feel so scary"). The coach then asked participants questions related to their own fears about heights. Participants answered the questions through a UI interface. They went through this interactive conversation at their own pace, which typically took around 4 min.

Outdoor scene. The outdoor scene was also a standing experience in which participants had to walk along an elevated walkway. They started in the middle of a virtual terrace to receive instructions from the virtual coach. The task involved stepping on the walkway, walk until reaching a circular platform, and return to the terrace. The scene concluded once the task was completed or if the participant decided to end it before completion.

We combined the use of motion capture, blend-shape and bone animation to create realistic facial expressions and nods for the virtual coach[24]. A female psychologist was invited as the voice and facial motion actor. The animations were recorded and processed using *Iclone7*[25] with the *LiveFace* plugin. We ran a pilot test with 12 individuals to verify our character animations of the warm facial expressions and affirmative nods.

Experimental procedures. Participants were invited for a single session at our VR lab. They were informed that they would try the introductory part of a VR therapy for fear of heights. After obtaining written consent to participate in the study, the researcher first demonstrated the use of VR and helped participants fit the VR headset. Later, the researcher selected the parameters for the VR experience according to each participant's condition group and they experienced the indoor scene. Once that stage ended, participants took the VR headset off and completed the measures of therapeutic alliance, warmness of voice, treatment credibility/expectancy, and presence. The outdoor scene was a virtual heights experience and could elicit anxious feelings for people with fear of heights. The researcher made sure that participants knew beforehand they could stop the VR scene at any time. Participants experienced the outdoor VR scene and then completed the presence and mistrust questionnaires.

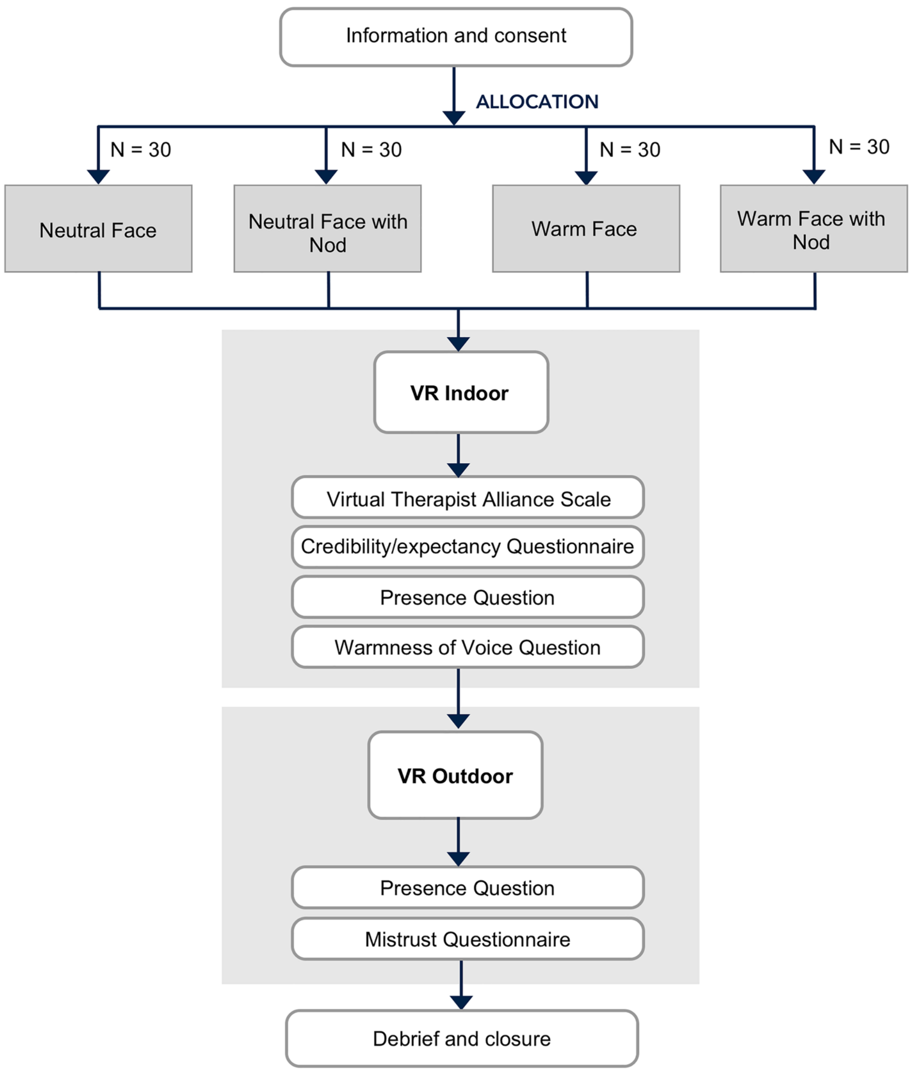


Figure 2. Study procedure.

Finally, they were fully debriefed about the purpose of the study. The entire session lasted approximately 45 min, and participants were reimbursed for their time. Figure 2 shows a summary of the procedure.

Measures. *Therapeutic alliance.* Alliance with the virtual coach was measured by the *Virtual Therapist Alliance Scale (VTAS)*[6]. It is a 17-item self-report questionnaire describing the perception and relationship with the therapist, such as "The way that the virtual coach communicated was captivating" and "The virtual coach gave me new perspectives on my troubles". All items are scored from 0 (Do not agree at all) to 4 (Agree completely) using the same response format with total scores ranging from 0 to 68. Higher scores reflect a stronger alliance with the virtual coach. The measure had very high internal reliability in this study (Cronbach's $\alpha = 0.94$, $N = 120$).

Treatment credibility/expectancy. Treatment credibility and improvement expectancy of the VR fear of heights treatment was measured by the *Credibility/expectancy questionnaire (CEQ)*[26]. It is a six-item questionnaire assessing two factors credibility (three items) and expectancy (three items) separately. Each item is rated in a Likert scale and computed to a score from 1 to 9 (responses to the fourth and the sixth item were linear interpolated from 0 to 100% to 1 to 9), with total scores ranging from 3 to 27 for each factor. Both factors had good internal reliability in this study (credibility: Cronbach's $\alpha = 0.81$; expectancy: Cronbach's $\alpha = 0.89$).

Mistrust. Level of mistrust was measured by *The Revised Green* et al., *Paranoid Thoughts Scale (R-GPTS)*[27]. It is an 18-item scale assessing ideas of persecution, such as "I have been thinking a lot about people avoiding me" and "I was certain people did things in order to annoy me". All items are scored from 0 (do not agree at all) to 4 (Totally) with total scores ranging from 0 to 72. Higher scores reflect higher levels of mistrust. The measure had very high internal reliability in this study (Cronbach's $\alpha = 0.92$).

Fear of heights. Fear of heights was measured by the *Heights Interpretation Questionnaire (HIQ)*[23]. It is a 16-item self-report questionnaire predicting subjective distress and avoidance of heights. The items assess people's anxious fears such as the fear of falling or getting hurt, when imagining two height situations (i.e. being on a ladder against a two-story house and on the balcony of a 15th-floor building). The total score ranges from 16 to 80. The measure had good internal reliability in this study (Cronbach's $\alpha = 0.88$).

Presence. We used a single item from the *Igroup Presence Questionnaire*[28] to measure sense of presence ("In the computer-generated world I had a sense of 'being there'"). This measure was simply used to check that both the scenarios led to participants feeling like they were in the virtual environment. The item is scored on a 5-point Likert scale, from 1 (Not at all) to 5 (Very much).

Warmness of voice. We used a single item to measure the perceived warmness of voice ("The voice of the virtual coach was warm and friendly"). The item is scored from 0 (Do not agree at all) to 4 (Agree completely).

VR behavioural data. We recorded participants' tracking data (position and rotation) in VR. For the virtual walking task in outdoor VR, we also marked the timestamp and duration corresponding to the key events (step on the walkway, reach the circular platform, back to the terrace).

Statistical methods. We first checked that the data were suitable for two-way analysis of variance (ANOVA), using Levene's test for homogeneity of variance and Shapiro–Wilk test of normality (see Supplementary Table S1). The homogeneity of variance was satisfied for all the variables, while the normality assumption was not met for therapeutic alliance, treatment expectancy, presence and warmness of voice. We maintained the original data without transformations due to the robustness of ANOVA to deviations from normality and the sufficient sample size[29].

To assess the effects of warm facial expressions and affirmative head nods on the therapeutic alliance, treatment credibility and expectancy and other subjective measures, we used a two-way ANOVA test with interaction. The partial eta-squared (eta_p^2) was computed to measure effect sizes. Tukey's honest significant difference test (Tukey's HSD) was used for multiple pairwise comparisons. All tests for significance were made at the $\alpha = 0.05$ level. We report the results as mean differences and 95% confidence interval (95% CI) of the difference between conditions.

To assess whether mistrust would moderate the effect of warm facial expressions and affirmative head nods on therapeutic alliance, we used a multiple regression model with the interaction $VirtualCoachAlliance = WarmFace + AffirmativeNod + AffirmativeNod \times Mistrust + Mistrust \times WarmFace$. We evaluated the moderating effect based on the significance of the regression coefficient for the interaction term.

Data cleaning and processing was performed using *Python*'s *Pandas* and *NumPy* libraries[30,31]. Analyses were conducted using *R* with *RStudio 1.4*[32].

Results

Figure 3 shows the raw data box plots for the primary measures of therapeutic alliance, treatment credibility, and expectancy. Descriptive statistics for the measures are shown in Table 2. Apart from two sets of incomplete responses for treatment expectancy items, there were no other missing data. The full details of the analyses can be found in Supplementary Tables S2–S7.

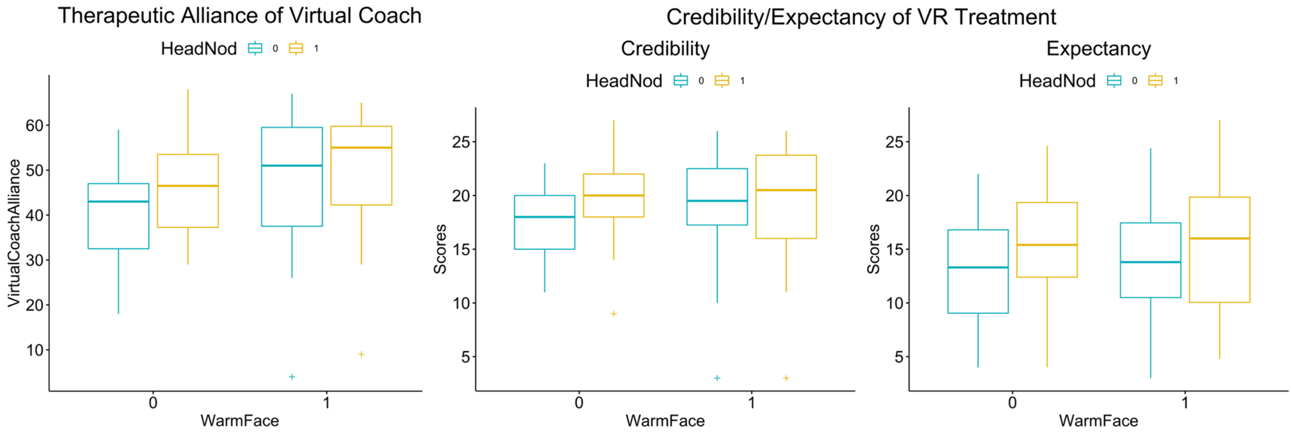


Figure 3. Boxplots of the therapeutic alliance, treatment credibility, and expectancy scores with the two-factor breakdown (0—without, 1—with). Crosses indicate the outlier points, detected at 1.5 times the interquartile range above the upper quartile and below the lower quartile.

	Neutral face	Neutral face with nod	Warm face	Warm face with nod
Measures	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**
Therapeutic alliance	40.1 (11.5)	46.2 (10.5)	47.8 (14.9)	50.5 (13.5)
Credibility	17.5 (3.4)	20.1 (4.0)	18.8 (4.9)	19.3 (5.1)
Expectancy	12.8 (5.0)	15.7 (4.9)	14.1 (4.8)	15.1 (6.0)
Presence	7.7 (1.7)	8.2 (1.4)	8.3 (1.2)	8.6 (1.4)
Warmness of the voice	2.8 (1.1)	3.1 (1.1)	3.2 (1.0)	3.5 (0.8)
Mistrust	10.5 (10.3)	9.4 (9.2)	13.8 (11.9)	8.5 (11.5)

Table 2. Descriptive data of measures by randomization group.

Therapeutic alliance. We removed two extreme outliers (< Q1–3 × IQR) before the two-way ANOVA statistical test. Simple main effects analysis showed that warm facial expressions (group difference = 7.44, 95% CI [3.25, 11.62], F(1, 114) = 12.389, p < 0.001, eta_p^2 = 0.10) and affirmative nods (group difference = 4.36, 95% CI [0.21, 8.58], F(1, 114) = 4.318, p = 0.040, eta_p^2 = 0.04) led to significant increases in therapeutic alliance. There was no significant interaction between warm facial expressions and affirmative nods (F(1, 114) = 0.705, p = 0.403, eta_p^2 = 0.01). Tukey's HSD Test for multiple comparisons found that therapeutic alliance was significantly greater in the warm face compared to the neutral face condition (p-adj = 0.014) and in the warm face with nod compared to the neutral face condition (p-adj < 0.001).

Treatment credibility and expectancy. Simple main effects analysis showed that affirmative nods (group difference = 1.76, 95% CI [0.34, 3.11], F(1, 113) = 6.11, p = 0.015, eta_p^2 = 0.05) led to significant increases in treatment credibility but that warm facial expressions did not (group difference = 0.64, 95% CI [– 0.75, 2.02], F(1, 113) = 0.833, p = 0.363, eta_p^2 = 0.01). There was no statistically significant interaction between warm facial expressions and affirmative nods (F(1, 113) = 3.293, p = 0.072, eta_p^2 = 0.03), although there was a trend in the direction of the combination leading to greater credibility ratings. Tukey's HSD Test for multiple comparisons found that credibility was significantly greater in the neutral face with nod condition compared to the neutral face condition (p-adj = 0.016).

Two participants had incomplete data completion for the expectancy items and were removed from the statistical analysis. Simple main effects analysis showed that affirmative nods (group difference = 2.28, 95% CI [0.45, 4.12], F(1, 114) = 6.055, p = 0.015, eta_p^2 = 0.05) led to a significant increase in expectancy but that warm facial expressions did not (group difference = 0.36, 95% CI [– 1.48, 2.20], F(1, 114) = 0.833, p = 0.700, eta_p^2 = 0.001). There was no statistically significant interaction between warm facial expressions and affirmative nods (F(1, 114) = 1.202, p = 0.275, eta_p^2 = 0.01). Tukey's HSD Test for multiple comparisons found that mean expectancy was not significantly different between the groups.

Moderator effect of mistrust. A multiple regression was used to predict therapeutic alliance by the variables of warm facial expression, affirmative nods, and their interaction with mistrust (F(5, 112) = 4.21, p = 0.002,

R^2 = 15.83). Both interaction terms *WarmFace*Mistrust* ($p = 0.961$) and *AffirmativeNods*Mistrust* ($p = 0.971$) were not statistically significant, suggesting mistrust did not moderate the effects.

Presence. A two-way ANOVA showed that warm facial expressions (group difference = 0.70, 95% CI [0.21, 1.19], $F(1, 113) = 8.119$, $p = 0.005$, $eta_p^2 = 0.07$) led to significantly higher levels of presence but that affirmative nods did not (group difference = 0.40, 95% CI [− 0.09, 0.89], $F(1, 113) = 2.649$, $p = 0.106$, $eta_p^2 = 0.02$). There was no significant interaction between warm facial expressions and affirmative nods ($F(1, 113) = 0.178$, $p = 0.674$, $eta_p^2 = 0.001$). Tukey's HSD Test for multiple comparisons found that the presence was significantly greater in the warm face with nod compared to the neutral face condition (p-adj = 0.011).

Warmness of voice. A two-way ANOVA showed that warm facial expressions (group difference = 0.45, 95% CI [0.16, 0.75], $F(1, 110) = 9.44$, $p = 0.003$, $eta_p^2 = 0.08$) and affirmative nods (group difference = 0.39, 95% CI [0.09, 0.67], $F(1, 110) = 6.54$, $p = 0.01$, $eta_p^2 = 0.06$) led to significantly higher ratings of voice warmness. There was no significant interaction for the combined effects of warm facial expressions and affirmative nods ($F(1, 110) = 1.579$, $p = 0.212$, $eta_p^2 = 0.01$). Tukey's HSD Test for multiple comparisons found that the warmness of the voice was significantly greater in the warm face with nod compared to the neutral face condition (p-adj < 0.001), the warm face condition compared to the neutral face condition (p-adj = 0.017) and the neutral face with nod condition compared to the neutral face condition (p-adj = 0.040).

Behavioural data. We conducted an exploratory analysis of participants' walking task performance across the virtual height. Table 3 shows the summary statistics. 82 out of 120 participants (68.33%) completed the task. The average time to move forward and step on to the walkway was 38.0 s (SD = 47.1), and the average duration spent in outdoor VR after the task brief was 113.1 s (SD = 82.0). We also calculated the normalized walking distance based on the horizontal distance of the virtual walkway. Two sets of data were excluded; one participant experienced a VR connection loss and another opted out of the walking task in the outdoor scene. A two-way ANOVA suggested that warm facial expressions ($p = 0.187$) and affirmative nods ($p = 0.374$) did not have statistically significant effects on walking distance. Similarly, warm facial expressions and affirmative nods did not have statistically significant effects on the time to step on to the virtual walkway (warm facial expressions: $p = 0.356$, affirmative nods: $p = 0.978$) and the time spent in the outdoor scene (warm facial expressions: $p = 0.732$, affirmative nods: $p = 0.511$).

Discussion

Virtual coaches are a key element in automated VR therapies for mental health disorders. We investigated whether introducing positive non-verbal behaviours to the coach increased the therapeutic alliance and treatment credibility and expectancy. Our results partly support our initial hypotheses. We hypothesised that warm facial expressions and affirmative head nods would enhance the therapeutic alliance, treatment credibility, and expectancy, and their combination would have the strongest impact. The results showed that warm facial expressions and affirmative head nods individually affected therapeutic alliance, and the impact of warm facial expressions was more substantial. Additionally, affirmative head nods increased people's beliefs in both the credibility of the treatment and the expectancy of good outcomes. Although there was no significant interaction between warm facial expressions and affirmative head nods, there was a trend in the direction that the combination led to greater treatment credibility. In essence, how a virtual coach is programmed affects the treatment experience and potentially therapeutic outcomes. In this study we showed that there is likely to be value in implementing facial expressions and positive non-verbal behaviours for the virtual coach.

The primary finding that warm facial expressions and affirmative head nods increase alliance is in line with previous studies of virtual humans outside of the context of VR mental health treatment[11,17,19,33,34]. Similar to the conclusion from Oh et al.[35] that virtual agents' facial expressions contribute more than body movements (such as raising of hands and head tilts), the effect size of warm facial expressions of the virtual coach in the current study on the therapeutic alliance was larger than affirmative head nods. Unexpectedly, we did not detect a main effect of warm facial expressions on treatment credibility or expectancy. However, when warm facial expressions were combined with affirmative head nods, there was a trend towards higher credibility ratings. This result might be due to the head nods giving the impression that the therapist was attentively listening and acknowledging participant responses[36]. Such an impression could have then enhanced the potential positive effects of warm facial

	Task completion	Normalized distance	Duration-StepOnWalkway (s)	Duration-outdoor (s)
	Number (%)	Mean (SD)	Mean (SD)	Mean (SD)
All groups	82 (68.33%)	4.65 (2.73)	38.0 (47.1)	113.1 (82.0)
Neutral face	21 (70.00%)	4.78 (2.68)	36.1 (60.9)	111.5 (93.8)
Neutral face with nod	17 (56.67%)	3.86 (3.00)	50.1 (53.1)	109.5 (71.1)
Warm face	22 (73.33%)	4.96 (2.62)	40.4 (40.1)	124.6 (87.0)
Warm face with nod	22 (73.33%)	4.99 (2.55)	27.5 (28.6)	106.7 (77.2)

Table 3. Summary statistics of the VR walking task.

expressions on treatment credibility when they were displayed simultaneously. Interestingly, positive non-verbal behaviours also led to positive voice perception, which highlights an interplay between perceptions of different sensory traits of virtual humans. In this study, we presented two plausible examples of virtual coach's behaviours (i.e. facial expressions and head nods) to demonstrate their impact on mental health treatment. Future research could examine other attributes (e.g. visual, auditory, and other non-verbal behaviours such as eye gaze and hand gestures) and their interactive effects.

Our main focus was the effect of characteristics of a coach on established proxies for good therapeutic outcomes. But we also took an exploratory look at potential effects on participants' behaviours in relation to virtual heights. Approximately one-third of participants did not complete the circuit out to the virtual height and back again. There was no significant difference in the task completion rate, or the distance covered, between the groups allocated to different virtual coach conditions. Since this was the participants' initial exposure to virtual heights, as opposed to the multiple immersions experienced during a full therapy session, we did not make any specific predictions. It would be plausible that the relationship with the virtual coach would make no noticeable difference as patients obtain their first experience of the treatment technique. Indeed, no group differences were detected in whether a person stepped onto the platform or the distance covered.

The study has several limitations. First, we do not know whether the effects of the non-verbal behaviours do translate to better outcomes. This would require a clinical trial to provide evidence. Our view is that using proxies of good outcomes such as therapeutic alliance and treatment credibility is a more sensible testing strategy than conducting multiple clinical trials on small changes to a programme. When such treatments get used at scale then it may be possible to look at outcome effects by programming modifications. Second, we only focused on the virtual coach's facial expressions and head nods and did not account for factors such as gender, ethnicity, and age of the participants. Previous research indicates that people tend to have stronger bonds with virtual humans with similar characteristics as the person[37]. In the future it is likely that people will be able to customize the appearance, style, and even animations of their virtual coach, which could be studied in relation to therapeutic alliance. Third, we used single-blind testing, with the experimenter being aware of a participant's allocated condition since there was only one experimenter running the study. This design choice may have introduced potential bias during the conduct of the experiment, including the experimenter's greeting style, which could have subsequently influenced participants' subjective ratings. Fourth, mistrust was measured at the end of testing, and this may have affected ratings, and therefore was not actually a true moderator variable. However, there was no clear evidence that mistrust was linked to perceptions of the therapeutic alliance or treatment credibility or expectancy. Finally, the violation of normality in the two-way ANOVA can result in overestimating test significance and increase the chances of Type I error. For example, the p-value of 0.04 for the relationship between nodding and alliance is close to the significance threshold, indicating that a larger sample size will be needed for more robust conclusions.

In this study we investigated the effects of a virtual coach's positive non-verbal behaviours during an automated VR consultation for the treatment of the fear of heights. The inclusion of warm facial expressions and affirmative head nods independently increased therapeutic alliance. Furthermore, affirmative head nods by the virtual coach improved perceptions of treatment credibility and positive outcome expectancy. The findings highlight the potential to enhance the experience and effectiveness of VR therapies through tailored VR character design. While our study focused on the cognitive treatment of fear of heights, further study is needed to examine the degree to which there is generalization to other mental health difficulties and different treatment techniques. The development of VR therapies would benefit from a systematic programme of research of the best attributes of virtual coaches, which may vary depending on the conditions and treatment techniques, and require strong collaborations between clinical staff, people with lived experiences, and software developers.

Data availability
Deidentified data are available from the corresponding authors on reasonable request and contract with the university.

Received: 17 November 2022; Accepted: 10 July 2023
Published online: 17 July 2023

References
1. Freeman, D. *et al.* Automated psychological therapy using immersive virtual reality for treatment of fear of heights: A single-blind, parallel-group, randomised controlled trial. *Lancet Psychiatry* **5**, 625–632 (2018).
2. Freeman, D. *et al.* Automated virtual reality therapy to treat agoraphobic avoidance and distress in patients with psychosis (gameChange): A multicentre, parallel-group, single-blind, randomised, controlled trial in England with mediation and moderation analyses. *Lancet Psychiatry* **9**, 375–388 (2022).
3. Ardito, R. B. & Rabellino, D. Therapeutic alliance and outcome of psychotherapy: Historical excursus, measurements, and prospects for research. *Front. Psychol.* **2**, 270 (2011).
4. Horvath, A. O. & Symonds, B. D. Relation between working alliance and outcome in psychotherapy: A meta-analysis. *J. Couns. Psychol.* **38**, 139–149 (1991).
5. Sagui-Henson, S. J. *et al.* Understanding components of therapeutic alliance and well-being from use of a global digital mental health benefit during the COVID-19 pandemic: Longitudinal observational study. *J. Technol. Behav. Sci.* https://doi.org/10.1007/s41347-022-00263-5 (2022).
6. Miloff, A. *et al.* Measuring alliance toward embodied virtual therapists in the era of automated treatments with the virtual therapist alliance scale (VTAS): Development and psychometric evaluation. *J. Med. Internet Res.* **22**, e16660 (2020).
7. Mitruț, O., Moldoveanu, A., Petrescu, L., Petrescu, C. & Moldoveanu, F. A review of virtual therapists in anxiety and phobias alleviating applications. In *Virtual, Augmented and Mixed Reality* (eds. Chen, J. Y. C. & Fragomeni, G.) 71–79 (Springer International Publishing, 2021). https://doi.org/10.1007/978-3-030-77599-5_6.
8. Thompson-Hollands, J., Bentley, K. H., Gallagher, M. W., Boswell, J. F. & Barlow, D. H. Credibility and outcome expectancy in the unified protocol: Relationship to outcomes. *J. Exp. Psychopathol.* **5**, 72–82 (2014).

9. Schulte, D. Patients' outcome expectancies and their impression of suitability as predictors of treatment outcome. *Psychother. Res.* **18**, 481–494 (2008).
10. Bickmore, T. W. & Picard, R. W. Towards caring machines. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* 1489–1492 (ACM, 2004). https://doi.org/10.1145/985921.986097.
11. Lawson, A. P. & Mayer, R. E. Does the emotional stance of human and virtual instructors in instructional videos affect learning processes and outcomes?. *Contemp. Educ. Psychol.* **70**, 102080 (2022).
12. ter Stal, S., Jongbloed, G. & Tabak, M. Embodied conversational agents in eHealth: How facial and textual expressions of positive and neutral emotions influence perceptions of mutual understanding. *Interact. Comput.* **33**, 167–176 (2021).
13. DeVault, D. *et al.* SimSensei kiosk: A virtual human interviewer for healthcare decision support. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems* 1061–1068 (International Foundation for Autonomous Agents and Multiagent Systems, 2014).
14. Lisetti, C., Amini, R., Yasavur, U. & Rishe, N. I Can Help You Change! An empathic virtual agent delivers behavior change health interventions. *ACM Trans. Manag. Inf. Syst.* **4**, 1–28 (2013).
15. Ranjbartabar, H., Richards, D., Bilgin, A. A. & Kutay, C. First impressions count! The role of the human's emotional state on rapport established with an empathic versus neutral virtual therapist. *IEEE Trans. Affect. Comput.* **12**, 788–800 (2021).
16. Elshan, E., Zierau, N., Engel, C., Janson, A. & Leimeister, J. M. Understanding the design elements affecting user acceptance of intelligent agents: Past, present and future. *Inf. Syst. Front.* **24**, 699–730 (2022).
17. Aburumman, N., Gillies, M., Ward, J. A. & de Hamilton, A. F. C. Nonverbal communication in virtual reality: Nodding as a social signal in virtual interactions. *Int. J. Hum.-Comput. Stud.* **164**, 102819 (2022).
18. Wu, Y., Wang, Y., Jung, S., Hoermann, S. & Lindeman, R. W. Using a fully expressive avatar to collaborate in virtual reality: Evaluation of task performance, presence, and attraction. *Front. Virtual Real.* **2**, (2021).
19. Volonte, M. *et al.* Effects of interacting with a crowd of emotional virtual humans on users' affective and non-verbal behaviors. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* 293–302 (2020). https://doi.org/10.1109/VR46266.2020.00049.
20. Trotta, A., Kang, J., Stahl, D. & Yiend, J. Interpretation bias in paranoia: A systematic review and meta-analysis. *Clin. Psychol. Sci.* **9**, 3–23 (2021).
21. Faul, F., Erdfelder, E., Buchner, A. & Lang, A.-G. Statistical power analyses using G*Power 31: Tests for correlation and regression analyses. *Behav. Res. Methods* **41**, 1149–1160 (2009).
22. Urbaniak, G. & Plous, S. Research randomizer (version 4.0)[computer software]. 2013. (2013).
23. Steinman, S. A. & Teachman, B. A. Cognitive processing and acrophobia: Validating the heights interpretation questionnaire. *J. Anxiety Disord.* **25**, 896–902 (2011).
24. Chapter Seven. Facial and Dialogue Animation. In *Digital Character Animation 3* vol. Chapter Seven. Facial and Dialogue Animation.
25. 3D Animation Software for Character Animator | iClone.
26. Devilly, G. J. & Borkovec, T. D. Psychometric properties of the credibility/expectancy questionnaire. *J. Behav. Ther. Exp. Psychiatry* **31**, 73–86 (2000).
27. Freeman, D. *et al.* The revised Green et al., Paranoid Thoughts Scale (R-GPTS): Psychometric properties, severity ranges, and clinical cut-offs. *Psychol. Med.* **51**, 244–253 (2021).
28. Schubert, T., Friedmann, F. & Regenbrecht, H. The experience of presence: Factor analytic insights. *Presence Teleoperators Virtual Environ.* **10**, 266–281 (2001).
29. Sawyer, S. F. Analysis of variance: The fundamental concepts. *J. Man. Manip. Ther.* **17**, 27E-38E (2009).
30. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585**, 357–362 (2020).
31. McKinney, W. Data structures for statistical computing in python. 56–61 (2010). https://doi.org/10.25080/Majora-92bf1922-00a.
32. RStudio Team. RStudio: Integrated Development Environment for R. (2021).
33. Oh, S. Y., Bailenson, J., Krämer, N. & Li, B. Let the avatar brighten your smile: Effects of enhancing facial expressions in virtual environments. *PLoS ONE* **11**, e0161794 (2016).
34. Osugi, T. & Kawahara, J. I. Effects of head nodding and shaking motions on perceptions of likeability and approachability. *Perception* **47**, 16–29 (2018).
35. Oh Kruzic, C., Kruzic, D., Herrera, F. & Bailenson, J. Facial expressions contribute more than body movements to conversational outcomes in avatar-mediated virtual environments. *Sci. Rep.* **10**, 20626 (2020).
36. Inoue, M., Irino, T., Furuyama, N. & Hanada, R. Observational and accelerometer analysis of head movement patterns in psychotherapeutic dialogue. *Sensors* **21**, 3162 (2021).
37. Gamberini, L., Chittaro, L., Spagnolli, A. & Carlesso, C. Psychological response to an emergency in virtual reality: Effects of victim ethnicity and emergency type on helping behavior and navigation. *Comput. Hum. Behav.* **48**, 104–113 (2015).

Acknowledgements
This research has been funded through the Oxford Cognitive Approaches to Psychosis (O-CAP) research fund. The script for the fear of heights consultation was used with permission from Oxford VR's fear of heights VR therapy. Oxford VR is a University of Oxford spin-out company. We thank André Lages for the virtual coach modelling, and Kira Williams for being the voice and motion actor. We also thank Mariagrazia Zottoli and Maria Christodoulou for advising on the statistical analysis. Daniel Freeman is an NIHR Senior Investigator. The work was also supported by the NIHR Oxford Health Biomedical Research Centre (BRC). The views expressed are those of the authors and not necessarily those of the National Health Service, NIHR, or the Department of Health.

Author contributions
S.W., D.F., and A.R. conceived the study. S.W. created the VR environments, S.W. completed recruitment and testing, and conducted the analysis. S.W. wrote the first draft of the manuscript. A.R. and D.F. supervised the research project and contributed to the writing of the manuscript.

Competing interests
Daniel Freeman is a founder of Oxford VR, a University of Oxford spin-out company, which commercialises automated VR therapies. The other authors do not have any competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-38499-7.

Correspondence and requests for materials should be addressed to S.W.

scientific reports

Check for updates

OPEN

A randomised controlled test in virtual reality of the effects on paranoid thoughts of virtual humans' facial animation and expression

Shu Wei[1,2]✉, Daniel Freeman[2,3], Victoria Harris[4] & Aitor Rovira[2,3]

Virtual reality (VR) is increasingly used in the study and treatment of paranoia. This is based on the finding that people who mistakenly perceive hostile intent from other people also perceive similar threat from virtual characters. However, there has been no study of the programming characteristics of virtual characters that may influence their interpretation. We set out to investigate how the animation and expressions of virtual humans may affect paranoia. In a two-by-two factor, between-groups, randomized design, 122 individuals with elevated paranoia rated their perceptions of virtual humans, set in an eye-tracking enabled VR lift scenario, that varied in facial animation (static or animated) and expression (neutral or positive). Both facial animation (group difference = 102.328 [51.783, 152.872], $p < 0.001$, $\eta_p^2 = 0.125$) and positive expressions (group difference = 53.016 [0.054, 105.979], $p = 0.049$, $\eta_p^2 = 0.033$) led to less triggering of paranoid thoughts about the virtual humans. Facial animation (group difference = 2.442 [−4.161, −0.724], $p = 0.006$, $\eta_p^2 = 0.063$) but not positive expressions (group difference = 0.344 [−1.429, 2.110], $p = 0.681$, $\eta_p^2 = 0.001$) significantly increased the likelihood of neutral thoughts about the characters. Our study shows that the detailed programming of virtual humans can impact the occurrence of paranoid thoughts in VR. The programming of virtual humans needs careful consideration depending on the purpose of their use.

Paranoia—perceiving hostile intent where there is none—is prevalent in the general population. Many individuals occasionally experience paranoid thoughts, while a smaller number of people frequently experience paranoid thoughts[1]. A recent survey of a representative group of ten thousand UK adults indicated that approximately one in six people wanted help to be more trusting of other people[2]. Virtual reality (VR) has been used to both study[3–7] and treat[8–11] paranoia. Freeman et al.[6] pioneered the use of VR to assess and understand paranoia by examining people's appraisals of neutral virtual humans. The insight was that if the characters are neutral but hostile intent is perceived, then this is clear evidence of paranoid thinking. Studies have shown that higher levels of paranoia in daily life are associated with experiencing higher levels of paranoia about virtual characters[12]. Qualitative findings indicate that close observations of the virtual humans may contribute to the occurrence of paranoid interpretations. As a participant in one study described: "I was just looking around, looking at people, just observing them…"[13]. This paper reports, for the first time, on the detailed characteristics of virtual humans that may affect their appraisal by people vulnerable to paranoia.

Previous VR studies outside of the topic of paranoia have shown that facial expressions and animations of virtual humans can significantly influence people's behavioural and psychological responses. For example, Geraets et al.[14] suggested that facial emotion cues are beneficial for non-clinical populations to accurately identify the emotions of virtual characters, with the recognition accuracy in VR comparable to that in photographs and videos. Additionally, it was observed that participants tended to focus more on the eye and nose areas when interpreting the emotions of the virtual humans. Bönsch et al.[15] looked at how the emotional expressions of virtual humans affect personal space preference, by examining responses to approaches by virtual men exhibiting happy,

[1]Department of Psychiatry, University of Oxford, Oxford, UK. [2]Department of Experimental Psychology, University of Oxford, Radcliffe Observatory Quarter, Oxford OX2 6GG, UK. [3]Oxford Health NHS Foundation Trust, Oxford, UK. [4]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK. ✉email: shu.wei@psych.ox.ac.uk

angry, or neutral expressions on static faces. The study showed that participants maintained larger distances from a virtual man with an angry face compared to a happy or neutral face. Moreover, Kimmel et al.[16] found that integrating facial animations, such as mouth and eye movements, not only enhanced the social presence in the scenario felt by participants but also made participants feel that the virtual humans had a better understanding of their emotions and attitudes.

Notably, recent investigations have shown the significant role of virtual human faces in shaping the perception of trust in VR[17–19]. Luo et al.[18] suggested that, compared to neutral or negative facial expressions, positive facial expressions promoted trust and willingness to cooperate in a VR game. Choudhary et al.[17] explored the impact of conflicting facial and vocal emotional expressions. They found that virtual humans with happy faces were perceived as happier and more trustworthy than virtual humans with unhappy faces. Although appraisals of trust became less predictable for mismatched expressions (e.g., a happy face with unhappy voice), facial expressions had a stronger impact than vocal tone. Wei et al.[19] found that adding positive facial expressions to a virtual coach in a VR phobia treatment significantly improved user connection with that coach. Furthermore, animated faces of VR characters have been found to be perceived as more natural and believable than static faces in VR social experiences[16,20].

Despite the evidence of the impact of virtual human characteristics on user perceptions, the influence of the detailed programming of virtual human faces on paranoia remains untested. In this study, we focused on two key elements: facial animation and facial expression. Previous research has indicated that facial animation can make virtual humans appear more empathetic and less strange[21], while positive facial expressions can promote trust[18]. Therefore, our primary hypothesis was that using facial animation or positive facial expression would reduce the likelihood of paranoia appraisals. VR eye-tracking enables capture of objective data of where a person is looking[14,22]. Hence, we also examined eye tracking data to provide information about how individuals vulnerable to paranoia may pay visual attention to virtual characters.

Methods

Experimental design

We used a two-by-two factorial, between-groups, randomised design to examine two aspects of facial programming: animation (static or animated) and expression (neutral or positive). Participants were randomised to one of the four experimental conditions (see Fig. 1a): all virtual humans having (1) static neutral (2) animated neutral (3) static positive or (4) animated positive faces. In all experimental conditions, the virtual humans had the same body animation. The study was conducted single-blind: participants were unaware of the study hypotheses and that they were randomized into one of the different versions. The randomization was conducted by an independent researcher using *Research Randomizer*[23].

Apparatus and VR scenario

We used a Windows 10 computer (Intel i7-8700K, Nvidia GeForce GTX 1080Ti, 32 GB RAM) to run the VR scenario and render it on a *Meta Quest Pro* (Meta, 2022) via Meta Air Link. The headset has a resolution of 1832 * 1920 pixels per eye and a field of view of 106° (horizontal) × 96° (vertical). The refresh rate was set at 90 Hz refresh rate.

The VR software was developed in Unity 2021.3.15 with Oculus' Movement SDK V1.3.2, presenting a virtual lift scenario. Participants began in a hallway waiting for the lift to arrive. The lift door opened automatically upon arrival, and participants were instructed to take the lift to the "sky lounge". Inside the lift were five virtual humans (three men and two women of different ages and ethnicities), as shown in Fig. 1b. The ride lasted three

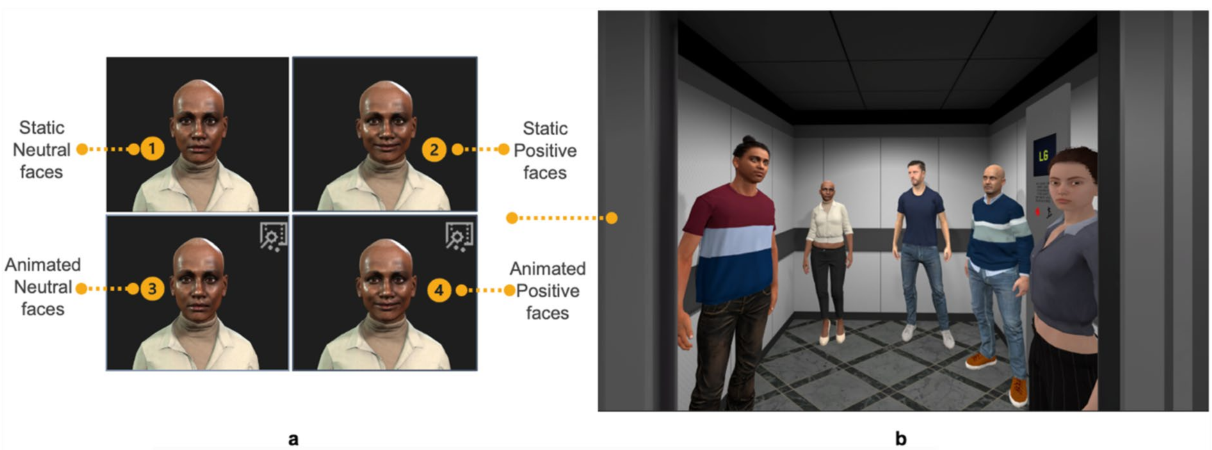


Figure 1. (**a**) Factorial design with virtual human faces varied in animation (static or animated) and expressions (neutral or positive) (**b**). VR lift environment: view of the lift when the door opened with five virtual humans standing inside.

minutes, and the VR scenario concluded when the lift reached the "sky lounge". A video of the VR experience is provided in the supplementary materials.

We implemented the facial animations by combining motion capture technologies and blend shape animation. Starting with motion-captured facial movements to establish a foundational animation, we then applied blend shapes for smoother transitions between expressions. The animations were recorded and refined using *Iclone7* with the *LiveFace* plugin (https://mocap.reallusion.com/iclone-motion-live-mocap/iphone-live-face.html). We also programmed the characters to occasionally look at the participants and make eye contact with them.

We used the integrated eye tracker in the *Meta Quest Pro* to monitor eye gaze behaviour in VR. We pre-defined a set of regions of interest (ROIs) in the VR environment to record the eye gazes directed at them. The ROIs included the faces of all five virtual humans and other specific areas where participants might look to avoid eye contact—the floor, the screen in the lift displaying the current floor, and the exit door. The software was adjusted to minimize false negatives when detecting eye fixation to any of the elements on the ROI list. We accounted for the accuracy reported of the Meta Quest Pro's eye tracker (accuracy: 1.652°; SD precision: 1.652°)[24].

Participants and recruitment

Participants were recruited via social media advertisements in Oxfordshire, United Kingdom. We screened for individuals vulnerable to paranoia using *The Revised Green* et al., *Paranoid Thoughts Scale (R-GPTS)*[25], with a Part B score greater than 5. This cut-off score captures elevated or higher levels of persecutory ideation. Exclusion criteria were individuals (a) under 18 years old, (b) reported photosensitive epilepsy in the past or a significant visual, hearing, or mobility impairment that would prevent them from using VR, or (c) currently under medication that could induce motion sicknesses. Participants requiring correction-to-normal lenses were requested to use contact lenses to avoid any potential discomfort wearing the VR headset and to preserve the quality of the data recorded with the eye tracker.

Ethical approval was received from the University of Oxford Medical Sciences Interdivisional Research Ethics Committee (R85111/RE001). The study was performed in accordance with relevant guidelines and regulations. Written informed consent was obtained from all participants.

1581 individuals completed the screening questionnaire, 296 were eligible (i.e. adults with elevated paranoia). 122 participants (female = 70, male = 52) attended the VR session (we had booked an extra two participants over the target sample size to account for potential cancellations).The average age of the participants was 36.2 years (SD = 14.8, range: 18, 76). The mean R-GPTS Part B score was 13.30 (SD = 7.70, range: 6, 37). The average previous experience of VR rated on a 5-point-scale (where 1 indicates "never tried VR" and 5 indicates "very experienced") was 2.00 (SD = 1.13). Table 1 provides a summary of the participants' demographic information and study-relevant data.

Experimental procedures

Each participant was invited to the university for a single session at our VR lab. The researcher provided an overview of the study procedure and informed participants that they would try out a VR social experience, during which eye gaze direction data would be collected (no pictures or videos of their eyes would be recorded). Participants gave written informed consent to participate. The researcher then helped the participants fit the VR headset and guided them through the eye tracker calibration process. The researcher then selected the parameters according to each participant's experimental condition group and they experienced the VR lift ride. Once the VR scenario ended, participants took the VR headset off and completed the paranoid thoughts visual analogue scales[26] and State Social Paranoia Scale[27]. Finally, they were debriefed about the full purpose of the study. The session lasted approximately 45 min, and participants were reimbursed for their time.

Measures

Baseline paranoia During screening participants completed *The Revised Green* et al., *Paranoid Thoughts Scale (R-GPTS)*[25]. We used Part B to assess ideas of persecution. There are 10 items such as "I was sure someone wanted

	Static neutral (n = 31)	**Static positive (n = 30)**	**Animated neutral (n = 30)**	**Animated positive (n = 31)**
Mean age in years (SD)	36.2 (16.3)	36.1 (13.6)	36.8 (13.9)	35.9 (15.9)
Gender				
Male (%)	12 (38.7%)	12 (40.0%)	15 (50.0%)	13 (41.9%)
Female (%)	19 (61.3%)	18 (60.0%)	15 (50.0%)	18 (58.1%)
Ethnicity				
White	25	18	22	25
Black/African American	0	1	0	0
Asian	6	7	6	6
Others	0	4	2	0
R-GPTS Part B (i.e. baseline paranoia score) (SD)	13.35 (7.71)	13.07 (7.31)	14.00 (8.18)	12.81 (7.89)
Previous experience of VR (SD)	2.00 (0.93)	2.07 (1.28)	2.40 (1.19)	2.19 (1.17)

Table 1. Participant information per group.

to hurt me" and "People have been hostile towards me on purpose". All items are scored from 0 (Not at all) to 4 (Totally), with total scores ranging from 0 to 40 (Cronbach's α = 0.86 in the current study, N = 122). Higher scores reflect higher levels of paranoia.

Paranoid thoughts visual analogue scales (VAS)[26] This was the *primary* paranoia outcome measure, assessing participants' appraisals of the virtual humans in the lift scenario. After the VR experience, participants rated 6 visual analogue scales concerning the VR humans ("Right now I feel suspicious of the people in the lift", "Right now I feel that people in the lift wanted to harm me", "Right now I feel like the people in the lift wanted to upset me", "Right now I feel like the people in the lift were against me", "Right now I am thinking that the people in the lift were trying to persecute me" and "Right now I feel like the people in the lift were hostile towards me"). Participants marked each item on a standard 10 cm visual analogue scale on paper from 0 (not at all anxious) to 100 (extremely), with total scores ranging from 0 to 600 (Cronbach's α = 0.935 in the current study, N = 122). Higher scores indicate higher levels of paranoia about the virtual humans.

State social paranoia scale (SSPS)[27] This provided a further assessment of paranoid thoughts about the virtual human and also neutral and positive appraisals. In the scale, each item is scored from 1 (Do not agree) to 5 (Totally agree). There are 10 items measuring paranoid thoughts (SSPSPersecutory) (range: 10, 50, Cronbach's α = 0.950 in the current study, N = 122), and 5 items each measuring neutral views (SSPSNeutral) (range: 5, 25, Cronbach's α = 0.825) and positive views (SSPSPositive) (range: 5, 25, Cronbach's α = 0.736) of the people in the VR social situation. Higher SSPS scores on each subscale indicate greater levels of persecutory or neutral or positive thinking.

Eye gaze data We recorded the raw eye tracker output and the detected eye gazes on ROIs (i.e. the start and end times for each gaze on an ROI object). Data were first screened for missing records by examining gaps in the raw data output file. We excluded the data from a participant if more than 15% of the data were missing, following the suggested practices in Holmqvist et al. and Schuetz & Fiehler[28,29]. We processed the data to calculate the duration for each eye gaze event, and identified fixations using a time threshold of 0.275 seconds[30]. We then aggregated these fixations for each ROI object per participant. The following variables were calculated for analysis:

- *Visual Attention to Virtual Humans*: total fixation time on virtual humans as the percentage of total fixation time on all ROI objects.
- *Visual Attention to Exit/Floor/Lift Screen (displays the current floor)*: total fixation time on lift exit door/floor/screen as the percentage of total fixation time on all ROI objects.
- *First Fixation Target:* The object participants looked at on the first fixation after entering the lift.

Statistical methods

We used two-way ANCOVA models, examining the effects on participant views of the virtual humans of facial animation and positive facial expression while controlling for baseline paranoia. We first checked the data against the assumptions of the ANCOVA model, using Levene's test for homogeneity of variance and Shapiro–Wilk test of normality. We also performed a log transformation on heavily skewed data (visual attention to the floor and screen) before analysis. Details of the assumption checking results are included in the supplementary materials. A similar approach was taken to analysing the eye gaze data.

All significance tests were made at the α = 0.05 level, and we calculated the partial eta-squared (η_p^2) to measure the effect sizes. Tukey's honest significant difference test (Tukey's HSD) was used for multiple pairwise comparisons with the adjusted *p* value. We report the results as mean group differences and 95% confidence interval (95% CI). Additionally, we conducted contrast tests in cases where a significant interaction was detected, assessing the impact of each factor at different levels of another factor, with estimates reported alongside their 95% confidence interval. Data cleaning and processing was performed using *Python*'s *Pandas* and *NumPy* libraries[31,32]. The statistical analysis was done in *R*.

To determine the target sample size for our experimental design, we aimed to detect a medium effect size of partial eta-squared = 0.06 and conventional values of power = 0.80 and α = 0.05 for a between-factors ANOVA using G * power 3.124. Thus, a total of 120 participants (30 per condition) would be required.

Results

Table 2 summarises scores for the paranoid thoughts VAS and the three SSPS subscales by randomised group. There were no missing data in these measures. Both the paranoid thoughts VAS and SSPS persecutory were used to assess participants' paranoid ideation and were positively correlated (Spearman r = 0.82, $p < 0.001$). Figure 2 shows summary scores for these two measures in the randomised groups.

	Static neutral (n = 31)	Static positive (n = 30)	Animated neutral (n = 30)	Animated positive (n = 31)
VAS Paranoia mean (SD)	285.00 (126.15)	195.33 (146.47)	145.20 (141.88)	132.16 (138.66)
SSPSPersecutory mean (SD)	28.71 (10.86)	20.20 (10.60)	23.03 (10.89)	19.52 (9.80)
SSPSNeutral mean (SD)	11.03 (4.56)	11.87 (4.26)	14.70 (5.08)	13.10 (5.19)
SSPSPositive mean (SD)	11.03 (3.69)	10.87 (4.36)	11.27 (4.17)	14.23 (4.28)

Table 2. Descriptive statistics of the appraisals of the virtual humans by randomisation group.

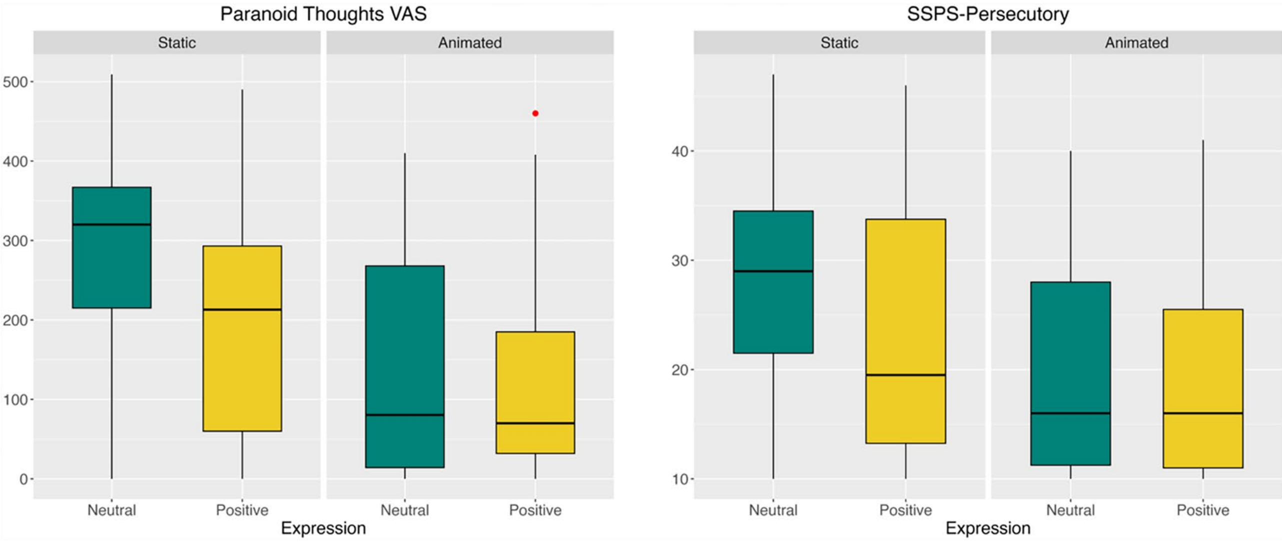


Figure 2. Box plots of the VAS *Paranoia* and *SSPS-Persecutory.* Red point indicates the outlier point.

Paranoid thoughts visual analogue scale (VAS paranoia)

A two-way ANCOVA model was used to assess the impact of facial animation and expressions while controlling for baseline paranoia. Simple main effects analysis showed that facial animation (group difference = 102.328, 95% CI = [51.783, 152.872], F(1, 117) = 17.071, $p < 0.001$, $\eta_p^2 = 0.125$) and positive expression (group difference = 53.016, 95% CI = [0.054, 105.979], F(1, 117) = 3.938, $p = 0.049$, $\eta_p^2 = 0.033$) led to less paranoid thinking about the virtual humans. There was no significant interaction between animation and positive expression (F(1, 117) = 2.519, $p = 0.115$, $\eta_p^2 = 0.021$). The effect of baseline paranoia was not significant (F(1, 117) = 2.678, $p = 0.104$, $\eta_p^2 = 0.022$). Tukey's HSD test for multiple comparisons showed there was a significant difference between the static neutral and animated neutral group (p-adj < 0.001) and between the static neutral and animated positive group (p-adj < 0.001).

SSPS-persecutory thoughts

Simple main effects analysis showed that facial animation (group difference = 6.066, 95% CI = [2.247, 9.885], F(1, 117) = 10.464, $p = 0.002$, $\eta_p^2 = 0.081$) but not positive expressions (group difference = 3.279, 95% CI = [− 0.650, 7.207], F(1, 117) = 2.534, $p = 0.114$, $\eta_p^2 = 0.021$) led to significantly lower levels of paranoia. There was no significant interaction between animation and positive expressions (F(1, 117) = 1.897, $p = 0.171$, $\eta_p^2 = 0.016$). The effect of baseline paranoia was not significant (F(1, 117) = 3.624, $p = 0.059$, $\eta_p^2 = 0.030$). Tukey's HSD test for multiple comparisons showed a statistically significant difference between the static neutral group and animated neutral group (p-adj = 0.008), and between the static neutral and animated positive group (p-adj = 0.005).

SSPS-neutral thoughts

Simple main effects analysis showed that facial animation (group difference = 2.442, 95% CI = [− 4.161, − 0.724], F(1, 117) = 7.843, $p = 0.006$, $\eta_p^2 = 0.063$) but not positive expressions (group difference = 0.344, 95% CI = [− 1.429, 2.110], F(1, 117) = 0.17, $p = 0.681$, $\eta_p^2 = 0.001$) led to a more neutral interpretation of the virtual humans. There was no significant interaction between animation and positive expressions (F(1, 117) = 1.914, $p = 0.169$, $\eta_p^2 = 0.016$). The effect of baseline paranoia was not significant (F(1, 117) = 0.442, $p = 0.508$, $\eta_p^2 = 0.002$). Tukey's HSD test for multiple comparisons showed there was a statistically significant difference between the static neutral group and animated neutral group (p-adj = 0.019).

SSPS-positive thoughts

There was a significant interaction between animation and positive expressions (F(1, 117) = 4.297, $p = 0.040$, $\eta_p^2 = 0.035$). Facial animation led to more positive thoughts when the expressions were positive (estimate = 3.358 [1.250, 5.460], SE = 1.060, $p = 0.002$), but not when expressions were neutral (estimate = 0.236 [− 1.870, 2.340], SE = 1.060, $p = 0.825$). Positive expressions led to more positive thoughts when the faces were animated (estimate = 2.956 [0.849, 5.060], SE = 1.060, $p = 0.006$), but not when faces were static (estimate = − 0.166 [− 2.271, 1.940], SE = 1.060, $p = 0.876$). Tukey's HSD test for multiple comparisons showed that there was a statistically significant difference between the static neutral and animated positive group (p-adj = 0.016), the static positive and animated positive group (p-adj = 0.011), and between the animated neutral and animated positive group (p-adj = 0.032).

Paranoid thinking in VR and baseline paranoia

The correlations between paranoid thinking in VR and baseline paranoia were examined in the different groups through spearman correlation (see Table 3). A moderate positive correlation was found between paranoid thinking in VR and baseline paranoia within the animated neutral group.

	Spearman correlation	*p* value
Static neutral (n = 31)	−0.110	0.554
Static positive (n = 30)	0.000	0.999
Animated neutral (n = 30)	0.385	0.036
Animated positive (n = 31)	0.194	0.296

Table 3. Correlation between VAS paranoia and baseline paranoia.

Visual attention

Five datasets from the eye-tracking analysis were excluded as more than 15% of the raw data were missing due to technical issues. Analysis of the remaining 117 participants during the lift ride showed an average total fixation duration of 112.59 s (SD = 28.42) and an average of 1.86 s (SD = 1.84) per fixation. Participants spent 31.95% of the fixation time on the virtual humans (SD = 24.68%). The average duration for these fixations was 5.87 s (SD = 5.84). The most common initial fixation targets were the male virtual human directly facing the lift entrance (27.4%), the exit (17.9%), and the floor (17.9%) (see Fig. 3). Descriptive statistics for the visual attention allocation are shown in Table 4.

Visual attention to the virtual humans was tested using a two-way ANCOVA model controlling for baseline paranoia. There was a marginally non-significant interaction between animation and positive expressions (F(1, 112) = 3.591, $p = 0.061$, $\eta_p^2 = 0.031$), suggesting a trend where animation and positive expression might jointly influence the amount of visual attention allocated to the virtual humans. Specifically, positive expressions led to a lower amount of visual attention on virtual humans when faces were static (estimate = −0.169 [−0.294, −0.043], SE = 0.064, $p = 0.010$), but not when faces were animated (estimate = 0.001 [−0.124, 0.127], SE = 0.063, $p = 0.983$). Facial animation did not affect visual attention to virtual humans either when the facial expressions were neutral (estimate = −0.060 [−0.184, 0.064], SE = 0.062, $p = 0.338$) or positive (estimate = 0.110 [−0.017, 0.238], SE = 0.064, $p = 0.090$). Tukey's HSD test indicated a significant difference between the static neutral and static positive group (*p*-adj = 0.044).

Visual attention to the environment The same two-way ANCOVA model was used to examine the extent of visual attention on the lift exit door, floor, and screen separately. For the exit, there was a significant interaction between animation and positive expressions (F(1, 112) = 8.026, $p = 0.005$, $\eta_p^2 = 0.067$). Facial animation led to a higher amount of visual attention on the exit when the expressions were neutral (estimate = 0.176 [0.027, 0.325], SE = 0.075, $p = 0.021$), but not when expressions were positive (estimate = −0.131 [−0.285, 0.023], SE = 0.078, $p = 0.095$). Positive expressions led to a higher amount of visual attention on the exit when faces were static (estimate = 0.182 [0.030, 0.334], SE = 0.077, $p = 0.020$), but not when faces were animated (estimate = −0.125

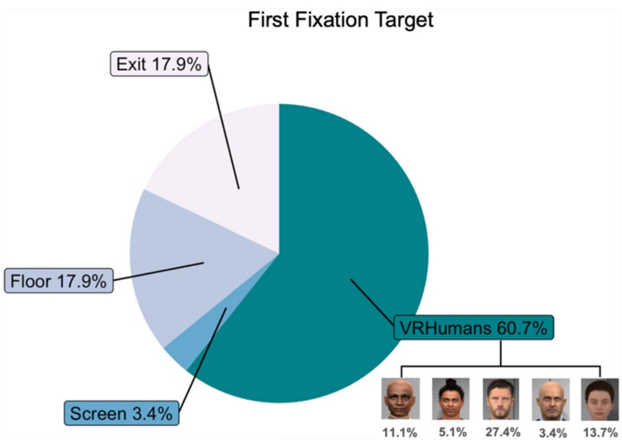


Figure 3. Distribution of the target of first fixation.

Visual attention allocation (%)	Static neutral (n = 30)	Static positive (n = 28)	Animated neutral (n = 30)	Animated positive (n = 29)
VR humans (SD)	38.8 (20.6)	21.9 (19.7)	32.6 (27.3)	33.8 (27.9)
Exit (SD)	23.8 (21.2)	41.8 (30.8)	41.9 (36.6)	26.7 (29.3)
Floor (SD)	9.7 (18.3)	8.1 (15.5)	7.4 (14.5)	9.0 (11.6)
Screen (SD)	25.7 (19.7)	26.5 (23.4)	16.8 (16.1)	29.1 (26.0)

Table 4. Visual attention to virtual humans, lift exit, lift floor, and lift screen.

[− 0.277, 0.027], SE = 0.077, $p = 0.105$). Tukey's HSD test indicated no statistically significant differences in pairwise comparisons.

The analysis was performed on log-transformed data for visual attention to the floor and the lift screen. For the floor, there was no significant interaction between animation and positive expressions ($F(1, 112) = 0.306$, $p = 0.581$, $\eta_p^2 = 0.003$) and there were no main effects from animation ($F(1, 112) = 0.025$, $p = 0.874$, $\eta_p^2 < 0.001$) or positive expression ($F(1, 112) = 0.011$, $p = 0.916$, $\eta_p^2 < 0.001$). The effect of baseline mistrust was not significant ($F(1, 112) = 2.493$, $p = 0.117$, $\eta_p^2 = 0.020$). For the lift screen, there was no significant interaction between animation and positive expressions ($F(1, 112) = 1.895$, $p = 0.171$, $\eta_p^2 = 0.017$) and there were no main effects from animation ($F(1, 112) = 0.846$, $p = 0.360$, $\eta_p^2 = 0.009$) or positive expression ($F(1, 112) = 2.496$, $p = 0.117$, $\eta_p^2 = 0.019$). The effect of baseline mistrust was not significant ($F(1, 112) = 1.250$, $p = 0.266$, $\eta_p^2 = 0.009$).

Correlation between visual attention and paranoia We computed Spearman correlations between visual attention and paranoia using all the retained eye-tracking data across different condition groups (N = 117, see Table 5). There was a positive correlation between the amount of visual attention to the virtual humans and the severity of paranoid thoughts in VR (VAS: r = 0.19, $p = 0.040$) and a positive correlation between the amount of visual attention to the lift exit and baseline paranoia (Baseline Paranoia: r = 0.21, $p = 0.023$). There were no significant correlations between the paranoid thoughts in VR/ baseline paranoia and the visual attention to other ROIs.

Discussion

Our study presented the first test of whether the detailed programming of virtual humans' facial features affects paranoid interpretations. The primary results supported the hypothesis that facial animation and positive facial expressions of virtual humans both reduce the likelihood of paranoid appraisals. In our study, facial animation and positive expression each independently led to people vulnerable to paranoia perceiving the virtual characters as less hostile. In contrast, paranoid thoughts were more likely to occur when faces were static or the expression was neutral. The sizes of the effects were moderate to large. Facial animation fostered more neutral perceptions of virtual humans too. An examination of the correlation between paranoid thinking in VR and baseline paranoia in each condition indicated that the animated neutral characters produced the strongest association between paranoia in day-to-day life and paranoia in VR. This means that animated neutral characters in VR may provide the most effective assessment test for paranoia. Overall, the study highlights the importance of considering how faces in VR are programmed when assessing or treating paranoia.

The findings align with prior research showing that facial-animated characters appear more natural and believable[20], and people are better at recognising emotions from them[33]. Similarly, the addition of positive emotion led to the perception of less negative intention or attributes from the virtual characters regardless of whether their faces were animated[14,19]. It was notable that the effects of these two features were independent of people's baseline paranoia, and in this study, animation had a more substantial effect (accounting for 13% of the variance in paranoia) compared to facial expressions (accounting for 3% of the variance). This might be attributed to our implementation of positive expressions as friendly faces with gentle, subtle smiles, to fit the neutral VR context. Such nuanced emotional expression typically requires accurate delivery with dynamic movement[34]; the absence of animations may lead to the "frozen face" effect, where a static human face appears less flattering than one with motion[35]. Additionally, the lack of dynamic information in the virtual human faces could render their expressions more ambiguous, leading individuals with elevated paranoia to interpret this ambiguity negatively and perceive the virtual humans as potentially hostile[36,37].

The programming of virtual human faces also influenced individuals' neutral and positive perceptions. According to Krumhuber et al.[38], dynamic information (e.g. animation) enhances emotion recognition, particularly when facial expressions are subtle or convey a neutral emotion. Consistent with this, animation led to a more neutral interpretation of the characters, and the animated neutral faces were rated as the most neutral. Interestingly, regarding positive thoughts about the characters, an interaction effect suggested that animation was critical for positive expressions to lead to stronger positive interpretations, while the static positive faces scored the lowest. This reduced likelihood of eliciting positive thoughts from static positive faces likely stems from the mismatch between expressed emotions and the absence of movement, making the characters appear less lively and emotionally inconsistent.

Examining visual attention to virtual humans, we found that positive expressions led to less visual attention when virtual human faces were static. The discrepancy between positive facial expressions and the absence of animation could lead to the characters being perceived as anomalous. Particularly, lack of eye movements, such as eye blinking or gaze, could cause smiles to seem eerie or ungenuine[39,40]. Exploring the link between visual attention and paranoia, we found a positive relationship between attention to virtual humans and paranoid

	Correlation with VAS paranoia		Correlation with baseline paranoia	
Visual attention allocation (%)	Spearman r	*p* value	Spearman r	*p* value
Virtual humans	0.190	0.040	− 0.046	0.621
Exit	− 0.152	0.101	0.210	0.023
Floor	0.117	0.208	− 0.101	0.277
Screen	− 0.040	0.668	− 0.079	0.398

Table 5. Correlation between visual attention and paranoia measures (N = 117).

thoughts in VR. This might imply that closely observing virtual humans could foster the development of paranoid ideations[13], or that individuals with heightened paranoia are more inclined to concentrate on these characters. In addition, there was also a positive relationship between people's focus on the lift exit and their levels of day-to-day paranoia. This behaviour aligns with the use of safety-seeking strategies in response to persecutory thoughts[41], verifying that VR elicits reactions similar to those in the real world. Such behaviour also coincides with the patterns found in social anxiety studies, where individuals often avoid eye contact with virtual humans and shift their attention to other areas of the virtual scene under distress[42,43]. Although the complex relationships between visual attention, character animation, and people's mental health states requires further exploration, the study of eye gaze behaviour might provide additional information to help understand paranoia.

There are several limitations to the study. First, our choice of the VR lift scenario limited participants to a close distance (less than 2 m) from the virtual humans, which may affect the generalizability of the results to scenarios involving greater social distances. A comparison of being in a lift to walking into a room would be of clear interest. Interpersonal distance in VR can affect people's emotional and behavioural responses[44]. Moreover, the neutral context of strangers in a lift ride may not produce the same results in scenarios with other social interactions. Second, we focused on two features—animation and positive expressions—but other characteristics such as eye gaze behaviour patterns and facial mimicry could also be important[45,46]. Third, we did not consider demographic (e.g. age, gender, ethnicity) similarities or differences between participants and the characters, nor their spatial arrangement. Fourth, our visual attention analysis focused only on spatial allocation and fixation-related metrics, excluding other relevant measures like saccades and gaze angles. Current technology provides limited capability to study whether participants were looking at someone from the corner of their eyes. Additionally, the resolution of the eye tracker was insufficient for a detailed analysis of which facial parts participants focused on (e.g. eyes or mouth). As prior research has shown that people direct their attention to different parts of virtual faces depending on the displayed emotions[14], further investigation could provide more comprehensive insights into participant behaviours.

Our findings provide evidence that character animations alter people's perceptions and experiences in VR. This may, for example, affect VR experiences focused on the understanding and treatment of paranoia. Therefore, careful consideration of character design and animation is likely to be important in developing future VR mental health applications.

Data availability

Deidentified data are available from the corresponding authors on reasonable request and contract with the university.

Received: 25 March 2024; Accepted: 12 July 2024
Published online: 24 July 2024

References

1. Freeman, D. *et al.* Psychological investigation of the structure of paranoia in a non-clinical population. *Br. J. Psychiatry* **186**, 427–435 (2005).
2. Freeman, D. & Loe, B. S. Explaining paranoia: Cognitive and social processes in the occurrence of extreme mistrust. *BMJ Ment. Health* **26**, e300880 (2023).
3. Freeman, D. *et al.* Virtual reality study of paranoid thinking in the general population. *Br. J. Psychiatry* **192**, 258–263 (2008).
4. Gorisse, G. *et al.* Self-observation of a virtual body-double engaged in social interaction reduces persecutory thoughts. *Sci. Rep.* **11**, 23923 (2021).
5. Valmaggia, L. R. *et al.* Virtual reality and paranoid ideations in people with an 'at-risk mental state' for psychosis. *Br. J. Psychiatry* **191**, s63–s68 (2007).
6. Freeman, D. *et al.* Can virtual reality be used to investigate persecutory ideation?. *J. Nerv. Ment. Dis.* **191**, 509–514 (2003).
7. Freeman, D. *et al.* The psychology of persecutory ideation II: A virtual reality experimental study. *J. Nerv. Ment. Dis.* **193**, 309–315 (2005).
8. Freeman, D. *et al.* Virtual reality in the treatment of persecutory delusions: Randomised controlled experimental study testing how to reduce delusional conviction. *Br. J. Psychiatry* **209**, 62–67 (2016).
9. Freeman, D. *et al.* Automated virtual reality cognitive therapy versus virtual reality mental relaxation therapy for the treatment of persistent persecutory delusions in patients with psychosis (THRIVE): A parallel-group, single-blind, randomised controlled trial in England with mediation analyses. *Lancet Psychiatry* **10**, 836–847 (2023).
10. Pot-Kolder, R. M. C. A. *et al.* Virtual-reality-based cognitive behavioural therapy versus waiting list control for paranoid ideation and social avoidance in patients with psychotic disorders: A single-blind randomised controlled trial. *Lancet Psychiatry* **5**, 217–226 (2018).
11. Freeman, D. *et al.* Agoraphobic avoidance in patients with psychosis: Severity and response to automated VR therapy in a secondary analysis of a randomised controlled clinical trial. *Schizophr. Res.* **250**, 50–59 (2022).
12. Freeman, D., Pugh, K., Vorontsova, N., Antley, A. & Slater, M. Testing the continuum of delusional beliefs: An experimental study using virtual reality. *J. Abnorm. Psychol.* **119**, 83–92 (2010).
13. Fornells-Ambrojo, M. *et al.* how do people with persecutory delusions evaluate threat in a controlled social environment? A qualitative study using virtual reality. *Behav. Cogn. Psychother.* **43**, 89–107 (2015).
14. Geraets, C. N. W. *et al.* Virtual reality facial emotion recognition in social environments: An eye-tracking study. *Internet Interv.* **25**, 100432 (2021).
15. Bönsch, A., Radke, S., Ehret, J., Habel, U. & Kuhlen, T. W. The impact of a virtual agent's non-verbal emotional expression on a user's personal space preferences. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents* 1–8 (Association for Computing Machinery, New York, NY, USA, 2020). https://doi.org/10.1145/3383652.3423888.
16. Kimmel, S., Jung, F., Matviienko, A., Heuten, W. & Boll, S. Let's face it: Influence of facial expressions on social presence in collaborative virtual reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* 1–16 (Association for Computing Machinery, New York, NY, USA, 2023). https://doi.org/10.1145/3544548.3580707.
17. Choudhary, Z. *et al.* Exploring the social influence of virtual humans unintentionally conveying conflicting emotions. In *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)* 571–580 (2023). https://doi.org/10.1109/VR55154.2023.00072.

18. Luo, L., Weng, D., Ding, N., Hao, J. & Tu, Z. The effect of avatar facial expressions on trust building in social virtual reality. *Vis. Comput.* https://doi.org/10.1007/s00371-022-02700-1 (2022).
19. Wei, S., Freeman, D. & Rovira, A. A randomised controlled test of emotional attributes of a virtual coach within a virtual reality (VR) mental health treatment. *Sci. Rep.* **13**, 11517 (2023).
20. Kullmann, P., Menzel, T., Botsch, M. & Latoschik, M. E. An evaluation of other-avatar facial animation methods for social VR. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* 1–7 (ACM, Hamburg Germany, 2023). https://doi.org/10.1145/3544549.3585617.
21. Parmar, D., Olafsson, S., Utami, D., Murali, P. & Bickmore, T. Designing empathic virtual agents: Manipulating animation, voice, rendering, and empathy to create persuasive agents. *Auton. Agents Multi-Agent Syst.* **36**, 17 (2022).
22. Shadiev, R. & Li, D. A review study on eye-tracking technology usage in immersive virtual reality learning environments. *Comput. Educ.* **196**, 104681 (2023).
23. Urbaniak, G. & Plous, S. Research randomizer (version 4.0) [computer software] (2013).
24. Wei, S., Bloemers, D. & Rovira, A. A preliminary study of the eye tracker in the meta quest pro. In *Proceedings of the 2023 ACM International Conference on Interactive Media Experiences* 216–221 (ACM, Nantes France, 2023). https://doi.org/10.1145/3573381.3596467.
25. Freeman, D. *et al.* The revised Green *et al.*, paranoid thoughts scale (R-GPTS): Psychometric properties, severity ranges, and clinical cut-offs. *Psychol. Med.* **51**, 244–253 (2021).
26. Freeman, D. *et al.* How cannabis causes paranoia: Using the intravenous administration of Δ^9—Tetrahydrocannabinol (THC) to identify key cognitive mechanisms leading to paranoia. *Schizophr. Bull.* **41**, 391–399 (2015).
27. Freeman, D. *et al.* A measure of state persecutory ideation for experimental studies. *J. Nerv. Ment. Dis.* **195**, 781–784 (2007).
28. Holmqvist, K., Nyström, M. & Mulvey, F. Eye tracker data quality: What it is and how to measure it. In *Proceedings of the Symposium on Eye Tracking Research and Applications* 45–52 (ACM, Santa Barbara California, 2012). https://doi.org/10.1145/2168556.2168563.
29. Schuetz, I. & Fiehler, K. Eye tracking in virtual reality: Vive pro eye spatial accuracy, precision, and calibration reliability. *J. Eye Mov. Res.* https://doi.org/10.16910/jemr.15.3.3 (2022).
30. Llanes-Jurado, J., Marín-Morales, J., Guixeres, J. & Alcañiz, M. Development and calibration of an eye-tracking fixation identification algorithm for immersive virtual reality. *Sensors* **20**, 4956 (2020).
31. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585**, 357–362 (2020).
32. McKinney, W. *Data Structures for Statistical Computing in Python* 56–61 (Austin, Texas, 2010). https://doi.org/10.25080/Majora-92bf1922-00a.
33. Faita, C. *et al.* Perception of basic emotions from facial expressions of dynamic virtual avatars. In *Augmented and Virtual Reality* (eds De Paolis, L. T. & Mongelli, A.) 409–419 (Springer International Publishing, 2015). https://doi.org/10.1007/978-3-319-22888-4_30.
34. Recio, G., Sommer, W. & Schacht, A. Electrophysiological correlates of perceiving and evaluating static and dynamic facial emotional expressions. *Brain Res.* **1376**, 66–75 (2011).
35. Post, R., Haberman, J., Iwaki, L. & Whitney, D. The frozen face effect: Why static photographs may not do you justice. *Front. Psychol.* **3**, 22 (2012).
36. Savulich, G., Freeman, D., Shergill, S. & Yiend, J. Interpretation biases in paranoia. *Behav. Ther.* **46**, 110–124 (2015).
37. Green, C. E. L. *et al.* Paranoid explanations of experience: A novel experimental study. *Behav. Cogn. Psychother.* **39**, 21–34 (2011).
38. Krumhuber, E. G., Skora, L. I., Hill, H. C. H. & Lander, K. The role of facial movements in emotion recognition. *Nat. Rev. Psychol.* **2**, 283–296 (2023).
39. Tinwell, A., Grimshaw, M., Nabi, D. A. & Williams, A. Facial expression of emotion and perception of the Uncanny Valley in virtual characters. *Comput. Hum. Behav.* **27**, 741–749 (2011).
40. Liew, T. W., Mat Zin, N. A., Ashaari, N. & Tan, S.-M. The effects of a pedagogical agent's smiling expression on the learner's emotions and motivation in a virtual learning environment. *Int. Rev. Res. Open Distrib. Learn.* **17**, 248–266 (2016).
41. Freeman, D. *et al.* Acting on persecutory delusions: The importance of safety seeking. *Behav. Res. Ther.* **45**, 89–99 (2007).
42. Reichenberger, J., Pfaller, M. & Mühlberger, A. Gaze behavior in social fear conditioning: An eye-tracking study in virtual reality. *front. Psychol.* **11**, 35 (2020).
43. Rubin, M. *et al.* Avoidance of social threat: Evidence from eye movements during a public speaking challenge using 360°-video. *Behav. Res. Ther.* **134**, 103706 (2020).
44. Kroczek, L. O. H., Pfaller, M., Lange, B., Müller, M. & Mühlberger, A. Interpersonal distance during real-time social interaction: Insights from subjective experience, behavior, and physiology. *Front. Psychiatry* **11**, 561 (2020).
45. Hale, J. & Hamilton, A. F. D. C. Testing the relationship between mimicry, trust and rapport in virtual reality conversations. *Sci. Rep.* **6**, 35295 (2016).
46. Raimbaud, P. *et al.* The stare-in-the-crowd effect in virtual reality. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* 281–290 (2022). https://doi.org/10.1109/VR51125.2022.00047.
47. Wei, S., Freeman, D. & Rovira, A. Visual attention and virtual human facial animations in virtual reality (VR): An eye-tracking study. In *2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* 891–892 (2024). https://doi.org/10.1109/VRW62533.2024.00242.

Acknowledgements
We thank André Lages for providing the virtual human models and Felicity Hudson for helping with the randomisation. Part of the eye-tracking analysis has been submitted to the IEEE VR 2024 conference as a poster presentation[47].

Author contributions
S.W., D.F., and A.R. conceived the study. S.W. programmed the VR software, S.W. completed recruitment and testing, and conducted the analysis. V.H. reviewed the statistical analysis. S.W. wrote the first draft of the manuscript. A.R. and D.F. supervised the research project and contributed to the writing of the manuscript. All authors reviewed the manuscript.

Funding
This research has been funded through the Oxford Cognitive Approaches to Psychosis (OCAP) research fund and the National Institute for Health and Care Research (NIHR) Oxford Health Biomedical Research Centre (BRC) (NIHR203316). DF is a NIHR Senior Investigator and is supported by the NIHR Oxford Health BRC. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. SW is partially supported by a Rabin Erza Scholarship from the Rabin Ezra Scholarship Trust.

Competing interests

Daniel Freeman is the scientific founder of Oxford VR, a University of Oxford spin-out company, which commercialises automated VR therapies. The other authors do not have any competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-67534-4.

Correspondence and requests for materials should be addressed to S.W.

Reprints and permissions information is available at www.nature.com/reprints.

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

A Preliminary Study of the Eye Tracker in the Meta Quest Pro

Shu Wei
Department of Psychiatry, University of Oxford
Oxford, UK
shu.wei@psych.ox.ac.uk

Desmond Bloemers
Independent Researcher
Oxford, UK
desmond.bloemers@googlemail.com

Aitor Rovira
Oxford Health NHS Foundation Trust
Oxford, UK
Department of Experimental Psychology, University of Oxford
Oxford, UK
aitor.rovira@psy.ox.ac.uk

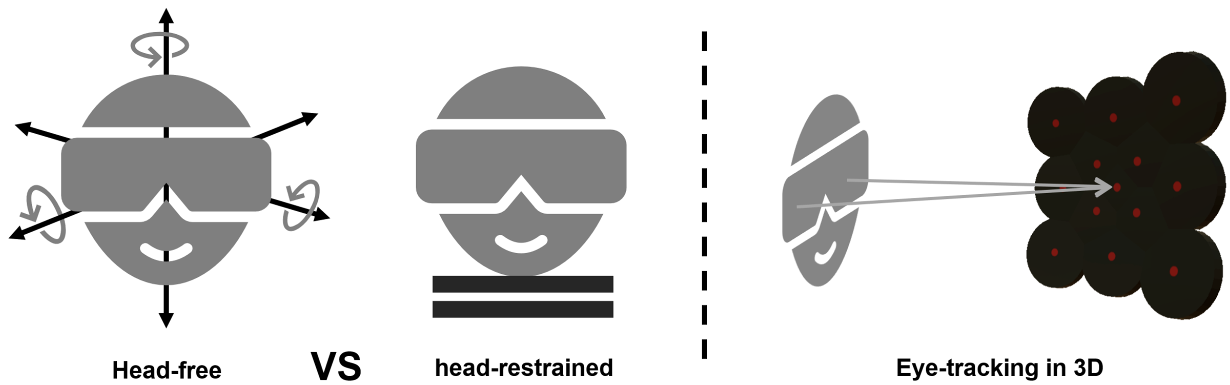


Figure 1: Study design. An eye-tracking task in head-free and head-restrained conditions.

ABSTRACT

This paper presents the preliminary results of an accuracy testing of the Meta Quest Pro's eye tracker. We conducted user testing to evaluate the spatial accuracy, spatial precision and subjective performance under head-free and head-restrained conditions. Our measurements indicated an average accuracy of 1.652° with a precision of 0.699° (standard deviation) and 0.849° (root mean square) for a visual field spanning 15° during head-free. The signal quality of Quest Pro's eye-tracker is comparable to existing AR/VR eye-tracking headsets. Notably, careful considerations are required when designing the size of scene objects, mapping areas of interest, and determining the interaction flow. Researchers should also be cautious about interpreting the fixation results when multiple targets are within close proximity. Further investigation and better specification information transparency are needed to establish its capabilities and limitations.

CCS CONCEPTS

• **Human-centered computing** → **Interaction devices**; **Virtual reality**.

IMX '23, June 12–15, 2023, Nantes, France

ACM ISBN 979-8-4007-0028-6/23/06.
https://doi.org/10.1145/3573381.3596467

KEYWORDS

Eye-tracking; Virtual Reality (VR); Accuracy; Precision; Meta Quest Pro

ACM Reference Format:
Shu Wei, Desmond Bloemers, and Aitor Rovira. 2023. A Preliminary Study of the Eye Tracker in the Meta Quest Pro. In *ACM International Conference on Interactive Media Experiences (IMX '23), June 12–15, 2023, Nantes, France.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3573381.3596467

1 INTRODUCTION

The integration of eye-tracking technology in virtual reality (VR) opens up the possibility of evaluating users' eye gaze behaviours in real time, providing valuable insights into how people respond in various social situations. An eye tracker enables researchers to observe which objects in the VR scene people focus on and for how long. It allows for empirical assessments of how people interact with virtual human characters and other elements within the VR environment. High-quality eye-tracking data can also help researchers to identify subtle changes in a user's cognitive and emotional states [16], enabling more nuanced interpretations of people's behavioural responses.

A growing body of psychology research has incorporated eye-tracking in VR social scenarios for behavioural insights [3, 4, 12, 13, 18]. As an early example, Han et al. [4] looked at the eye gaze patterns of patients with schizophrenia. They developed a three-party conversation scenario and presented it on a three-degree-of-freedom (3DoF) VR headset with an attached ViewPoint binocular eye tracker. They found that patients with schizophrenia actively avoided eye contact when interacting with two other virtual humans. Similarly, Raimbaud and colleagues [12] tested participants' gaze behaviours in front of a group of virtual audiences with different gaze patterns, using the FOVE VR (https://fove-inc.com/) headset's built-in eye-tracking system. Their results suggested that participants' eye gaze behaviours can be influenced by the eye gaze pattern of virtual human. People with higher social anxiety tended to have shorter dwell time towards virtual characters with directed gazes. In another study, Geraets et al. [3] developed a VR task for participants to recognise virtual humans' emotions. The eye tracking data from an integrated SensoMotoric Instruments (SMI) eye tracker showed that participants focused more on virtual humans' eyes and nose than the mouth for emotion recognition. Moreover, Rubin et al. [13] and Wechsler et al. [18] proposed that VR applications with eye-tracking could provide social anxiety intervention through attention guidance training. By using real-time eye-tracking data, people can follow professional suggestions to navigate their visual attention to specific areas in VR social scenarios, thereby reducing social anxiety.

It is essential to understand the hardware capabilities and limitations to achieving high accuracy in VR eye-tracking measurements when using eye-tracking for scientific research. While some eye-tracker manufacturers provide data quality estimates in VR for development guidance (e.g. Tobii reported an accuracy of 0.5°–1.1° for the eye-tracker in HTC Vive Pro Eye [17]), these estimated accuracies are typically based on the ideal performance. This may not be achievable in an actual experimental study. The quality of the eye-tracking signal is influenced by various factors, including the hardware manufacture, deployed platform, and the participants and experimental environment [8, 9]. Only a few studies have evaluated the accuracy and usability of VR eye-tracking headsets [14, 15], which has revealed discrepancies between official values and their actual performance in specific research use cases. Therefore, in-depth user testing is critical to establishing a solid foundation for researchers to design experimental studies. Meta Quest Pro is a VR headset with an integrated eye tracker released in October 2022. It estimates the direction of the user's eye gaze using an infrared camera and processes these data as abstracted gaze data (i.e., eye gaze direction) in real time [10], enabling eye-gaze data collection in VR. It allows VR applications to capture eye gaze data both in PC and standalone mode. It has a coherent development pipeline and workflow similar to previous Oculus VR devices like Quest 2. However, there is currently no public information available regarding its eye-tracking accuracy. The sparsity of empirical data presents a challenge to researchers wishing to design virtual scenarios that can produce reliable eye-tracking data.

To validate the use of Meta Quest Pro for VR eye-tracking research, we tested the usability (i.e. spatial accuracy, spatial precision, self-rated accuracy) of Quest Pro's eye-tracking capabilities in the head-free and head-restrained conditions. We carried out the testing with users with varying vision conditions. We intend to utilise the results of our initial analysis to inform the design decisions and technical implementation for designing VR eye-tracking studies.

2 METHOD

2.1 Test Design

12 volunteers (8 female, 4 male, average age: 27.92 [22, 39]) took part in this initial testing. Five of them conducted the VR task without vision correction (prescription range: [-2.5, 0]) and the other seven wore contact lenses (prescription range: [-4, -2.5]) during the data collection.

We used a within-subject study design to test spatial accuracy, spatial precision and self-rated performance for the eye-tracking task in head-free and head-restrained conditions. Each volunteer experienced the VR eye-tracking test in both conditions with a counter-balanced order.

2.2 Apparatus

We used a Windows 10 computer (Intel i7-8700K, NVIDIA GeForce GTX 1080, 32 GB RAM) to run the VR experience and display it on a Meta Quest Pro (Model: DK94EC) through a wire connection (Quest Link). Meta Quest Pro has a resolution of 1832*1920 pixels per eye and a claimed field of view of 106° (horizontal) × 96° (vertical). We recorded the data using the Quest Pro's built-in eye-tracking feature, and the headset was set up at a 90Hz refresh rate.

To minimize people's head movements during the head-restrained condition, we also designed a chinrest, where people placed their chin during the VR session.

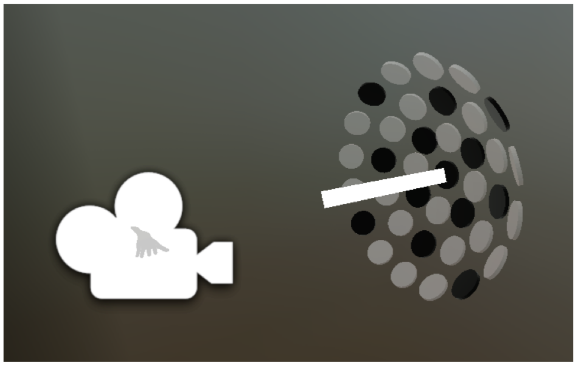
2.3 Software

We developed the VR stimulus in Unity 2021.3.15 and used Oculus' Movement SDK V1.3.2 [11]. We designed a 3D spherical object (Fig.2-a), rather than a plane, to place the target points so that all the targets have the same distance from the VR camera. The VR camera was placed in the origin (0,0,0) in Unity's world coordinate system. The 13 targets were placed in a visual field spanning ±15° horizontally and vertically° with a 5° interval (Fig.2-b). Each target was shown as a red dot measuring around 0.7° at a 1-meter viewing distance on a non-reflective dark environmental background, which helped reduce visual distraction.

For the head-free condition, the target points were fixed in the world coordinate system, and volunteers could rotate their heads freely to position the targets at the centre of their visual fields. For the head-restrained condition, the target points were attached to the camera and headset positional tracking was disabled so that the targets remained at a fixed position relative to the people's heads.

2.4 Procedure

Volunteers were invited to the VR lab for a single-time visit. They were first instructed to fit the Quest Pro headset and adjust the interpupillary distance (IPD) setting for an optimal view. After this, they completed a built-in 9-point eye-tracking calibration and verified the calibration results. Next, they performed the formal eye-tracking task in head-free and head-restrained conditions. The order of these two sessions was counter-balanced to control for any potential order effects. During each session, volunteers focused on 13 targets for 5 seconds each. They could control the pace of the test by pressing a button on the hand controller to proceed through



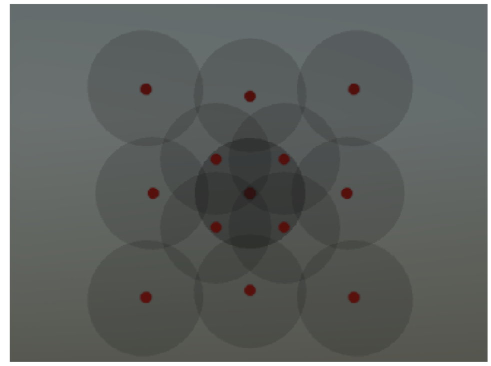


Figure 2: Eye Tracking Targets. a) A 3D spherical object to place the target points, all of which have the same distance with the camera; b) The 13 targets points in 2D view (in a visual field spanning ±15°)

targets. Each session took 2-3 minutes, after which they were asked to provide a subjective rate of their performance.

2.5 Measures

As the primary measures of eye-tracking data quality, we computed the spatial accuracy and spatial precision based on the definition from Holmqvist et al. [7, 8].

Spatial accuracy refers to the average angular offsets between the measured fixation positions and target positions among a series of samples as the generic formula (1). A smaller angular value indicates a higher eye-tracking data accuracy. Following the practice of calculating the angular offset in a 3D environment [1, 15], we computed the accuracy of each sample as the angular offset (in degree) between the 3D vector $\vec{t}$ (from the centre of the eyes to the target position) and the normalised 3D vector $\vec{g}$ (from the centre of eyes to the reported gaze position) as formula (2).

$$\theta_{\text{Offset}} = \frac{1}{n}\sum_{i=1}^{n}\theta_i \tag{1}$$

$$\theta = \arccos\left(\frac{\vec{t}\cdot\vec{g}}{||\vec{t}||||\vec{g}||}\right) * 180/\pi \tag{2}$$

Spatial precision measures the stability of individual measured gaze samples over time for a fixation target. The two common ways to compute spatial precision are based on standard deviation (SD) and the inter-sample root mean square (RMS) of gaze error. In this study, we examined both SD precision and RMS precision to evaluate eye-tracking performance. A smaller value of spatial precision indicates a higher quality of eye-tracking signal for both methods.

In addition to the spatial quality metrics, we also examined the volunteers’ subjective performance ratings in the eye-tracking task. After each session, they rated their accuracy in focusing on the targets on a scale from 1 (very poor) to 5 (very good). A higher rating indicated better subjective eye-tracking accuracy.

2.6 Data Processing and Analysis

We computed the eye-tracking signal quality after data aggregation at multiple levels. Data were pre-processed for all fixation targets within each condition for each volunteer. Following the standard practice to account for saccades latency and noises [7, 14], we excluded the first 100 samples for each target. We then removed the top and bottom 5% outliers. After the pre-processing, we had an average of around 650 samples per target point per person. Data filtering and processing were performed using MATLAB R2022b.

We compared the eye-tracking data quality in head-free and head-restrained conditions using the paired sample t-test after the assumption check. We reported the results with the effect size (cohen’s d) and 95% confidence interval (95% CI) of the difference between conditions. Analyses were conducted using R with RStudio 1.4.

3 RESULTS

We first screened data for noise and missing records. We excluded the data from one person because more than 15% of the data consisted of invalid samples. The remaining data were included in the analysis. We presented the aggregated eye-tracking data results of spatial accuracy and precision, with key descriptors and interquartile range (IQR). The results were summarised as the values across all fixations from all volunteers.

3.1 Spatial Accuracy

Table 1 presents the spatial accuracy results after outlier correction. The average spatial accuracy for the head-free condition was 1.652° (median: 1.635; SD: 0.496). The average spatial accuracy for the head-restrained condition was 2.162° (median: 2.015; SD: 0.692). A paired samples t-test revealed that the eye-tracking accuracy was significantly better in the head-free condition than in the head-restrained condition ($t(10) = -2.246$, $p = 0.049$, 95% CI = [-1.017, -0.004]).

Figure 3 displays the resulting frequency density histograms and the accuracy heat map based on the fixation targets span 15°.

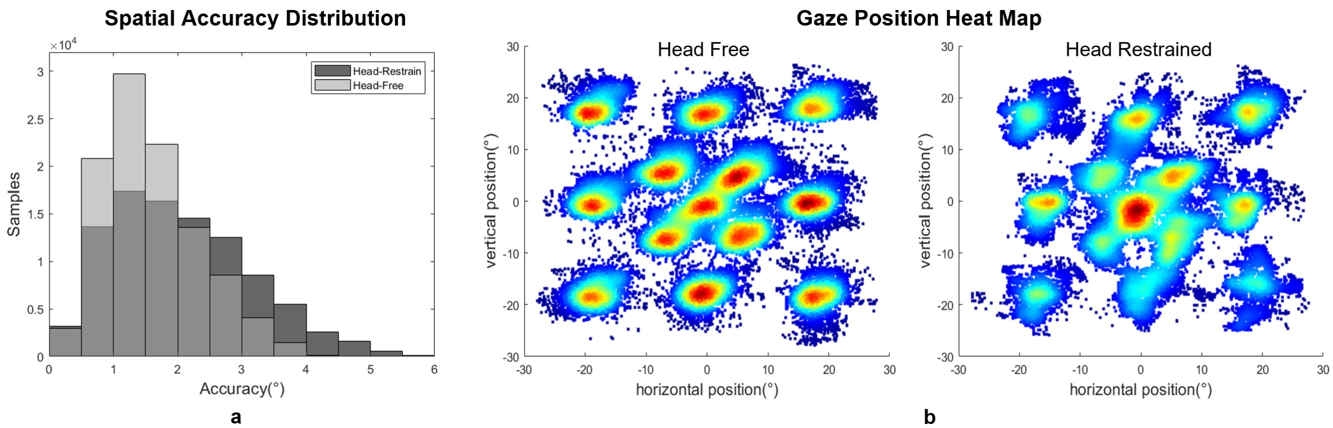


Figure 3: Eye Tracking Accuracy Result. a) Spatial accuracy distribution after outlier correction for head-free and head-restrained conditions; b) Gaze position heatmap for the 13 eye-tracking targets (graph created using the Flow Cytometry Data Reader and Visualization package[5]).

Table 1: Spatial Accuracy after Outlier Correction

	Spatial Accuracy (°)			
	Mean	*Median*	*SD*	*IQR*
Head-Free	1.652	1.635	0.496	[1.249, 1.813]
Head-Restrained	2.162	2.015	0.692	[1.785, 2.563]

3.2 Spatial Precision

We computed both the SD and RMS spatial precision shown in Table 2. For the head-free condition, the average SD precision was 0.699° (median: 0.702; SD: 0.198) and the average RMS precision was 0.849° (median: 0.898; SD: 0.173). For the head-restrained condition, the average SD precision was 0.673° (median: 0.694; SD: 0.167) and the average RMS precision was 0.772° (median: 0.632; SD: 0.461).

The paired samples t-test suggested there was no significant difference between the head-free and head-restrained conditions both in the SD precision ($t(10) = 0.745$, $p = 0.473$) and RMS precision ($t(10) = 0.692$, $p = 0.504$).

3.3 Self-evaluation and Eye-tracking Accuracy

Self-evaluation scores average was 4.73 (SD = 0.467) for head-free and 3.00 (SD = 0.894) for head-restrained conditions. A paired samples t-test suggested the subjective performance for the head-free condition was significantly better than the head-restrained condition ($p < 0.001$, Cohen's d = 0.905, 95% CI = [1.120, 2.335]).

A Spearman's correlation test was used to determine the relationship between the self-evaluation scores of the eye-tracking task and their performance. There was a significant correlation between self-evaluated performance and accuracy ($R_s = -0.843$, $p = 0.001$) specifically in the head-restrained condition, indicating that people who reported higher subjective performance also tended to exhibit higher actual accuracy during head-still.

4 DISCUSSION

This study presents the early results of evaluating Meta Quest Pro's eye-tracking feature by examining its spatial accuracy and precision. Volunteers with varying vision conditions performed a 13-point VR eye-tracking task both in head-free and head-restrained conditions and rated their performance. We tested the head-free condition to consider the effects of head movement and data spillage, and it is the most common interactive VR experience. We also examined the head-still condition to reference eye-tracking signal quality independent of head movement compensation. To our knowledge, this is the first empirical test to evaluate the accuracy of the eye tracker integrated in the Meta Quest Pro VR headset.

Our measurements indicated an average accuracy of 1.652° and an average precision of 0.699 (P_{SD}) and 0.849 (P_{RMS}) for a visual field spanning 15° for head-free VR experience. While there is no Quest Pro manufacturer data to benchmark with, relevant research has reported measured average accuracies of 1.08° for HTC Vive Pro Eye [14], 2.66° for HoloLens 2 [1], and 1.588° for an eye-tracking enabled Meta Quest 2 [2]. Our results are comparable to these VR/AR eye-tracking user studies with similar settings. Additionally, the eye-tracking quality of Quest Pro was generally better when targets were positioned at the centre of the field of view given all the targets lie at the same distance. This pattern aligns with the decreasing gaze accuracy for peripheral eye-tracking [8, 14, 15] for screen-based and HMD devices.

When comparing the head-free results against the head-restrained condition, we found both the measured and subjective eye-tracking accuracies were better when people could move their heads freely. The spatial precision was similar in both conditions. This finding suggests that using head movement to focus on areas of interest (AOI) manually provides an effective compensation mechanism for accurate fixation in VR. Interestingly, people tended to evaluate their eye-tracking accuracy well in head-restrained, but not head-free condition. This could be attributed to the noise and distortion of the eye-tracking signal during free movement. Also, it reflected the

Table 2: SD and RMS Spatial Precision

	Spatial Precision (°)							
	SD				RMS			
	Mean	*Median*	*SD*	*IQR*	*Mean*	*Median*	*SD*	*IQR*
Head-Free	0.699	0.702	0.198	[0.567, 0.818]	0.849	0.898	0.173	[0.670, 0.951]
Head-Restrained	0.673	0.694	0.167	[0.540, 0.789]	0.772	0.632	0.461	[0.415, 1.053]

potential influence of the interaction methods on user perception of the task and their ability to control their gaze accurately.

Overall, our preliminary testing suggests that Meta Quest Pro can be a capable tool for studying people’s visual attention in VR. Based on our measured error of 1.249°-1.813° (IQR) and a precision of over 0.5° at a 1-meter distance, researchers are suggested to carefully design the size of scene objects and scale the detection range of AOI accordingly. Additionally, researchers should consider the proximity between participants and key scene objects when designing the interaction flow, such as the interaction input method, locomotion, and the range of movement from the early development stage. This is necessary to manage the effect of viewing distance on eye-tracking accuracy. For instance, to determine whether a person is looking at the face of a real-life scaled VR human character, we recommend a distance of less than 3.5 meters or occupying an angular area of 2.5° to achieve reliable eye-tracking results. Moreover, researchers should be cautious when interpreting fixation results, especially when different AOIs are within close proximity. This is important for VR studies using Quest Pro to investigate subtle visual attention changes.

This preliminary test has several limitations and areas for future research. First, we combined the binocular eye gaze direction instead of computing separate measures for each eye or axis (horizontal versus vertical). As an early adoption of the eye-tracking API, we used the ray cast projection to compute the converged gaze points and calculated real-time gaze error. Future studies should evaluate eye-tracking signals for more detailed eye separation and HMD region. Second, we focused solely on the spatial quality of eye-tracking data and did not examine device latency and temporal errors. In order to assess the temporal accuracy, we would need to find the sample rate, as this is not provided in the hardware specifications sheet from the manufacturer. Third, we only conducted eye-tracking tests on PC VR with Oculus link. Further investigation into eye-tracking performance on Android standalone mode might lead to different results than those we obtained in the tethered mode and help inform decisions when choosing the deployment platform for eye-tracking VR. Fourth, each preliminary test was a brief session taking less than 4 minutes, and we did not look at how accuracy changes throughout an extended duration. Finally, we did not investigate the user experience, such as eye strain and comfort level of Quest Pro during the eye-tracking task. This is a critical human factor in understanding and enhancing the headset’s usability in VR research [6].

5 CONCLUSIONS

Our study evaluated the spatial accuracy and precision of the Meta Quest Pro eye-tracking signal. Our findings suggest it is a viable option for understanding user behaviours in immersive VR, as its signal quality is comparable to existing AR/VR eye-tracking headsets. These results highlight the importance of carefully designing the size of scene objects, mapping the areas of interest, and determining the interaction flow. Special caution is also needed when interpreting fixation results for the visual targets within a small proximity. We hope our preliminary investigation will inform future study designs for eye-tracking VR research and encourage the device manufacturer to provide better information transparency.

ACKNOWLEDGMENTS

The authors would like to thank Professor Daniel Freeman for his supervision. The authors would also like to thank Dr Michael Burch for his introduction to eye-tracking research, and Mr Dilmer Valecillos for his suggestions on OVR eye-tracking-related API.

REFERENCES

[1] Samantha Aziz and Oleg Komogortsev. 2022. An assessment of the eye tracking signal quality captured in the hololens 2. In *2022 Symposium on Eye Tracking Research and Applications*. 1–6.

[2] Ajoy S Fernandes, T Scott Murdison, and Michael J Proulx. 2023. Leveling the Playing Field: A Comparative Reevaluation of Unmodified Eye Tracking as an Input and Interaction Modality for VR. *IEEE Transactions on Visualization and Computer Graphics* (2023).

[3] CNW Geraets, S Klein Tuente, BP Lestestuiver, M Van Beilen, SA Nijman, JBC Marsman, and W Veling. 2021. Virtual reality facial emotion recognition in social environments: An eye-tracking study. *Internet interventions* 25 (2021), 100432.

[4] Kiwan Han, Jungeun Shin, Sang Young Yoon, Dong-Pyo Jang, and Jae-Jin Kim. 2014. Deficient gaze pattern during virtual multiparty conversation in patients with schizophrenia. *Computers in biology and medicine* 49 (2014), 60–66.

[5] Robert Henson. 2005. Flow cytometry data reader and visualization. *MATLAB Central File Exchange. Retrieved October* 9 (2005), 2015.

[6] Teresa Hirzle, Fabian Fischbach, Julian Karlbauer, Pascal Jansen, Jan Gugenheimer, Enrico Rukzio, and Andreas Bulling. 2022. Understanding, addressing, and analysing digital eye strain in virtual reality head-mounted displays. *ACM Transactions on Computer-Human Interaction (TOCHI)* 29, 4 (2022), 1–80.

[7] Kenneth Holmqvist, Marcus Nyström, Richard Andersson, Richard Dewhurst, Halszka Jarodzka, and Joost Van de Weijer. 2011. *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.

[8] Kenneth Holmqvist, Marcus Nyström, and Fiona Mulvey. 2012. Eye tracker data quality: What it is and how to measure it. In *Proceedings of the symposium on eye tracking research and applications*. 45–52.

[9] Anthony J Hornof and Tim Halverson. 2002. Cleaning up systematic error in eye-tracking data by using required fixation locations. *Behavior Research Methods, Instruments, & Computers* 34, 4 (2002), 592–604.

[10] Meta. 2022. Eye Tracking Privacy Notice. https://www.meta.com/en-gb/help/quest/articles/accounts/privacy-information-and-settings/eye-tracking-privacy-notice/

[11] Meta. 2022. Movement SDK for Unity. https://developer.oculus.com/documentation/unity/move-overview/

[12] Pierre Raimbaud, Alberto Jovane, Katja Zibrek, Claudio Pacchierotti, Marc Christie, Ludovic Hoyet, Julien Pettre, and Anne-Hélène Olivier. 2022. The

Stare-in-the-Crowd Effect in Virtual Reality. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 281–290.

[13] Mikael Rubin, Karl Muller, Mary M Hayhoe, and Michael J Telch. 2022. Attention guidance augmentation of virtual reality exposure therapy for social anxiety disorder: a pilot randomized controlled trial. *Cognitive Behaviour Therapy* 51, 5 (2022), 371–387.

[14] Immo Schuetz and Katja Fiehler. 2022. Eye tracking in virtual reality: Vive pro eye spatial accuracy, precision, and calibration reliability. *Journal of Eye Movement Research* 15, 3 (2022).

[15] Alexandra Sipatchin, Siegfried Wahl, and Katharina Rifai. 2021. Eye-tracking for clinical ophthalmology with virtual reality (vr): A case study of the htc vive pro eye's usability. In *Healthcare*, Vol. 9. Mdpi, 180.

[16] Vasileios Skaramagkas, Giorgos Giannakakis, Emmanouil Ktistakis, Dimitris Manousos, Ioannis Karatzanis, Nikolaos Tachos, Evanthia Eleftherios Tripoliti, Kostas Marias, Dimitrios I Fotiadis, and Manolis Tsiknakis. 2021. Review of eye tracking metrics involved in emotional and cognitive processes. *IEEE Reviews in Biomedical Engineering* (2021).

[17] Tobii. 2023. Hardware Accuracy. https://developer.tobii.com/xr/learn/eye-behavior/hardware-accuracy/

[18] Theresa F Wechsler, Michael Pfaller, Rahel E van Eickels, Luise H Schulz, and Andreas Mühlberger. 2021. Look at the audience? a randomized controlled study of shifting attention from self-focus to nonsocial vs. social external stimuli during virtual reality exposure to public speaking in social anxiety. *Frontiers in Psychiatry* 12 (2021), 751272.