UNIVERSITA' DEGLI STUDI "G. d'Annunzio"

Chieti-Pescara

SCUOLA SUPERIORE "G. d'Annunzio"

EARTHQUAKE AND ENVIRONMENTAL HAZARDS

Psycho-Sociological Analysis of Hazard Perception and Risk Communication

PhD Course

37° Cycle

Department of Psychological, Health and Territorial Sciences

SSD: M-PSI/08

THE USE OF VIRTUAL REALITY EXPOSURE THERAPY FOR THE TREATMENT OF PTSD FOLLOWING A SEISMIC EVENT

PhD Candidate

Dr. Marco Colasanti

3224249

Coordinator

Prof. Rita de Nardis

Tutor

Prof. Maria Cristina Verrocchio

Academic Years 2021-2024


BY

The Use of Virtual Reality Exposure Therapy for the Treatment of PTSD following a Seismic Event

Table of Contents

Table of Contents 3
Self-Presentation 4
Abstract 6
Thesis Presentation 8
Chapter 1 – Virtual Reality and Memory

1. Introduction 11
2. Materials & Method 14
3. Analysis and Results 23
4. Discussion 28

Chapter 2 – Virtual Reality and Stimuli Emotional Features

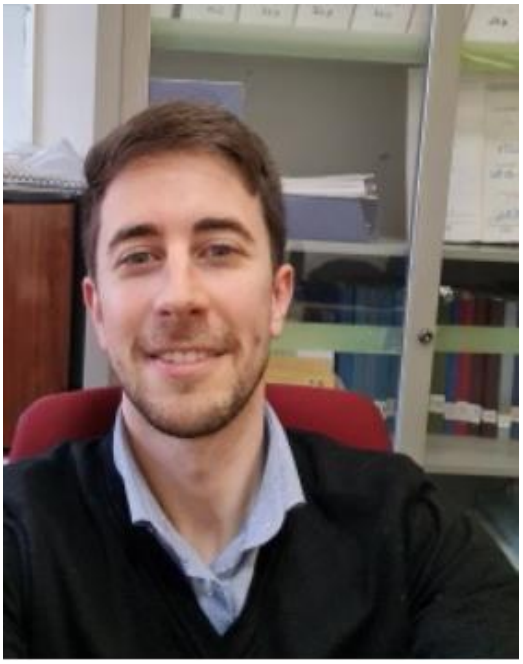
1. Introduction 31
2. Materials & Method 35
3. Results 44
4. Discussion 51

Chapter 3 – Exposure to a Virtual Seismic Event

1. Introduction 55
2. Materials & Method 58
3. Results 68
4. Discussion 71

General Results and Conclusions 77
Software and Resources 79
References 80
Supplementary Materials 99
List of Ph.D. Student Papers 104

Self-Presentation



Ph.D. Candidate: Marco Colasanti

Tutor: Prof. Maria Cristina Verrocchio

SSD: M-PSI/08

Characterizing Topic: Psycho-Sociological Analysis of Hazard Perception and Risk Communication

Contact: marco.colasanti@unich.it

I earned a bachelor's degree in Psychology and Health in 2016 and a master's degree in Forensic Psychology in 2018 at Sapienza University of Rome. After completing my professionalizing internship and registering with the Order of Psychologists of Lazio, I was awarded research grants from Sapienza University of Rome and “G. d'Annunzio” University of Chieti-Pescara, where I conducted research from 2019 to 2021. Since 2021, I have been pursuing a Ph.D. in Earthquake and Environmental Hazards. My research interests span forensic and clinical psychology, with a particular focus on the applications of virtual reality and emerging technologies in these fields. I am especially interested in their therapeutic potential within clinical settings, as I am currently training to become a cognitive-behavioural psychotherapist.

During my Ph.D., I completed a research period abroad at Erasmus University Rotterdam, where I had the opportunity to work within an international research environment. This experience allowed me to engage with researchers from different

disciplines and backgrounds, gain new perspectives, and appreciate the value of international collaboration in tackling complex scientific challenges.

To date, I have co-authored twenty-one research articles and book chapters, and contributed to twenty presentations at national and international conferences. Throughout my academic and professional journey, I have developed essential skills in interdisciplinary collaboration, research methodology, and scientific communication. The multidisciplinary environment of the EEH doctoral program has further strengthened my ability to work across disciplines, engaging with researchers from diverse fields to address complex challenges. I am confident that these experiences have prepared me for a career that integrates psychological research with innovative technologies to improve mental health outcomes.

Abstract

Introduction: The psychological impact of earthquakes extends far beyond immediate physical destruction, often leading to long-term trauma, including Post-Traumatic Stress Disorder (PTSD). Virtual Reality Exposure Therapy has been proposed as a promising intervention for trauma-related disorders, offering a controlled, immersive alternative to traditional exposure therapy. However, for VR to be effective, it is essential to understand how it compares to real-life experiences in terms of memory processing, emotional engagement, and physiological arousal. This research explores whether VR can serve as an ecologically valid substitute for real-world experiences and how its unique properties can be leveraged strategically for PTSD treatment.

Methods: The thesis comprises three studies. Chapters 1 and 2 examine memory encoding and recall in different presentation modalities (real life, VR, and 2D images) in a sample of 119 participants. Memory performance was assessed using free recall, visual recognition, and suggestibility tasks, with additional analyses of emotional valence and object typicality. Chapter 3 investigates participants' emotional, cognitive, and physiological responses to a VR-simulated earthquake scenario in a sample of 72 individuals. Emotional responses were assessed via self-reports, while physiological responses were measured through heart rate variability.

Results: Findings from Chapters 1 and 2 indicate that memory performance in VR differs significantly from real-life exposure, with real-world experiences leading to superior recall and VR responses being more comparable to those observed in a traditional 2D format. Emotional valence and object typicality influenced memory encoding differently across conditions, suggesting that VR does not inherently enhance memory accuracy or recall compared to 2D images. Chapter 3 revealed that VR exposure to a simulated earthquake successfully induced emotional and physiological arousal, with significant increases in state anxiety, heart rate, and sympathetic nervous system activation. However, no significant associations were found between individual psychological traits (e.g., trait anxiety, emotion regulation) and physiological responses.

Conclusion: Although VR does not fully replicate real-world memory processing, its controlled and moderate emotional engagement can be therapeutically

advantageous. Rather than overwhelming patients with highly realistic trauma recall, VR allows for a more gradual and controlled exposure experience, making it particularly valuable in the early stages of PTSD treatment. These findings provide a foundation for optimizing VR-based interventions, ensuring that they are tailored to balance emotional activation and patient safety in PTSD treatment.

Thesis Presentation

The psychological impact of natural disasters, particularly earthquakes, is an area of growing concern in disaster research and mental health interventions. While the immediate consequences of seismic events, such as infrastructural damage and economic losses, are often the primary focus, the long-term psychological effects on survivors cannot be overlooked. Many individuals develop Post-Traumatic Stress Disorder (PTSD) and other trauma-related disorders, experiencing recurrent distress, avoidance behaviours, and hyperarousal symptoms that can persist for years after the event. Developing effective and accessible treatment options for earthquake-related PTSD is, therefore, a crucial challenge in both clinical psychology and disaster response planning.

This thesis explores an innovative approach to trauma treatment by investigating the feasibility of Virtual Reality Exposure Therapy (VRET) for earthquake-related PTSD. While exposure therapy is widely recognized as one of the most effective treatments for trauma, it has notable limitations, including high patient dropout rates and patients' reluctance or difficulty in recalling the traumatic event. VR technology offers a compelling alternative by providing a controlled, immersive, and highly realistic simulation of traumatic events, allowing patients to confront their fears in a safe and therapeutic setting.

The research presented in this thesis unfolds in several key phases: first, evaluating memory performance in VR versus real-life conditions to assess its ecological validity; then, examining participants' response, both through self-reports and physiological measurements, to a simulated seismic event, experienced immersively through a head-mounted display. Studying memory performance in VR is a crucial first step because memory plays a central role in both trauma development and its treatment: PTSD is characterized by intrusive, distressing recollections of traumatic events, often influenced by distortions. Exposure therapy, including VRET, relies on the accurate reconstruction of traumatic experiences to facilitate emotional processing and desensitization. If memory functions differently in VR compared to real life, whether in terms of accuracy, retention, or susceptibility to misinformation, this has significant implications for the effectiveness of VR-based interventions. By first investigating how memory

encoding and recall operate in VR, this research establishes a necessary foundation for understanding whether VR exposure can effectively simulate real-world traumatic experiences and support therapeutic outcomes.

Chapter 1 presents a research article, published in *Virtual Reality*, that investigated the impact of different presentation modalities (i.e., real-life exposure, virtual reality, and 2D images) on memory performance and suggestibility. A total of 119 participants experienced the same environment through one of these mediums and completed free recall, visual recognition, and suggestibility tasks. Results showed that real-life exposure led to the highest memory performance, with no significant differences between VR and 2D conditions in most measures. This research provides critical insights into the effectiveness of VR as a tool for cognitive and psychological studies. By directly comparing real-life, VR, and 2D conditions, the study highlights potential limitations of VR in replicating real-world memory recall. It underscores the importance of considering not only presence but also interactivity and multisensory engagement when designing VR environments for cognitive tasks.

Chapter 2 presents a research article, published in *Psychological Research*, examining how stimuli emotional valence, emotional arousal, and typicality influence memory performance in different encoding contexts: real-life exposure, virtual reality, and 2D images. Using the same experimental approach as Chapter 1, findings revealed that in VR and 2D conditions, positive items were recalled more frequently than negative ones, whereas typicality influenced recall only in real-life settings, where atypical objects were better remembered. The findings emphasize the importance of emotional valence and object typicality in memory encoding, showing that VR and 2D contexts may engage different cognitive mechanisms than real-world settings. Taken together, Chapter 1 and 2 raise important implications for forensic and educational applications, where VR is increasingly used as a substitute for real-life experiences.

Chapter 3 introduces the experimental study investigating the feasibility of using VR as an exposure tool for simulated seismic events, aiming to lay the groundwork for its potential application in PTSD treatment. This research assesses the realism, usability, and emotional impact of a VR earthquake scenario, as well as the role of

individual differences, such as trait anxiety, personality traits, and cognitive emotion regulation strategies, in shaping participants' responses. A sample of 72 participants underwent a controlled VR exposure to an earthquake simulation while their emotional, cognitive, and physiological responses were assessed. Results indicated that participants perceived the VR scenario as moderately realistic and representative, with significant increases in state anxiety, heart rate, and sympathetic nervous system activation during exposure. However, no significant associations were found between individual psychological factors and physiological responses. This research provides crucial insights into the application of VR as a tool for trauma-related interventions, particularly for earthquake-related PTSD. By demonstrating that VR can elicit significant emotional and physiological responses while maintaining minimal adverse effects, the study supports the potential of VR-based exposure therapy as a complement to traditional methods. The findings highlight the need for further research into optimizing VR scenarios to enhance realism, immersion, and emotional engagement in clinical populations, and the necessity to leverage VR's unique properties, such as interactivity and multisensoriality.

The EEH Ph.D. program focuses on multidisciplinary research addressing the impacts of seismic and environmental hazards on human systems. Within this framework, this thesis contributes to the intersection of psychology and disaster response by exploring how virtual reality can be used to mitigate the psychological consequences of earthquakes. By leveraging VR's highly immersive and controlled environments, this study aligns with EEH's mission to develop innovative, interdisciplinary solutions to environmental hazards. The findings contribute to the understanding of human responses to seismic events, offering insights that are relevant not only for clinical interventions but potentially for disaster preparedness, emergency training, and resilience-building strategies. The findings also lay the groundwork for future clinical applications of VR-based therapy, potentially transforming the way PTSD related to natural disasters is treated. This work highlights the necessity of multidisciplinary collaboration in tackling complex global challenges, reinforcing the importance of innovative, technology-driven approaches in environmental hazard research.

Chapter 1

Virtual Reality and Memory

Monaro, M., Mazza, C., Colasanti, M. et al. (2024). Testing memory of a VR environment: comparison with the real environment and 2D pictures. *Virtual Reality*, 28, 100. https://doi.org/10.1007/s10055-024-00999-w

1. Introduction

Over the past decade, the increasing availability and decreasing cost of virtual reality (VR) technology has led to a considerable rise in research exploring VR applications. VR has become an especially popular research modality in experimental psychology, due to the almost limitless possibilities for creating complex and realistic scenarios with a high degree of experimental control while, supposedly, providing higher ecological validity than laboratory settings. Indeed, research has assumed that VR and reality are somewhat comparable, as long as certain conditions are met. In the early implementation of VR research designs, it was hypothesized that the similarity of users' responses between real-world and VR environments would be proportional to the degree to which the VR setting simulated "naturalistic" experiences (Bell et al., 2001)—in other words, its degree of realism (i.e., how faithfully it represented real-world input on all sensory channels and the fidelity of its environmental responses; Freeman et al., 2000). To this end, research has highlighted the role of a series of constructs in determining users' perception of VR environments. Among these, two closely related concepts—presence and immersion—seem significant. Presence indicates the subjective "sense of being there" (Freeman et al., 1999), while immersion is related to the objective properties of a system (Slater, 2009), in terms of its replication of vivid, multisensory perceptions. On this point, it should be noted that the umbrella term "VR" may apply to a variety of devices (e.g., flat screen three-dimensional [3D] environments, head-mounted displays [HDM], room-based systems such as the Cave Automatic Virtual Environment [CAVE]) offering different levels of immersion. It has been suggested that VR can elicit a sense of presence that is akin to being present in real life, and thus greater than the sense of presence that can be achieved through interaction with traditional 2D media (e.g., Wagler & Hanus, 2018).

While the literature seems to agree on this high sense of subjective presence in VR settings, the implications of this are debated. For instance, some researchers (e.g., Hodges et al., 1994; North et al., 1998) have linked the sense of presence in VR settings with increased emotional arousal, but not to improved task performance. Mania and Chalmers (2001) compared recollection performance of a 15-minute seminar delivered across four different conditions (i.e., in person, on a 3D desktop, through a 3D HMD, via audio), finding that level of presence was not associated with accurate memory recall, and that recall was significantly higher in the real-life condition, compared to the VR condition. Similarly, Slater et al. (1996) found that immersion—but not presence—increased task performance, which involved comprehension and memory of a complex 3D object. However, other studies investigating performance differences between VR and real-life environments have found contradictory results. For instance, Hu-Au and Okita (2021), in a study assessing environmentally-related learning differences, found comparable learning of general content knowledge in VR and real-life conditions. Conversely, Taylor and Dando (2018) compared episodic retrieval performance during interviews in a virtual avatar-to-avatar environment (i.e., with both interviewer and interviewee represented by avatars) and a traditional face-to-face environment, finding that participants in the avatar-to-avatar interview had significantly better recall. Research has also investigated user experience and performance differences between two-dimensional (2D) display environments and 3D virtual environments, finding that: at the lowest level, the main difference between these environments is that VR provides users with a sense of depth and proportion that is lacking in traditional 2D media; and at higher levels, VR generally induces a stronger sense of presence and engagement (e.g., Radianti et al., 2020). Indeed, the international literature highlights that VR movies generate different EEGs and greater emotional arousal in viewers compared to 2D movies (Tian & Whang, 2021). Similarly, men (but not women) find VR pornography more sexually arousing than its 2D counterpart (Elsey et al., 2019). Some authors (see, e.g., Elmquaddem, 2019) advocate for the use of VR as a learning tool, positing that VR "can improve and facilitate learning, increase memory capacity and make better decisions while working in entertaining and stimulating conditions" (p. 237). There is some

corroboration for this claim. For instance, Schöne et al. (2017) reported that participants who watched a motorcycle ride via VR not only rated their experience as more realistic but also performed twice as well in a memory task than participants who experienced the same motorcycle ride via 2D video. Likewise, Krokos et al. (2019) found superior memory recall with an HMD compared to a traditional 2D desktop computer. Similarly, Norman et al. (2020) found a greater skin conductance response (taken to indicate recognition) to a mock crime scene presented in VR compared to 2D. Surprisingly, most studies investigating recall following exposure in VR in comparison to other modalities have not examined suggestibility. From a forensic perspective, interrogative suggestibility (i.e., the extent to which, within a closed social interaction, messages communicated during formal questioning are accepted, with a subsequent effect on behavioural responses; Gudjonsson & Clark, 1986) could be an interesting variable to consider, as insight into this factor might contribute to the development of ecological mock crime scenarios.
However, two main problems arise: on the one hand, VR results are not consistent, as there are reports of both similar and worse memory performance based on interaction with VR, compared to 2D media (e.g., Ernstsen et al., 2019; Makransky et al., 2019; Kisker et al., 2021b). Furthermore, an important caveat of research applying VR to learning is that, for VR to be effective, it must leverage its unique advantages, which include both presence and immersivity and embodiment and agency (e.g., Johnson-Glenberg, 2019; Johnson-Glenberg et al., 2021). Indeed, research has confirmed that the utility of VR is dependent on a high degree of presence, immersion, and interactivity (e.g., Sutcliffe et al., 2005).
In light of this observation and the inconsistent findings on the differences between VR versus other media, the present study aimed at gaining an understanding of how the modality through which stimuli are presented (i.e., 2D vs. VR vs. real life) impacts memory recollection and suggestibility. For this purpose, three groups of participants were exposed to, respectively, a room in real life, the same room in VR, and 2D pictures of the same room captured from different angles. The following hypotheses were formulated: (a) participants in the VR condition would perform similarly to participants in real-life condition on free recall, visual recognition, non-suggestive visual and verbal questions, and resistance to suggestibility (verbal and

visual questions) tasks; and (b) participants in the VR condition would perform significantly better than participants in the 2D picture condition on free recall, visual recognition, non-suggestive visual and verbal questions, and resistance to suggestibility (verbal and visual questions) tasks.

2. Materials & Method

2.1 Participants

A total of 123 participants were volunteers who responded to a social media advertisement or were located near Sapienza. The inclusion criteria were: (a) aged at least 18 years; and (b) excellent comprehension of the Italian language. Four participants (3.25%) were excluded due to set-up issues related to the Meta Quest 2 device used in the study, which invalidated the procedure. The final sample comprised 119 participants, of whom 62 were male (52.1%) and 57 were female (47.9%), aged 18–35 years ($M = 24.20$, $SD = 4.13$). The majority of the sample were students ($N = 77$, 64.7%), educated to a high school level ($N = 67$, 56.3%), Italian citizens ($N = 118$, 99.2%), living in central Italy ($N = 107$, 89.9%), and experiencing no visual impairments ($N = 61$, 51.3%). A post hoc power analysis was computed using G*Power (Faul et al., 2007): a sample size of 119 resulted to be sufficiently large to achieve a statistical power (1-β) of at least 0.90 in a testing involving three groups, given a significance level of 0.05 and a large effect size (0.40).

Participants were randomly assigned to three groups using the Excel RAND function, according to a manipulated variable (i.e., the modality in which they visited the target room; see the "Measures" section for detailed information):

-Group 1 (G1) ($M_{age} = 26.53$, $SD = 3.60$) was composed of 40 participants who visited the target room in real life.

-Group 2 (G2) ($M_{age} = 23.25$, $SD = 4.19$) was composed of 40 participants who visited the target room in VR, using a Meta Quest 2 device.

-Group 3 (G3) ($M_{age} = 22.79$, $SD = 3.60$) was composed of 39 participants who observed 2D pictures of the target room (captured from different angles) on a computer.

The mean age was statistically different between the three groups ($F(2,116) = 11.34$; $p < .001$).

Table 1 reports the descriptive statistics for all of the characteristics considered, for each group and for the entire sample.

Table 1

Descriptive Statistics of the Sample and each Group.

	G1 *n* (%)	G2 *n* (%)	G3 *n* (%)	Total *N* (%)	χ^2	*p*
Biological sex					1.23	.540
Female	22 (55.0)	18 (45.0)	17 (43.6)	57 (47.9)		
Male	18 (45.0)	22 (55.0)	22 (56.4)	62 (52.1)		
Educational level					18.51	**<.001**
≤ High school diploma	12 (30)	27 (67.5)	29 (74.4)	68 (57.1)		
≥ Bachelor's degree	28 (70)	13 (32.5)	10 (25.6)	51 (42.9)		
Occupation					15.31	**.004**
Unemployed	1 (2.5)	1 (2.5)	2 (5.1)	4 (3.4)		
Student	17 (42.5)	30 (75)	30(76.9)	77 (64.7)		
Employed	22 (55)	9 (22.5)	7 (18)	38 (31.9)		
Vision impairment					.67	.717
No	19 (47.5)	20 (50)	22 (56.4)	61 (51.3)		
Yes	21 (52.5)	20 (50)	17 (43.6)	58 (48.7)		
Use of glasses / contact lenses					.22	.894
No	23 (57.5)	21 (52.5)	22 (56.4)	66 (55.5)		
Yes	17 (42.5)	19 (47.5)	17 (43.6)	53 (44.5)		
Medical/psychological/psychiatric diagnosis					2.10	.351
No	38 (95)	40 (100)	37 (94.9)	115 (96.6)		
Yes	2 (5)	0 (0)	2 (5.1)	4 (3.4)		
Prior experience of VR					1.07	.586
No	30 (75)	26 (65)	26 (66.7)	82 (68.9)		
Yes	10 (25)	14 (35)	13 (33.3)	37 (31.1)		

2.2 Measures

2.2.1 Measures Used in Phase 1 (see "Experimental Procedure" section)

Sociodemographic Questionnaire. Participants were administered a questionnaire to collect personal sociodemographic information on biological sex, age, education, occupational status, region of residence, citizenship, medical diagnoses, visual impairments, and prior experience with VR.

Rivermead Behavioural Memory Test (RBMT-III), "Figure Recognition" Subtest. The Rivermead Behavioural Memory Test (Wilson et al., 1985; Italian validation: Beschin et al., 2013) is an ecological assessment instrument that evaluates respondents' ability to use memory in everyday situations. Showing good ecological validity, it has great value in predicting real-life behavior and deficits outside the evaluation situation. The measure is composed of 14 subtests, aimed at evaluating visual memory, verbal memory, and recall memory aspects, both immediate and delayed. The present study administered the "Figure Recognition" subtest, which aims at testing respondents' ability to recall previously displayed images from a larger set.

Corsi Block-Tapping Test. The Corsi block-tapping test (De Renzi & Nichelli, 1975; Spinnler, 1987) is one of the most popular and widely used tests for measuring the quantity of information that can be held in short-term memory, otherwise known as visuospatial memory span. The stimulus is a board (32×25 cm) on which nine black cubes ($4.5 \times 4.5 \times 4.5$ cm) are attached asymmetrically. The cubes are progressively numbered on the face displayed to the examiner, who sits opposite the participant. The examiner taps the cubes in a prearranged sequence of increasing length (tapping a cube every 2 seconds). Immediately after the examiner finishes the sequence, the participant is asked to reproduce it, touching the cubes in the same order. The length of the sequence varies from 3 (the shortest) to 10 (the longest), and for each length, there are two prearranged sequences. If the participant correctly reproduces one of the sequences shown, the examiner progresses to a longer sequence. The participant's visuospatial memory span is reflected by the number of cubes related to the longest series correctly reproduced. The average visuo-spatial memory span is five (Spinnler, 1987).

2.2.2 Measures Used in Phase 3 (see "Experimental Procedure" section)

Free Recall Task. Participants were asked to write down all of the objects they remembered seeing in the target room. Specifically, the instructions were: "You have just seen a room. Please list in writing (without describing) all of the objects in the room. Let the examiner know when you have finished". Subsequently, participants were given 1 point for each object (out of 50) they were able to recall. The free recall task total score was based on how many objects of the target room

participants were able to recall. The 50 items are listed in the Supplementary Materials (Table S1).

Ad-hoc Questionnaire. An ad-hoc questionnaire was created and administered through the online software Qualtrics. The questionnaire included 60 questions related to the target room, meant to detect visual recognition, non-suggestive verbal and visual questions, and suggestibility (through verbal and visual suggestive questions). The questionnaire comprised two sections:

Section 1, Visual Recognition Task. Participants were shown a picture of an object/furniture item/painting and asked if they previously saw it in the room. Ten pictures represented items that were actually in the target room, while an additional 10 pictures depicted items that were not in the target room. Specifically, the question was as follows: "*Was the object/furniture item/painting in the room?*" Participants were allowed to write their answer, and they did not necessarily have to answer "*Yes*" or "*No*". Figure 1 displays two items included in the visual recognition task.

Figure 1

Two Items from the Visual Recognition Task Note



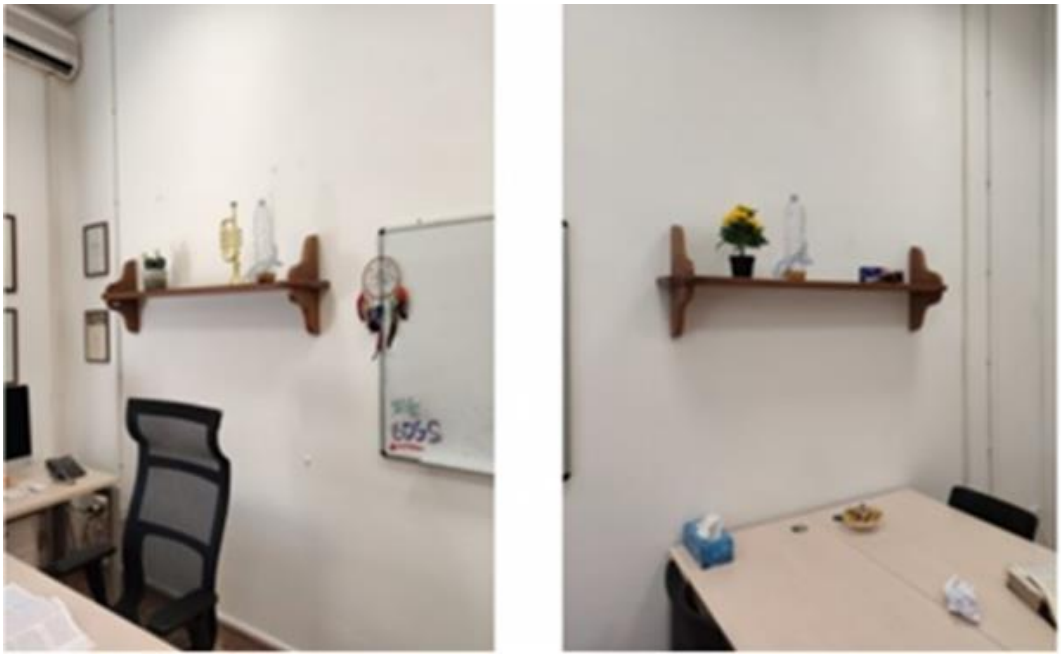
Note. The image on the left was on the wall in the target room, while the image on the right was not present in the target room.

Section 2, Suggestibility Task. The suggestibility task of the present study mimicked the structure of the Gudjonsson Suggestibility Scale-2 (GSS-2; Gudjonsson, 1997) – adapting it to the research purpose and including a visual task

– and presented participants with suggestive and non-suggestive (verbal and visual) questions related to the target room. The GSS-2 is a tool designed to measure interrogative suggestibility, which represents a person's propensity to accept information communicated during formal questioning with a subsequent influence on their responses. Specifically, the suggestibility task employed in the present study comprised: (a) 10 non-suggestive verbal questions, (b) 5 non-suggestive visual questions; (c) 15 suggestive verbal questions; and (d) 10 suggestive visual questions. The non-suggestive verbal questions included five questions concerning true details of the target room (e.g., "*Was there a calendar in the closet?*") and five questions concerning false details (e.g., "*Was the mini fridge green?*"; note that there was a fridge in the room, but it was blue). The non-suggestive visual questions showed five pairs of photographs in which only one of each paired alternative represented reality (see Figure 2 for an example question). The suggestive verbal questions asked about objects/furniture items/details that were not present in the target room (e.g., "*Was the backpack on the chair broken?*"; "*Was the carpet red or green?*"; note that there were no broken chairs or carpets in the room). Finally, the suggestive visual questions presented 10 pairs of photographs, both depicting in different locations an object that was not in the target room. Participants were, then, asked which of the two photographs represented the object's actual position, despite neither alternative was correct (see Figure 3 for an example question). For this task and to allow participants to choose neither, the response mode was open (i.e., participants were not forced to answer "*Yes*" or "*No*" or "*True*" or "*False*", as in the GSS-2).

Figure 2

One Pair of Alternatives Presented in the Non-suggestive Visual Section



Note. The question related to this stimulus was: "*Was the bottle near the trumpet or near the flower pot*?". note that the correct answer was "*near the flower pot*".

Figure 3

One Pair of Alternatives Shown in the Suggestive Visual Section



Note. The question related to this stimulus was: "*Was the backpack on the floor or on the chair*?" note that there was no backpack in the target room.

2.3 Experimental Procedure

Data were collected in October 2021. The experimental procedure was conducted during working hours (9:00–17:00) to ensure adequate lighting conditions and took place in a neutral room and a target room of the Department of Human Neuroscience, "Sapienza" University of Rome. The experiment was designed in accordance with the Declaration of Helsinki and approved by the local ethics committee (Board of the Department of Human Neuroscience, Faculty of Medicine and Dentistry, Sapienza University of Rome). The experimental procedure lasted approximately 30 min and consisted of three phases: (1) assessment of participants' visual-spatial memory through neuropsychological tests, (2) exposure to the target room (real life vs. in VR vs. through 2D pictures), and (3) completion of the free recall, visual recognition, and suggestibility tasks in relation to the target room.

2.3.1 Phase 1

After providing written informed consent, participants completed the sociodemographic questionnaire and underwent a visual-spatial memory assessment through the Rivermead Behavioural Memory Test III "Figure Recognition" subtest and the Corsi block-tapping test (see the "Measures" section). These tests were useful to check the cognitive abilities of the participants, making sure that they had no visual-spatial memory deficits that could interfere with performance in the experimental task (phases 2 and 3).

2.3.2 Phase 2

Participants were randomly assigned to one of three experimental conditions: real life, VR, and 2D pictures.

Group 1 (G1): Real-life Condition. Participants were taken into a target room of the Department of Human Neuroscience and positioned in the middle of the room. They were asked to memorize as many objects as possible over a period of 2 min, after which they were called by the experimenter. Participants had to stay in the middle of the room and could only rotate their body. Specifically, the instructions were as follows: "*We will now get you into a room. Your task is to observe the room carefully for 2 minutes, trying to memorize as many details as possible. We ask you to stand still at this point. You can turn your head in all directions and rotate around yourself*." Figure 4 presents an image of the target room.

Figure 4

Picture of the Target Room



Group 2 (G2): VR Condition Using a Meta Quest 2. Participants were accompanied into a neutral room and asked to wear a Meta Quest 2 visor in order to visit the target room in VR. The Meta device has a 72 Hz LCD screen with a resolution of 1832 × 1920 pixels per eye. The visor is worn in front of the eyes and covers the entire field of vision. It also comprises two hand-held knobs that simulate hands. In the present study, only one knob was used, in order to permit participants to virtually access the room. Through the Meta Quest 2, participants were shown a panorama 360° picture of the target room, taken by a professional photographer with a Lapbano Pilot One EE. The 360° picture was taken from the same point where participants in Group 1 were standing.

The target room was the same as the room Group 1 explored in real life. After receiving guidance on the use of the Meta Quest 2 (e.g., that they should not move outside the planned area, and that they had to physically turn their head and body to see all parts of the room), participants were asked to memorize as many objects as possible over a period of 2 min, after which they removed the visor. Specifically, the instructions were as follows: "*Now we are going to give you a virtual reality experience. You will be in a room that you can visually explore in 360°. You will use this visor. Your task is to observe the room carefully for 2 minutes, trying to memorize as many details as possible. We ask you to stand still at this point. You can turn your head in all directions and rotate around*."

Following this step, participants were asked to answer two questions to assess their sense of presence inside the virtual environment. The first question (i.e., "*I felt completely immersed*") was adapted from Jennett et al.'s (2008) scale, as previously applied in other studies (e.g., Hudson et al. 2019); participants responded using a 7-point Likert scale ranging from 1 (*Not at All*) to 7 (*Very Strongly*). The second question (i.e., "*I felt like I was inside the room*") was adapted from Wagler and Hanus's (2018) scale of spatial presence, and was measured on a 7-point Likert scale ranging from 1 (*Not at All*) to 7 (*A Lot*).

Group 3 (G3): 2D picture condition. Participants were accompanied into a neutral room with a computer. They were asked to sit in front of the computer and look at some 2D pictures of the target room on the computer screen. Participants had 2 minutes to look at these pictures and memorize as many objects as possible. The eight pictures (2048 × 1537 pixels; see Supplementary Materials, Figure S1) showed different parts of the target room from the same point of view of participants in the real life and VR conditions. The pictures were sequentially shown on a 27" computer monitor, and participants could scroll across them as they wanted (see for example Figure 5). Specifically, the instructions were as follows: "*Now you will be shown some pictures of a room. Your task is to observe the room carefully for 2 minutes, trying to memorize as many details as possible*."

Figure 5

Picture of the Target Room Used in the 2D Picture Condition



2.3.3 Phase 3

After exposure to the target room, all participants were taken into a neutral room and administered the free recall task (see the "Free Recall Task" section). Then, they completed an ad-hoc questionnaire about the room they observed on a 27" personal computer, with no time limit (see the "Ad-hoc Questionnaire" section). It was underlined that participants could answer according to their preference, and did not need to indicate "*Yes*" or "*No*."

3. Analysis and Results

3.1 Data Analysis

One-way independent ANOVA models were run to test performance differences between the three experimental groups (i.e., G1, G2, G3) in free recall, visual recognition, non-suggestive verbal questions, non-suggestive visual questions, suggestive verbal questions, and suggestive visual questions. The effect sizes of the score differences between groups were reported; with respect to magnitude, η^2 =

0.01 was considered indicative of a small effect, $\eta^2 = 0.06$ a medium effect, and $\eta^2 = 0.14$ a large effect (Cohen, 1988). To address the problem of multiple testing, Bonferroni correction was applied, dividing the p-value by the number of tested variables (n = 6) and setting the significance level to 0.008 (Shaffer, 1995).

As the three groups differed for age, educational level, and occupational status, to minimize any bias coming from the differences in covariates across groups, we employed two matching algorithms (i.e., the nearest neighbour matching and the coarsened exact matching) able to balance covariance discrepancies across groups through weights. We evaluated the performance of both algorithms and given the poor performance of nearest neighbour matching, we reported results using the coarsened exact matching (Iacus et al., 2012). The descriptive statistics and the density plots of the covariates before (pre-) and after (post-) matching procedure are reported in Supplementary Materials (Table S2, Figure S2). We then used regressions of each outcome on the experimental condition and covariates including the matching weights to estimate the average effects of the experimental manipulation and tested the null hypothesis of no effect of the experimental manipulation. We included the covariates in the final regression as they can provide additional robustness to imbalances remaining after matching and can augment precision.

All analyses were performed using the SPSS v.28 software (IBM, 2021) and R (R Core Team, 2021).

3.2 Results

3.2.1 Memory Performance

Table 2 reports each group's average scores and standard deviations, and the ANOVA results. The ANOVAs generated significant results with respect to the free recall task, and the visual recognition task. Moreover, the ANOVAs indicated a significant effect of the experimental manipulation for the non-suggestive verbal questions, suggestive verbal questions, and suggestive visual questions. No significant results emerged from the ANOVA that explored differences between groups in relation to the non-suggestive visual questions.

Table 2

Average Scores (M) and Standard Deviations (SD) for Each Experimental Group (G1, G2, G3) on the Free Recall, Visual Recognition, Non-suggestive (Verbal and Visual) and Suggestive (Verbal and Visual) Tasks, and the Results of the One-Way Independent ANOVA Models (F-test, p-value, η^2).

	G1	**G2**	**G3**	***F***	***p-value***	***η²***
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	df (2,116)		
Free recall task	17.95 (3.75)	10.95 (3.04)	11.54 (3.48)	50.97	<.001	0.47
Visual recognition task	17.18 (1.57)	14.18 (2.05)	13.80 (2.31)	34.24	<.001	0.37
Non-suggestive questions						
Verbal	9.28 (0.82)	8.43 (1.06)	7.85 (1.51)	15.12	<.001	0.21
Visual	4.70 (0.52)	4.48 (0.96)	4.13 (0.98)	4.59	0.012	0.07
Suggestive questions						
Verbal	1.75 (1.50)	3.05 (1.96)	4.08 (2.36)	13.92	<.001	0.19
Visual	3.10 (2.59)	6.83 (2.85)	7.21 (2.52)	28.97	<.001	0.33

Note. G1: Real life Condition; G2: VR Condition; G3: 2D picture condition.

Results using the matching approach, which considers the differences in covariates across groups in age, educational level, and occupation, are mainly consistent with the ANOVAs results (without considering any covariate). Table 3 reports the difference in mean outcomes (average effect of the experimental condition) between the participants assigned to the different groups, after applying the coarsened exact matching. The main results are the following.

Table 3

Difference in Mean Outcomes (Average Effect of the Experimental Condition) between the Participants Assigned to the Different Groups, after Applying the Coarsened Exact Matching

	Estimate	**Pr(>|z|)**	**2.50%**	**97.50%**
Free recall task				
G2 vs G1	-5.9	<0.001	-7.77	-4.04
G3 vs G1	-5.71	<0.001	-7.16	-4.26
G3 vs G2	0.726	0.299	-0.643	2.09
Visual recognition task				
G2 vs G1	-2.53	<0.001	-3.29	-1.76
G3 vs G1	-2.76	<0.001	-3.19	-2.34
G3 vs G2	0.0391	0.945	-1.08	1.16
Non-suggestive verbal questions				

G2 vs G1	-0.661	0.0441	-1.31	-0.0176
G3 vs G1	-1.32	<0.001	-1.42	-1.21
G3 vs G2	-0.491	0.00787	-0.853	-0.129
Non-suggestive visual questions				
G2 vs G1	-0.409	0.189	-1.02	0.202
G3 vs G1	-0.699	<0.001	-0.86	-0.538
G3 vs G2	-0.402	0.0509	-0.806	0.00166
Suggestive verbal questions				
G2 vs G1	0.603	0.0568	-0.0175	1.22
G3 vs G1	2.2	<0.001	1.8	2.59
G3 vs G2	0.834	0.0497	0.000931	1.67
Suggestive visual questions				
G2 vs G1	3.09	<0.001	1.7	4.48
G3 vs G1	4.52	<0.001	3.91	5.12
G3 vs G2	-0.353	0.627	-1.78	1.07

Free Recall Task. It emerged a statistically significant difference between G1 and G2 and between G1 and G3. In contrast, no differences emerged between G2 and G3. This indicates that participants who saw the room in real life had better recall of the room details compared to participants who explored the same room in VR or through 2D pictures.

Visual Recognition Task. The analysis revealed a statistically significant difference between G1 and G2 and between G1 and G3. In other words, participants who were exposed to the room in real life performed better on the visual recognition task than participants who saw the same room in VR or through 2D pictures. No significant differences emerged between G2 and G3.

Suggestibility Task.

Non-suggestive verbal questions. There was a statistically significant difference between G1 and G3 and between G2 and G3. In contrast, no difference emerged between G1 and G2. These results indicate that participants who were exposed to the target room in real life had more accurate verbal recall than participants who were exposed to the same room through 2D pictures but not than participants who were exposed to the same room in VR. Moreover, participants who were exposed to the target room in VR had a more accurate performance compared to those exposed to 2D pictures.

Non-suggestive visual questions. The analysis revealed a statistically significant difference between G1 and G3. This indicates that participants who were exposed

to the target room in real life had more accurate visual recall than participants who were exposed to 2D pictures. No significant differences emerged between G1 and G2, and between G2 and G3.

Suggestive verbal questions. The analysis showed a statistically significant difference between G1 and G3 suggesting that participants in real life were significantly more resistant to verbal suggestions than those in the 2D condition. There was no significant difference between G2 and G1 and G2 and G3. Suggestive visual questions. Statistically significant differences emerged between G1 and G3, and between G1 and G2, whereas no difference emerged between G2 and G3, suggesting that participants in real-life condition were significantly more resistant to visual suggestions than those in the other two experimental conditions.

3.2.2 Cognitive Ability

To rule out the possibility that the three groups differed for cognitive ability, rather than the experimental condition, one-way independent ANOVAs were run to compare the performance of the three groups (i.e., G1, G2, G3) on the visual-spatial memory tests administered in Phase 1 of the experimental procedure (i.e., RMBT-III "Figure Recognition" subtest, Corsi block-tapping test). No significant results emerged for either the RMBT-III "Figure Recognition" subtest ($F_{(2,116)} = 0.95$, $p = .390$, $\eta^2 = .016$) or the Corsi block-tapping test ($F_{(2,116)} = 0.72$, $p = .490$, $\eta^2 = .012$), suggesting that there were no differences between groups in terms of basic memory skills (i.e., visual recognition, visual span).

3.2.3 Sense of Presence

From the analysis of the questionnaire, it emerged that participants in the VR condition reported an appropriate sense of presence in response to both questions ("*I felt completely immersed*": M = 5.48; SD = 1.43; "*I felt like I was inside the room*": M = 5.52; SD = 1.71). A single sample t-test found that the means of both questions significantly differed from the central value of 4 (first question: $t(30) = 5.76$, $p < .001$; second question: $t(30) = 4.94$, $p < .001$).

4. Discussion

The present study aimed to examine how memory and suggestibility are affected by the media in which stimuli are presented. In more detail, three experimental groups were asked to memorize as many objects contained in a target room, shown respectively in real life, through a Meta Quest 2 HMD, and on a 2D desktop computer. Memory was assessed using free recall and visual recognition tasks, while suggestibility via verbal and visual tasks.

Compared to 2D, participants in the real-life condition remembered significantly more details during free recall, made fewer errors in visual recognition and in both the non-suggestive verbal and visual tasks, and were more resistant to suggestive verbal and visual questions. These results highlight that viewing the stimuli in real life or in 2D might yield different performances in both memory and suggestibility tasks.

Similarly, compared to VR, participants in the real-life condition remembered significantly more details during free recall and made fewer errors in visual recognition, hinting at the possibility that the memory performance in these two conditions is not comparable. Conversely, in relation to suggestibility, the performance between real life and VR did not significantly differ, except for the suggestive visual questions to which participants in real life were more resistant. While the impact of VR on suggestibility is still an under researched topic, these results indicate that the ability to resist to suggestive questions might be somewhat similar in these conditions.

Additionally, VR participants obtained memory and suggestibility performances similar to those in the 2D condition, with the exception of making significantly fewer errors when answering non-suggestive verbal questions.

These results make a valuable contribution to the literature, emphasizing that users' performance in real life is not necessarily comparable to their performance in a VR setting, and that for a VR environment to elicit a lifelike response, it must do more than merely provoke a strong sense of presence (e.g., Mania & Chalmers, 2001). The results are also partially aligned with previous studies finding no differences in performance between VR and 2D settings (e.g., Ernstsen et al., 2019; Makransky et al., 2019; Kisker et al., 2021b).

At least two hypotheses could be formulated to explain why participants in the VR condition showed an overall worse performance on memory tasks than participants in the real-life condition and similar performance to participants in the 2D condition. First, the VR stimuli used in the present study was a panorama 360° picture, rather than a computer-generated scenario; for this reason, although participants reported an appreciable sense of presence, other VR elements were missing, including interactivity and multisensoriality. Therefore, it may be the case that, for VR to be effective, it must leverage all of its unique advantages (i.e., embodiment and agency), as proposed by Johnson-Glenberg et al. (2021). However, the aim of the present study was to assess performance differences in memory tasks related to media presentation, net any other variable; thus, the stimuli presented in the three conditions were kept as similar as possible. Furthermore, the real-life condition also lacked interactivity, as the participants were not allowed to freely explore the room.

A second hypothesis could explain the lack of performance differences on most tasks between the VR and 2D picture conditions: the literature indicates that memory performance improves when participants recall information in the same context in which the information was originally presented (e.g., Godden & Baddeley, 1975). Therefore, considering that all participants completed the memory tasks on a computer, the feature similarity between the context in which participants in the 2D picture condition memorized the stimuli and carried out the memory tasks might have increased their performance to a level that was similar to that of participants in the VR condition. However, it should be also noted that participants in both the VR and real-life conditions performed the memory tasks in a different environment than the one in which the information was learned.

There are three main limitations of the present study. First, considering that more than two-thirds of the sample had no prior experience with VR, more time could have been spent on the participants' training phase with the Meta Quest 2, in order to help participants become accustomed to the virtual environment. Second, as already mentioned, some advantages of VR (e.g., interactivity) were not leveraged, and this may have decreased the performance of participants in the VR condition. Finally, the third limit concerns the experimental stimulus, which consisted of a

single item (i.e., the target room) shown in the three experimental conditions, thus the results should be interpreted with caution.

In conclusion, future research employing experimental paradigms in a VR environment should be careful in assuming that performance in a VR setting is comparable to performance in real life, and that VR environments are more ecological than traditional 2D media. Future research should also investigate the role of media characteristics on suggestibility. More research is needed to guide researchers in building VR environments with the aim of simulating real settings and measuring performance with high ecological validity. For example, future studies should investigate whether interactivity and multisensoriality are essential for VR environments to facilitate cognitive performance at real-life levels.

Chapter 2

Virtual Reality and Stimuli Emotional Features

Ceccato, I., Ricci, E., Mazza, C. et al. Influence of stimuli emotional features and typicality on memory performance: insights from a virtual reality context. *Psychological Research* 88, 257–270 (2024). https://doi.org/10.1007/s00426-023-01850-8

1. Introduction

1.1 Research Comparing Memory Performance Between Real-life, VR, and 2D Contexts

Memory describes the capacity to successfully encode, store, and retrieve the information that has been previously experienced (Zlotnik & Vansintjan, 2019). Memory is not a monolithic construct, but it includes different systems, such as sensory, short-term, and long-term memory. A further classification within the long-term memory system includes declarative (semantic, episodic, and autobiographical) and procedural memory (the reader interested in a detailed description of memory dimensions and mechanisms can refer to the extensive work of Baddeley et al. 2020). In the current study we focused on episodic memory and information retrieval. Episodic memory can be defined as the memory for specific events that occurred at a particular time and place (Anderson, 2020). Information retrieval is performed through the processes of recall and recognition. Recall refers to the ability to remember—and thereby reproduce—previously learned information. The simplest form of recall is free recall, which describes the process of recalling as many stimuli as possible, in no precise order, with no external cues. Recognition, on the other hand, refers to the ability to recognize that a particular stimulus or event (often presented within a set of distractors) has been previously experienced (Baddeley, 2004). Recognition performance, which is commonly measured in reference to signal detection theory (SDT; Green & Swets, 1966), is associated with recognition confidence (i.e., confidence in one's ability to recognize stimuli). Consistently, recognition memory is based on two fundamental and different components: recollection and familiarity (see, for a review, Aggleton & Brown, 2006; Diana et al., 2007; Yonelinas, 2002). Recollection describes the retrieval of a specific, previously experienced stimulus, while familiarity represents

a more global measurement of memory strength and confidence with respect to a specific event or stimulus (Yonelinas et al., 2010).

As demonstrated in the voluminous literature, memory performance, or the ability to correctly recall or recognize previously encoded information, is also intrinsically context-related (Smith, 1979; Smith & Vela, 2001), as retrieval is enhanced when the information is recalled in the same context as that in which it was encoded (see, e.g., Godden & Baddeley, 1975). However, it remains controversial whether memory performance in virtual reality (VR) environments is similar to memory performance in real life—and, if so, to what extent (Smith, 2019; van Helvoort et al., 2020). For instance, a recent study compared physiological features related to memory retrieval in real-life, VR, and conventional experimental conditions. Participants were shown the same event of a car driving in 2D video format, in 3D VR video, or as a physical enactment. No significant differences were detected in the recognition memory task between the real-life and VR conditions, but significant differences emerged between these conditions and the 2D video condition (Schöne et al., 2017), suggesting that VR and real-life contexts may be somewhat comparable and both better than 2D condition. Conversely, a study by Mania and Chalmers (2001) investigating memory recall in four experimental learning conditions (i.e., in-person, on a 3D desktop, with a 3D head-mounted display [HMD], via audio) found significantly superior recall in the in-person condition, compared to the VR condition. Likewise, Flannery and Walles (2003) compared item recognition for objects between participants in a virtual versus a real-life office. Scores were significantly greater for the real-life condition, suggesting non-comparable performance between real and virtual learning.

A parallel body of research has explored differences in memory performance between 2D and VR learning environments. Krokos et al. (2019) investigated memory in participants using virtual memory palaces (i.e., method of loci) presented through an HMD and participants using a traditional 2D desktop display. Participants in the HMD condition demonstrated higher memory recall compared to participants in the 2D desktop condition. However, other research comparing VR and 2D environments has produced contrasting results. For instance, Barreda-Ángeles et al. (2021) conducted a within-subjects study in which participants were

asked to watch eight journalistic pieces on a 360°-video and on a 2D screen. After the video presentation, participants completed a recognition task and, a week later, free recall and cued recall tests. The authors found that the VR and immersive condition elicited lower recognition and cued recall of information, while no significant differences were found in the free recall task. Similarly, a study by Kisker et al. (2021b) exposed participants to a video in a PC condition and a VR condition, followed by an unannounced recognition memory test. The results showed that both groups performed equally well on the unannounced recognition memory test.

1.2 The Impact of Stimuli's Emotional Valence, Emotional Arousal, and Typicality on Items Retrieval

The impact of coding context on memory performance may be partially influenced by the specific features of the memory stimulus, including its emotional valence and typicality. In fact, research has shown that stimuli emotional features (i.e., valence, arousal) have a significant impact on memory (for a review, see Kensinger & Schacter, 2016). Emotional valence concerns an individual's definition of a stimulus as attractive or aversive (i.e., positive or negative, respectively). Several studies conducted in laboratory settings have shown that both positive and negative emotional stimuli are remembered better and more vividly than non-emotional stimuli (i.e., neutral stimuli; Kensinger, 2004; Kensinger & Corkin, 2004; Palumbo et al., 2023). Concerning arousal, which reflects an individual's level of energy in response to a stimulus, research has shown that emotionally arousing items carry greater weight in cognitive processes and are better recalled than neutral stimuli (e.g., Kensinger & Corkin, 2003).

To date, few studies have evaluated differences in memory performance related to coding context (i.e., real-life, 2D, VR), even examining the influence of stimuli emotional valence and arousal. For example, it has been demonstrated that VR movies produce greater emotional arousal than 2D movies (Tian & Zhang, 2021). With respect to typicality, previous studies have shown that stimuli that are easily distinguished from (and thus not in alignment with) their context are significantly better recalled (Bylinskii et al., 2015; Vogt & Magnussen, 2007; Watier & Collin,

2012). Mourkoussis et al. (2010) created two versions of a virtual academic office: one with highly detailed objects and another with only plain objects. Furthermore, items placed in the virtual office were either consistent (e.g., books, chair) or inconsistent with the environment (e.g., snake, dart board). Participants explored the virtual office using an HMD and were subsequently administered an old/new item recognition test. The results revealed no main effect concerning visual detail, but significantly better memory for the inconsistent items in the highly detailed condition, and no influence of visual detail on memory for the consistent items.

1.3 Premises, Aims, and Hypotheses of the Present Research

Overall, the inconsistency encountered in the literature about memory performance is likely due to the wide variety of devices classified as VR. For example, Smith has illustrated three main subtypes of VR systems: Desktop-VR, Headset-VR (i.e., Head-Mounted Display, HMD), and Simulator-VR, each with different technical features (for a detailed description, see the review of Smith, 2019). Consequently, mixed results may depend on VR systems' properties, such as visual fidelity and visual detail of the Virtual Environment (VE), and the elicited sense of presence (i.e., the feeling of "being there" subjects experience in VR) and immersiveness, which is determined by objective characteristics of the VR system. Also differentiating the studies is the methodology employed, which may include: (a) active vs. passive viewing; (b) interaction with the VE; (c) multimodal sensory stimuli. Furthermore, recent studies moved the focus of the study on memory processes, rather than performance, exploring electrophysiological correlates in the VR setting (Johnsdorf et al., 2023). Hence, a plain comparison between studies is difficult.

Given the contradictory results reported in the literature, the first aim of the present study was to gain insight into the impact of coding context on memory performance. Specifically, three groups of young adults were exposed to the same environment (i.e., an academic office) in three different coding contexts: (a) a real office; (b) a panorama 360° photo of the office loaded inside a Head-Mounted Display (HMD, i.e., Meta Quest 2); and (c) 2D pictures of the office. In doing that, we considered different facets of memory performance and investigated both free recall and

recognition tasks. Although the literature is inconsistent, we hypothesized that the memory performance of participants exposed to a real-life context would be comparable to that of the subjects in VR; conversely, the 2D condition would lead to worse performance. Furthermore, as recognition is generally easier than recall with no external cues (Baddeley et al., 2020; Rhodes et al., 2019), we hypothesized that potential differences between coding contexts would be magnified in the free recall task compared to the recognition task.

The second aim of the study was to examine the impact of stimuli features (emotional valence, emotional arousal, and typicality) on item retrieval, comparing within the same study three experimental coding conditions (i.e., real-life, 2D, and HMD-VR). Finally, we tested recognition confidence to evaluate whether the coding contexts may differently affect the subjective assessment of memory performance, possibly revealing a discrepancy with the objective performance. To the best of our knowledge, few studies (e.g., Mania et al., 2003) have studied recognition confidence related to VR environments. Findings showed that participants in the real-life condition achieved the most accurate and confident recognition performance, followed by participants in the VR condition, and, lastly, participants in the 2D condition, who showed the lowest accuracy and confidence recognition.

2. Materials & Method

2.1 Participants

The participant sample in this study is the same as that described in Chapter 1. All individuals were recruited from the same pool and completed the experiments as part of a broader research project. One hundred and twenty-three young adults were recruited for and voluntarily participated in the study. The inclusion criteria were: (i) aged 18 years and older and (ii) able to comprehend the Italian language perfectly. Four participants (3.25%) were excluded due to an invalid procedure, caused by problems associated with the VR equipment (i.e., HMD-VR Meta Quest 2). Thus, the final sample was comprised of 119 participants (see Table 4): 62 male (52.1%) and 57 female (47.9%), aged 18–35 years ($M = 24.20$, $SD = 4.13$). Most participants were Italian citizens ($n = 118$, 99.2%) and living in central Italy

($n = 107$, 89.9%). They were predominantly students ($n = 77$, 64.7%) whose highest educational achievement was high school ($n = 67$, 56.3%). The majority had no visual impairment (51.4%). Sensitivity analyses were conducted using G*power (version 3.1.9.7), to compute the minimum effect size that could be detected given the alpha, power, and sample size (Perugini et al., 2018). The results for a one-way ANCOVA involving three groups with alpha = 0.05, power = 0.80, and 119 participants indicated sufficient power to detect a medium effect size, partial $\eta_2 \geq 0.08$. The results for a repeated measure within-between interaction ANOVA involving three groups and two measures as within-subject factors, given alpha = 0.05, power = 0.80, and 119 participants, indicated sufficient power to detect a medium effect size, partial $\eta_2 \geq 0.02$.

Table 4

Descriptive Statistics of the Sample

	Real life	**HMD-VR**	**2D pictures**	**Total**
	n(%)	*n*(%)	*n*(%)	*N*(%)
Biological sex				
Female	22 (55.0)	18 (45.0)	17 (43.6)	57 (47.9)
Male	18 (45.0)	22 (55.0)	22 (56.4)	62 (52.1)
Citizenship				
Italian	39 (97.5)	40 (100)	39 (100)	118 (99.2)
Other	1 (2.5)	0 (0)	0 (0)	1 (0.8)
Region				
North	0 (0)	2 (5)	1 (2.6)	3 (2.5)
Central	36 (90)	36 (90)	35 (89.7)	107 (89.9)
South	4 (10)	2 (5)	3 (7.7)	9 (7.6)
Educational level				
≤ High school diploma	12 (30)	27 (67.5)	29 (74.4)	68 (57.1)
≥ Bachelor's degree	28 (70)	13 (32.5)	10 (25.6)	51 (42.9)
Occupation				
Unemployed	1 (2.5)	1 (2.5)	2 (5.1)	4 (3.4)
Student	17 (42.5)	30 (75)	30(76.9)	77 (64.7)
Employed	22 (55)	9 (22.5)	7 (18)	38 (31.9)

Vision impairment				
No	19 (47.5)	20 (50)	22 (56.4)	61 (51.3)
Yes	21 (52.5)	20 (50)	17 (43.6)	58 (48.7)
Use of glasses/contact lenses				
No	23 (57.5)	21 (52.5)	22 (56.4)	66 (55.5)
Yes	17 (42.5)	19 (47.5)	17 (43.6)	53 (44.5)
Prior experience of VR				
No	30 (75)	26 (65)	26 (66.7)	82 (68.9)
Yes	10 (25)	14 (35)	13 (33.3)	37 (31.1)

Participants were randomly allocated to one of three experimental conditions according to a manipulated variable (i.e., the coding environment; see the "Experimental procedure" section for a detailed description):

- Real-life (Mage = 26.53, SD = 3.60): group 1 was composed of 40 participants who visited the academic office in vivo.
- HMD-VR (Mage = 23.25, SD = 4.19): group 2 was composed of 40 participants who visited the academic office in 3D VR, using a Meta Quest 2.
- 2D pictures (Mage = 22.79, SD = 3.60): group 3 was composed of 39 participants who observed 2D pictures of the academic office on a computer.

2.2 Measures

A socio-demographic questionnaire was administered to all participants to gather personal information about biological sex, age, education, occupational status, region of residence, citizenship, and medical diagnoses. Data concerning visual impairment and prior experience of VR were also retrieved.

The experimental memory tasks were the following:

2.2.1 Free recall task

Participants were given 4 min to write down all of the objects in the room they visually explored. Globally, 53 objects were identified across all groups. The free recall total score was calculated as a ratio of the number of objects recalled to the total number of recallable objects (i.e., $N = 53$). In a subsequent phase, a rating study

was conducted online on an independent sample of young adults ($N = 138$)—not undergoing the memory assessment—to evaluate each object according to three dimensions: valence (i.e., from negative/low to positive/high), arousal (i.e., from low to high), and typicality (i.e., commonality of the object in an office setting; from low to high). More details on the rating study are provided in the Supplementary Information (Table S1).

Based on the average ratings, each object was categorized as either low or high on each dimension. Of note, the three dimensions were treated independently. For example, a dirty water bottle was classified as high on the arousal dimension but low on the valence and typicality dimensions. Subsequently, six free recall subscores were calculated: two (i.e., low and high) for each dimension (i.e., valence, arousal, typicality). Subscores were calculated as a ratio of the number of objects recalled to the total number of objects in a given category. For example, the high typicality category comprised 34 objects (out of 53). If a participant recalled eight objects in this category, they obtained a high typicality subscore of 0.23. The analyses of these three dimensions provided insight into the impact of coding context (i.e., real-life, HMD-VR, 2D pictures) on memory performance, considering stimulus emotional valence and typicality.

2.2.2 Recognition task

Participants were shown 20 pictures of objects on a 27″ personal computer, with no time limit. Ten of these pictures depicted objects that were present in the room the participants visually explored (i.e., "old items"). The remaining 10 pictures depicted objects that were not present in the room (i.e., "new items"). Figure 1 presents two pictures included in the visual recognition task. Note that new items were not always perfectly paired with old items. Participants saw one picture at a time and were instructed to indicate if the object depicted in the picture was present in the room they visually explored, by selecting one of five options: 1 (surely I did not see it in the room), 2 (maybe I did not see it in the room), 3 (I do not know if I saw it in the room), 4 (maybe I saw it in the room), and 5 (surely I saw it in the room). This response scale was designed to analyze both accuracy (i.e., according to SDT; Stanislaw & Todorov, 1999) and recognition confidence.

In line with the SDT framework, recognition performance was explored through the calculation of several variables: First, the number of hits (i.e., number of correctly recognized old items) and the complementary number of misses (i.e., number of unrecognized old items); for example, out of 10 old items, a participant with a hit score of 7 had a miss score of 3. Second, the number of false alarms (i.e., number of misrecognized new items) and the complementary number of correct rejections (i.e., number of correctly identified new items). Third, a discriminability parameter (d′), reflecting the extent to which two "things" (conventionally labelled "signal" and "noise") could be distinguished. In the recognition task, old items were the "signal" to be detected, while new items were the "noise" to be ignored. The discriminability parameter d′ was thus calculated as the difference between the z-transformed scores of the distributions of hits and false alarms, ranging from 0 (i.e., no discrimination) to infinity (i.e., perfect discrimination). And fourth, a decision bias parameter I, indexing the propensity to label an object as old (or new). The C parameter was calculated as the sum of the z-transformed scores of the distributions of hits and false alarms, divided by 2 and multiplied by − 1. $C = 0$ reflected a "neutral" response tendency, with no decision bias towards either response type; $C < 0$ indicated a bias towards "old" responses (resulting in more hits, as well as more false alarms); and $C > 0$ indicated a preference for "new" responses (resulting in fewer hits and fewer false alarms). For old items, responses 4 and 5 were coded as correct recognition (i.e., hits), while responses 1, 2, and 3 were coded as misses. For new items, responses 1 and 2 were coded as correct rejection, while responses 3, 4, and 5 were coded as false alarms. Subsequently, recognition confidence for the correct identification as old (i.e., hit) or new (i.e., correct rejection), as well as for false alarms and misses, was calculated by recoding each answer, as described in Table 5. Average scores were then computed. Hit confidence and correct rejection confidence ranged from 0 to 2, while false alarm confidence and miss confidence ranged from 1 to 2.

Table 5

Coding Criteria for Confidence in the Recognition Task

	Response				

		1 **Surely new**	2 **Maybe new**	3 **Do not know**	4 **Maybe old**	5 **Surely old**
Item	Old	Miss confidence = 2	Miss confidence = 1	Hit confidence = 0	Hit confidence = 1	Hit confidence = 2
	New	Correct rejection confidence = 2	Correct rejection confidence = 1	Correct rejection confidence = 0	False alarm confidence = 1	False alarm confidence = 2

2.2.3 Sense of Presence in the VR Environment

Two individual items were employed to measure the sense of presence, with the goal of ensuring that the VR environment had elicited the experience of "being there" in participants. Specifically, two questions were asked: the first (i.e., "*I felt completely immersed*") was adapted from Jennett et al.'s (2008) scale and assessed on a 7-point Likert scale ranging from 1 (*Not at All*) to 7 (*Very Strongly*), as also applied in previous research (e.g., Hudson et al., 2019). The second question (i.e., "*I felt like I was inside the room*") was adapted from Wagler and Hanus's (2018) scale of spatial presence and was assessed on a 7-point Likert scale ranging from 1 (*Not at All*) to 7 (*A Lot*).

Furthermore, to control for baseline memory performance the following measures were administered:

2.2.4 Rivermead Behavioural Memory Test

The Rivermead Behavioural Memory Test (RBMT-III; Wilson et al., 1985; Italian validation: Beschin et al., 2013) is an ecological instrument that evaluates respondents' capacity to use memory in ordinary circumstances. The RBMT is composed of 14 subtests, which assess visual memory, verbal memory, and recall memory, both immediate and delayed. In the present study, only the Figure Recognition subtest was used to assess participants' ability to recall previously displayed images from a larger set. This subtest consists of 15 drawings of ordinary objects that are shown sequentially, each for 3 s. In part 1 of the subtest, respondents are asked to name the object represented in the 15 drawings; in part 2, measuring deferred recognition, they are instructed to identify the original objects from a group of 30 drawings (i.e., a set that includes 15 distractor drawings). If respondents make a mistake in their naming of an object, the examiner corrects them and notes the potential presence of language or perceptual problems. The RBMT has been shown

to have good ecological validity and to successfully predict real-life behavior and deficits outside the evaluation situation. The present study considered the number of images correctly recognized (ranging from 0 to 15).

2.2.5 Corsi Block-Tapping Test

The Corsi block-tapping test (CBT; De Renzi & Nichelli, 1975; Spinnler, 1987) measures the quantity of information that can be held in visuo-spatial memory (i.e., short-term memory). The stimulus consists of a wood board (32 × 25 cm) upon which nine progressively numbered cubes (4.5 × 4.5 × 4.5 cm) are asymmetrically placed. Each cube displays a number, which faces the examiner (who sits across from the participant). The examiner demonstrates a sequence of numbers by tapping on the cubes at a rate of 2 s per cube. Subsequently, the respondent must replicate the sequence by touching the cubes in the same order. This process is repeated, with the sequence length increasing from a minimum of 3 to a maximum of 10. For every sequence length, there are two sequences performed. When the participant correctly reproduces a sequence, the examiner moves on to the next longest sequence length. The participant's visuo-spatial memory span corresponds to the longest sequence of numbers reproduced correctly.

2.3 Experimental Procedure

The present study comprised part of a wider research project. The experimental procedure was conducted in the Department of Human Neuroscience laboratories at the "Sapienza" University of Rome, in October 2021. It was divided into three stages, each lasting approximately 30 min. The study was approved by the local ethics committee (Board of the Department of Human Neuroscience, Faculty of Medicine and Dentistry, Sapienza University of Rome—n° 15/2021), according to the Declaration of Helsinki.

2.3.1 Stage 1

After arriving at the laboratory, participants provided written informed consent, filled out the socio-demographic questionnaire, and completed the RBMT-III

Figure Recognition subtest and Corsi block-tapping text (see the "Measures" section).

2.3.2 Stage 2

Participants were randomly allocated to one of three experimental conditions. All groups were asked to observe an academic office and memorize as many objects as possible over a period of 2 min. However, the three experimental conditions were differentiated according to the coding environment in which the procedure took place (i.e., real-life, HMD-VR, 2D pictures).

- Real-life group: Participants were positioned in the middle of the academic office—a room within the university—and instructed to remain stationary and only rotate to observe the room and the objects within it. Specifically, the instructions were as follows: "*We will now take you into a room. You must observe the room carefully for 2 min, memorizing as many objects as possible. You must remain in the spot in which we position you. However, you can turn your body in all directions*".
- HMD-VR group: Participants were taken to a neutral room and instructed to put on a Head-Mounted Display (HMD; i.e., Meta Quest 2), which displayed a panorama 360° picture of the academic office, taken by a professional photographer with a Lapbano Pilot One EE. The 360° picture was taken from the same point where participants in the Real-life group were standing. Participants were instructed on the use of the Meta Quest 2 (e.g., not moving outside the planned area, physically turning on themselves to see all parts of the room) and asked to observe the academic office and memorize as many details as possible over a period of 2 min. Specifically, the instructions were as follows: "*Now you are going to participate in a virtual reality experience, wearing this visor. You will enter a room and explore it in 360 degrees. Observe the room carefully for 2 min, memorizing as many objects as possible. You must remain still, with the exception that you can turn your head and rotate your body in all directions*". Following this step, participants' sense of presence was assessed through two questions (see the "Sense of presence" paragraph).

- 2D pictures group: Participants were taken to a neutral room, where they were seated in front of a 27″ computer monitor upon which eight 2D pictures (2048 × 1537-pixel, 3 megapixels) displayed the same academic office from the same point of view of participants in the Real-life and HMD-VR conditions—in Figure 5 is reported one of the pictures used in the 2D condition (see Supplementary Information, Figure S1, for the remaining). Participants could scroll back and forward across the pictures as they wanted, and they were told that they had 2 min to look at and memorize as many objects as possible in the academic office. Specifically, the instructions were as follows: "*Now we will show you some pictures of a room. Observe the room carefully for 2 min, memorizing as many objects as possible*."

We loaded a panorama 360° photo into HMD Meta Quest 2 to avoid using a computer-generated scenario, representing an academic office in an artificial way, to make the VR stimulus as similar to reality as possible. Additionally, in the 2D condition was not shown the panorama 360° picture used in the HMD-VR condition as its resolution did not allow to see clearly all the objects in the office. Indeed, this picture was developed for a VR device and thus was unadaptable for a computer monitor.

2.3.3 Stage 3

Participants in all groups returned to the original reception room, where they completed the two experimental tasks, starting with the free recall task. Subsequently, they completed the recognition task (see the "Measures" section) on a 27″ personal computer, with no time limit.

2.4 Statistical Analyses

One-way ANOVAs and chi-square tests, depending on the nature of the variable, were run to test group differences in socio-demographic features and general memory performance (i.e., scores on the RBMT and CBT). A single sample t-test was run to test the sense of presence. Zero-order correlations were computed between socio-demographic variables, general memory performance, free recall,

and recognition. A univariate ANCOVA using group (i.e., real-life, HMD-VR, 2D pictures) as an independent factor was performed on the free recall total score. Subsequently, three repeated-measure ANCOVAs on the three recall subscores (i.e., valence, arousal, typicality) were run, entering the low versus high level of each dimension as a within-subject factor and group as a between-subject factor. Finally, univariate ANCOVAs were computed considering group (i.e., real-life, HMD-VR, 2D pictures) as an independent factor and SDT recognition parameters and recognition confidence (i.e., number of hits, number of false alarms, d′, C, hit confidence, correct rejection confidence, miss confidence, false alarm confidence) as dependent variables. Age and education were entered as covariates. All post hoc analyses were Bonferroni adjusted. All analyses were performed using SPSS version 27 software.

3. Results

3.1 Preliminary Analyses

Analyses were run to determine whether the three experimental groups were comparable in terms of socio-demographic features and memory performance (see Table 6). No differences in gender distribution or general memory performance emerged between groups. Scores on the RBMT Figure Recognition subtest and CBT for the three groups corresponded to general population averages (Beschin et al., 2013; Spinnler, 1987). However, significant differences in age and education were found, with participants in the real-life group older and more educated than participants in the other conditions. Hence, subsequent analyses controlled for age and education. Table 7 presents the zero-order correlations between all of the considered variables.

Table 6

Descriptive Information for the Three Experimental Groups and Results of the Group Comparison

	Real-life (n = 40)	**HMD-VR (n = 40)**	**2D pictures (n = 39)**	**Group comparison**	**Post hoc**
Age	26.53 (3.60)	23.25 (4.19)	22.79 (3.60)	***F*(2,116) = 11.34 *p* < 0.001**	Real-life > HMD-VR, 2D pictures

				Partial $\eta^2 = 0.16$	HMD-VR = 2D pictures
Gender	M = 55% F = 45%	M = 45% F = 55%	M = 44% F = 56%	$\chi^2(2) = 1.23$ $p = 0.540$	–
Education	3.78 (0.66)	3.33 (0.47)	3.36 (0.67)	**$F(2,116) = 6.80$** **$p = 0.002$** **Partial $\eta^2 = 0.10$**	Real-life > HMD-VR, 2D pictures HMD-VR = 2D pictures
RBMT	14.73 (0.64)	14.60 (0.67)	14.74 (0.55)	$F(2,116) = 0.62$ $p = 0.537$ Partial $\eta^2 = 0.01$	–
CBT	5.98 (1.35)	6.28 (1.47)	5.92 (1.42)	$F(2,116) = 0.72$ $p = 0.490$ Partial $\eta^2 = 0.01$	–

Note. RBMT Rivermead Behavioural Memory Test—Figure Recognition subtest, CBT Corsi Block-Tapping test of visuo-spatial memory span. Means (and standard deviations) are reported for all continuous variables. For gender, frequencies (in %) are reported. Post hoc analyses were Bonferroni adjusted. Statistically significant effects ($p < 0.05$) are in bold.

Table 7

Correlations between Socio-Demographic Characteristics, Memory Performances, Free Recall, and Recognition

	Gender	**Education**	**RBMT**	**CBT**	**FRT**	**Hit**	**False alarm**	***d'***	**C**	**Hit Conf**	**CR Conf**	**FA Conf**	**Miss Conf**
Age	− 0.01	0.58***	0.10	− 0.14	0.38***	0.27**	− 0.10	0.29**	− 0.16$^+$	0.26**	0.17$^+$	0.21*	0.08
Gender		− 0.17$^+$	− 0.07	0.10	− 0.24**	− 0.01	0.03	− 0.01	− 0.03	0.01	− 0.05	− 0.21*	0.01
Education			0.09	− 0.09	0.42***	0.23*	− 0.01	0.19*	− 0.17$^+$	0.15	0.16$^+$	0.15	0.03
RBMT				0.04	0.13	0.05	− 0.04	0.07	− 0.01	0.02	0.13	0.17$^+$	0.12
CBT					− 0.01	0.04	− 0.07	0.07	− 0.01	− 0.11	− 0.13	− 0.09	− 0.10
FRT						0.39***	− 0.01	0.33***	− 0.29**	0.38***	0.25**	0.42***	0.19*
Hit							− 0.05	0.77***	− 0.79***	0.55***	0.13	0.22*	0.13
False alarm								− 0.59***	− 0.64***	0.05	0.05	− 0.19*	− 0.11
d'									− 0.24**	0.40***	0.09	0.32***	0.17$^+$
C										− 0.45***	− 0.15	− 0.06	− 0.05
Hit Conf											0.40***	0.52***	0.30***
CR Conf												0.24**	0.71***

FA Conf		0.21*

Note. RBMT Rivermead Behavioural Memory Test, CBT Corsi block-tapping test of visuo-spatial memory span, FRT free recall total score, Hit Conf hit confidence, CR conf correct rejection confidence, FA conf false alarm confidence, Miss conf miss confidence Pearson's r is reported for all variables but gender. Point-biserial correlations are reported for the associations between gender and other variables

+$p < 0.10$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$

3.1.2 Sense of Presence

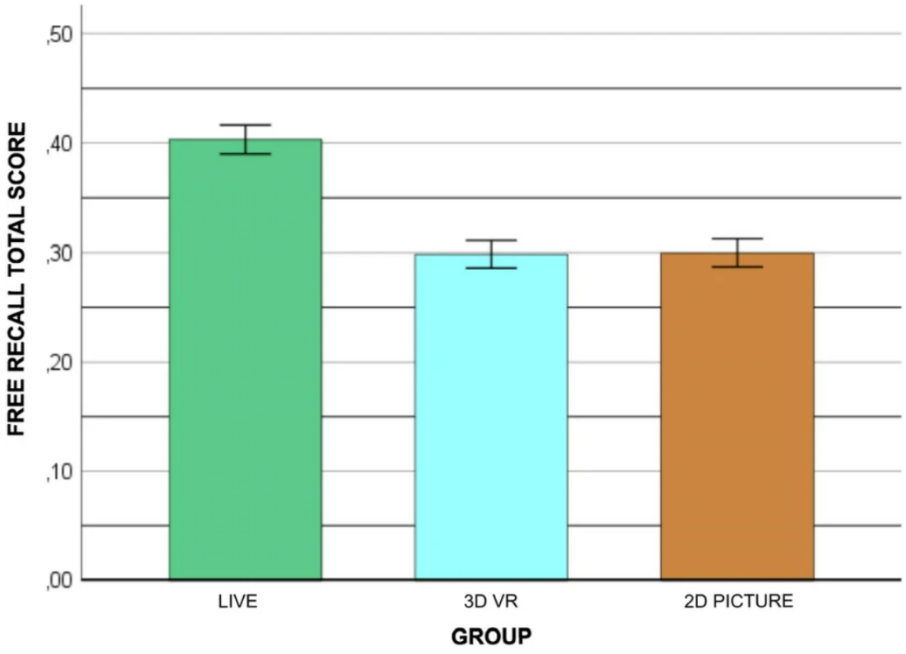
Results revealed that participants in the HMD-VR condition experienced an appropriate sense of presence in response to both questions ("*I felt completely immersed*": M = 5.48; SD = 1.43; "*I felt like I was inside the room*": M = 5.52; SD = 1.71). A single sample t-test showed that the average score of both questions significantly differed from the central value of 4 (first question: $t(30) = 5.76$, $p < .001$; second question: $t(30) = 4.94$, $p < .001$).

3.2 Group Differences in Free Recall

A univariate ANCOVA with group as the independent factor (three levels: real-life, HMD-VR, 2D pictures) was performed on the free recall total score. Age and education were entered as covariates. The results revealed a main effect of group, $F(2,114) = 19.44$, $p < .001$, partial $\eta_2 = .25$. Post hoc analyses showed that the real-life group outperformed both the HMD-VR and the 2D picture groups (Figure 6). Education was a significant covariate, $F(1,114) = 7.44$, $p = .007$, partial $\eta_2 = .06$, but not age, $F(1,114) = 0.16$, $p = .688$, partial $\eta_2 = .001$.

Figure 6

Free Recall Performance in Each Experimental Group



Note. Error bars represent ± 1 SE. Covariates were evaluated at the following values: age = 24.20; education = 3.49.

3.3 Effects of Valence, Arousal, and Typicality on Free Recall

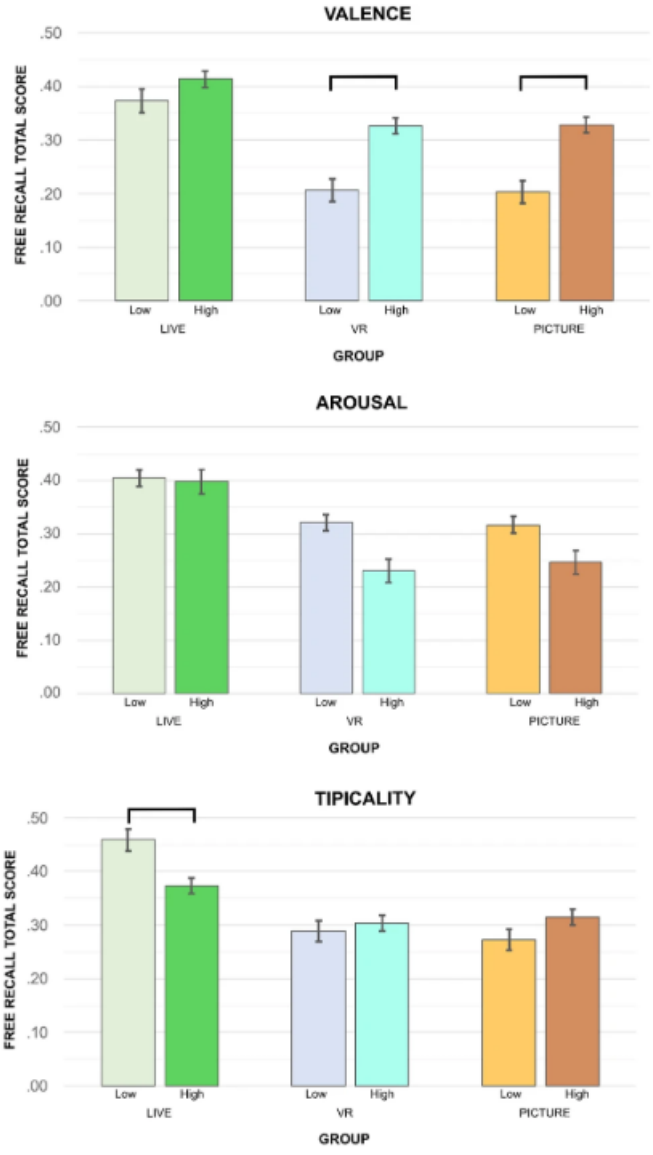
Analyses aimed at examining the role of object characteristics on recall, in terms of the three dimensions of valence, arousal, and typicality. Three repeated measure ANCOVAs were run on recall scores, entering dimension level (i.e., low, high) as a within-subject factor and group (i.e., real-life, HMD-VR, 2D pictures) as a between-subject factor.

For valence, the main effect of the experimental group was significant, $F(2,114) = 26.56$, $p < .001$, partial $\eta_2 = .32$, with participants in the real-life group outperforming the other two experimental groups. The main effect of level was not significant, $F(1,114) = 2.49$, $p = .117$, partial $\eta_2 = .02$. Crucially, a significant interaction effect was found between group and level, $F(2,114) = 3.09$, $p = .049$, partial $\eta_2 = .05$. Post hoc analyses (Bonferroni adjusted) indicated that, in the real-life condition, there were no differences in the recall of low valence (M = 0.39, SD = 0.17) versus high valence objects (M = 0.42, SD = 0.10, $p = .131$). In contrast, in both the HMD-VR and the 2D picture conditions, recall was higher for objects with high valence (HMD-VR: M = 0.32, SD = 0.09; 2D pictures: M = 0.32, SD = 0.09) than for objects with low valence (HMD-VR: M = 0.20, SD = 0.10; 2D

pictures: M = 0.19, SD = 0.11), $p < .001$ (Figure 7). Education was a significant covariate, $F(1,114) = 3.98$, $p = .048$, partial $\eta_2 = .03$, while age was not, $F(1,114) = 1.59$, $p = .210$, partial $\eta_2 = .01$.

Figure 7

Free Recall Performance (Valence, Arousal, and Typicality Dimensions) as a Function of Group (Real-life, HMD-VR, 2D Pictures) and Level (Low, High)



Note. Error bars represent ± 1 SE. Squared brackets indicate statistically significant differences ($p < 0.001$). Covariates were evaluated at the following values: age = 24.20; education = 3.49.

For arousal, only the main effect of the experimental group emerged, $F(2,114) = 23.59$, $p < .001$, partial $\eta_2 = .29$, indicating the superiority of the real-life condition in recall performance. Level had no main, $F(1,114) = 2.10$, $p = .150$, partial $\eta_2 = 0.02$, or interaction, $F(2,114) = 2.26$, $p = .109$, partial $\eta_2 = .04$, effects on recall. Education was a significant covariate, $F(1,114) = 4.04$, $p = .047$, partial $\eta_2 = 0.03$, while age was not, $F(1,114) = 1.69$, $p = .196$, partial $\eta_2 = 0.01$. For typicality, a significant main effect of the experimental group emerged, $F(2,114) = 23.99$, $p < .001$, partial $\eta_2 = 0.30$, confirming the real-life group's superiority, while level had no main effect on recall, $F(1,114) = 2.46$, $p = .120$, partial $\eta_2 = .02$. Of interest, a significant interaction effect between group and level was found, $F(2,114) = 8.25$, $p < .001$, partial $\eta_2 = 0.13$. Post hoc analyses showed that, in the real-life condition, low typicality objects ($M = 0.48$, $SD = 0.12$) were recalled more frequently than high typicality objects ($M = 0.38$, $SD = 0.10$), $p < .001$. In contrast, in the HMD-VR condition, no difference ($p = 0.513$) between low ($M = 0.28$, $SD = 0.12$) and high typicality ($M = 0.30$, $SD = 0.09$) objects emerged. In the 2D picture condition, the difference between low and high typicality objects approached statistical significance ($p = .053$) (low: $M = 0.26$, $SD = 0.12$; high: $M = 0.31$, $SD = 0.08$), in the direction of a better recall of high typicality objects. Education was a significant covariate, $F(1,114) = 5.65$, $p = .019$, partial $\eta_2 = .05$, while age was not, $F(1,114) = 0.94$, $p = .334$, partial $\eta_2 = 0.01$.

3.4 Group Differences in Recognition

A series of univariate ANCOVAs using the experimental group (i.e., real-life, HMD-VR, 2D pictures) as the independent factor was performed on the four SDT parameters (i.e., number of hits, number of false alarms, d′, C). Age and education were entered as covariates. Table 8 presents the means and standard deviations.

Table 8

Means (and Standard Deviations) for Recognition

	Real life *M* (SD)	**HMD-VR** *M* (SD)	**2D pictures** *M* (SD)
Number of hits	8.30 (1.14)	6.43 (1.95)	6.03 (1.72)
Number of false alarms	1.00 (1.68)	1.08 (1.05)	1.28 (1.10)

d′	2.28 (0.68)	1.67 (0.68)	1.44 (0.57)
C	0.14 (0.37)	0.41 (0.36)	0.44 (0.36)
Hit confidence	1.83 (0.15)	1.53 (0.23)	1.55 (0.26)
Correct rejection confidence	1.68 (0.22)	1.46 (0.31)	1.40 (0.30)
Miss confidence	1.76 (0.27)	1.59 (0.28)	1.55 (0.33)
False alarm confidence	1.92 (0.11)	1.64 (0.26)	1.73 (0.26)

For the number of hits, a main effect of the experimental group emerged, $F(2,114) = 15.63$, $p < .001$, partial $\eta_2 = .22$. Specifically, post hoc analyses showed that the real-life group outperformed both the HMD-VR and the 2D picture groups. In contrast, group had no effect on the number of false alarms, $F(2,114) = 0.27$, $p = .765$, partial $\eta_2 = 0.005$. A significant effect of group was found for d′, $F(2,114) = 12.51$, $p < .001$, partial $\eta_2 = .18$, again showing that participants in the real-life group were more able to discriminate between new and old items than participants in the other two groups. Finally, for the decision bias C, a significant effect of group emerged, $F(2,114) = 6.49$, $p = .002$, partial $\eta_2 = 0.10$. Post hoc analyses indicated that, although all groups showed a tendency to classify objects as new, this bias was significantly higher in the HMD-VR ($p = .009$) and 2D picture groups ($p = .004$), compared to the real-life group. Age and education were not significant covariates in any of the four models described, Fs$(1,114) \leq 1.22$, ps $\geq .272$, partial $\eta_2 \leq .01$. To explore recognition confidence, four ANCOVAs were conducted, one for each confidence score. Group was entered as the independent variable, while age and education were entered as covariates. In all four models, the main effect of the experimental group was significant (i.e., hit confidence: $F(2,114) = 19.23$, $p < .001$, partial $\eta_2 = .25$; correct rejection confidence: $F(2,113) = 8.57$, $p < .001$, partial $\eta_2 = 0.13$; miss confidence: $F(2,112) = 5.15$, $p = .007$, partial $\eta_2 = .08$; false alarm confidence: $F(2,113) = 13.22$, $p < .001$, partial $\eta_2 = .19$). Post hoc analyses showed that confidence was always higher in the real-life group compared to both the HMD-VR and the 2D picture groups, $p \leq .040$. Age and education were not significant covariates in any of the four models, Fs$(1,114) \leq 0.99$, ps $\geq .322$, partial $\eta_2 \leq .01$. For the number of hits and the number of false alarms, scores ranged from 0 to 10. For hit confidence and correct rejection

confidence, scores ranged from 0 to 2. For miss confidence and false alarm confidence, scores ranged from 1 to 2.

4. Discussion

The present study tested the impact of coding context on memory performance, as well as the influence of stimuli emotional valence, emotional arousal, and typicality on recollection. Three groups of participants were exposed to the same coding environment through different modalities: in vivo exposure (i.e., real-life), 3D virtual reality (i.e., Head-Mounted Display VR), and 2D pictures. Memory performance was assessed using both free recall and recognition tasks, while also considering recognition confidence. Overall, participants who performed the experimental task in the real-life encoding context showed higher memory performance in free recall, recognition, and recognition confidence than those who completed the task in the HMD-VR and 2D picture encoding contexts. In addition, memory performance in the HMD-VR and 2D image encoding contexts was comparable. Interestingly, emotional valence influenced the free recall of items in HMD-VR and 2D picture coding contexts, with positive items recalled more often. Finally, typicality influenced the free recall of items only in the real-life coding context: items less common within an academic office were recalled more often. The first aim of the present study was to gain insight into the impact of coding context on memory performance. The findings underlined that free recall and recognition were significantly better for the real-life coding context, with participants remembering and recognizing more objects, on average, compared to participants in the HMD-VR and 2D picture conditions. These results align with the works of Mania and Chalmers (2001) and Flannery and Walles (2003), who found that memory recall and recognition performance, respectively, were significantly higher in an "in-person" coding condition compared to an HMD-VR coding condition. Of interest, in the present study, participants in the HMD-VR coding environment did not show enhanced memory performance compared to participants in the 2D picture coding condition; rather, memory performance between these conditions was similar. Hence, the present results are in line with previous findings showing no significant differences in recognition performance between participants

exposed to 2D and VR coding environments (Barreda-Ángeles et al., 2021; Kisker et al., 2021b).

The second aim of the present research was to explore the influence of emotional valence, arousal, and typicality on memory performance. Interesting effects emerged regarding emotional valence and typicality. Specifically, emotional valence impacted the free recall of objects only in the HMD-VR and 2D picture contexts, whereby positive items were more accurately recalled than negative ones. These findings add to the large body of literature on the impact of emotional stimuli on memory performance (e.g., Fairfield et al., 2017; Kensinger & Schacter, 2016; Mammarella et al., 2016; Penolazzi et al., 2010). Moreover, objects connoted by a positive valence tended to be better recalled by participants in both the HMD-VR and the 2D picture contexts, compared to participants in the real-life context. This suggests that the emotional connotation of stimuli may play a differential role in memory performance, depending on the type of learning environment. However, the influence of stimuli emotional valence on memory performance in VR settings requires further investigation. Indeed, some research suggested that perceived emotional valence may change depending on the coding context, with the same stimulus eliciting even opposite emotions between VR and 2D conditions (Schöne et al., 2021).

Conversely, the present results highlight that typicality was relevant only in the real-life coding context, in which objects that were less commonly found within an office were recalled more often. These findings align with the results of previous research showing that stimuli that stand out from (and are thus not aligned with) their context are significantly better recalled (Bylinskii et al., 2015; Vogt & Magnussen, 2007; Watier & Collin, 2012). However, in the present study, typicality did not influence memory performance in the VR condition. This contradicts the findings of Mourkoussis et al. (2010), which showed better recall for non-typical items in a VR scenario. It could be argued that the degree of visual detail is superior in real-life contexts, relative to VR and 2D picture environments. Thus, the influence of object typicality on memory performance in VR environments requires further exploration.

Finally, the present study examined recognition confidence, finding this to be higher in the real-life condition than the HMD-VR and 2D picture conditions, with respect to both correct (i.e., hits, correct rejections) and incorrect responses (i.e., misses, false alarms). These results align with those of Mania et al. (2003), who compared participants in three different conditions (i.e., a real room, an HMD virtual room, a desktop display of a room) with respect to accuracy and confidence, finding that participants in the real room condition had the most accurate and confident memory performance, whereas participants in the HMD virtual room were slightly less accurate and confident, and participants who were exposed to the desktop display of the room were the least accurate and confident. Overall, it is worth mentioning that all the findings' comparisons in VR research should be interpreted with caution, considering the different technical features of the systems and the methodological choices by researches, as highlighted by Smith (2019) in his review.

Before drawing conclusions, some limitations of the present study must be mentioned. First, to maintain stimuli similarity across the three conditions, the HMD-VR stimulus consisted of a panorama 360° picture of the academic office and not an ad hoc, 3D, VR-generated scenario. Therefore, the type of VR stimulus might have jeopardized memory performance in the HMD-VR coding condition, making it more comparable to the 2D picture coding condition than to the real-life coding condition. Indeed, Johnson-Glenberg et al. (2021) claimed that, for VR to be effective, it should preserve some critical characteristics, such as a feeling of presence and interactivity with the environment. As the present study aimed at investigating differences in memory performance related to the coding context and the influence of stimuli emotional features and typicality on memory performance, no other factors related to the VR environment were specifically examined. However, it should be noted that participants in the HMD-VR coding condition rated an appropriate sense of presence and that in all experimental conditions participants could not interact with the environment. Second, as demonstrated in the literature (Godden & Baddeley, 1975), memory performance is superior when information is recalled in the same context as that of its coding. Thus, as all groups performed the memory recognition task on a computer, it is possible that the

performance of participants in the VR condition was further decreased by the different retrieval context that, instead, was the same coding environment for participants in the 2D picture condition. Lastly, in the VR coding context, participants had to wear a Meta Quest 2 device, with which most of the sample reported no prior experience. Thus, although participants were advised during the training phase to wear the device and adjust it for comfort, the device may have still affected participants' attention during the coding process.

Taken together, the results of the present study highlight that, although VR may contribute to the development of more ecologically valid experimental settings, to date, memory performance in VR settings cannot be considered fully comparable to that of real-life contexts, in which both recall, and recognition are superior. Nevertheless, it is important that future studies take into account also the differences in memory processes when examining memory performance in VR compared to in-vivo and 2D conditions. Indeed, previous studies suggested that VR and 2D conditions differ in the electrophysiological correlates elicited (Kisker et al., 2021a, 2021b). Furthermore, in the present study, an interesting effect of emotional valence emerged in the VR condition, particularly for objects with positive valence. This finding supports previous research underlining the crucial role of emotional valance on memory performance, and it opens up for future research on this topic in the field of VR.

In conclusion, further research is needed to create VR environments that are sufficiently comparable to real-life contexts to obtain reliable measures of cognitive performance. Importantly, such research should also consider critical characteristics of VR, including interactivity.

Chapter 3

Exposure to a Virtual Seismic Event

1. Introduction

Natural disasters pose significant threats to the environment, society, and individuals, causing widespread destruction and long-term consequences. On an environmental level, disasters such as earthquakes, hurricanes, and floods lead to habitat destruction, soil degradation, and loss of biodiversity (Gill & Malamud, 2017). Societally, they disrupt economies, displace populations, and strain healthcare and infrastructure systems, leading to prolonged recovery periods (Hallegatte & Przyluski, 2010). For individuals, natural disasters can have severe psychological repercussions, including anxiety, depression, and post-traumatic stress disorder (PTSD) (Neria et al., 2008). Among natural disasters, seismic events are one of the most frequent and destructive, affecting millions of people worldwide each year (e.g., United Nations Office for Disaster Risk Reduction, 2022).

Earthquakes strike suddenly and without warning, leading to significant casualties, infrastructure collapse, and economic devastation. Their unpredictable nature and high-impact consequences contribute to psychological distress in affected populations, with long-term mental health issues persisting long after the physical damage has been addressed (Galea et al., 2005).

One of the most severe psychological effects of seismic events is PTSD. PTSD is a psychiatric disorder that can develop following exposure to a traumatic event and is characterized by intrusive memories, avoidance behaviours, negative mood alterations, and hyperarousal symptoms (American Psychiatric Association, 2022). Earthquake survivors frequently exhibit PTSD symptoms, with an estimated prevalence of 23.66% (Dai et al., 2016), depending on the severity of exposure and individual risk factors (Hong & Efferth, 2015; Hosseinnejad et al., 2023). Without proper intervention, PTSD can severely impair daily functioning and quality of life.

While PTSD is a common consequence of traumatic events, individual responses to trauma can vary significantly based on psychological and personality-related factors. Trait anxiety, a stable dispositional tendency to experience heightened

worry and stress, has been shown to increase vulnerability to PTSD symptoms following traumatic events (Hensley & Varela, 2008; Knowles & Olatunji, 2020). Similarly, personality traits outlined by the Five-Factor Model (McCrae & John, 1992) also influence how individuals cope with trauma. High neuroticism has been consistently linked to increased PTSD severity, as individuals with this trait are more prone to emotional instability and maladaptive coping (Kotov et al., 2010). Conversely, higher conscientiousness and extraversion are often associated with resilience, as they contribute to adaptive coping mechanisms and social support-seeking behavior (Weinberg et al., 2021). In addition to personality factors, cognitive emotion regulation strategies play a crucial role in trauma adaptation. Adaptive strategies, such as positive reappraisal, acceptance, and perspective-taking, have been associated with reduced post-traumatic growth whereas maladaptive strategies like rumination, catastrophizing, and self-blame have been linked to more persistent PTSD symptoms and greater emotional distress (Seligowski et al., 2015). Understanding these individual differences is essential for tailoring PTSD interventions.

Among the many therapeutic approaches for PTSD, exposure therapy is considered the gold standard (Rauch et al., 2012), regardless of the type of trauma. Exposure therapy in imagination involves "*the graded and repeated imaginal reliving and narrative recounting of the traumatic event by the client within the therapeutic setting*" (Rizzo et al., 2015, p.256). Despite its effectiveness, traditional exposure therapy presents several challenges. Many patients find it distressing to relive their traumatic experiences, leading to reluctance or refusal to engage in treatment or struggle to vividly recall or reconstruct the traumatic event (Tarrier et al., 2000; van Minnen et al., 2002); and up to 41% of patients discontinue treatment prematurely (for a review, see Imel et al., 2013). This is crucial since the inability to emotionally engage is a predictor for negative treatment outcome (Jaycox et al., 1998). Virtual reality (VR) has been proposed as a way to overcome these limitations, offering a controlled, immersive, and interactive medium that enhances the effectiveness of exposure therapy (Botella et al., 2017). Unlike traditional methods that rely on imagination, VR provides a visual representation of the traumatic event, making it particularly useful for patients struggling with recall (Difede et al., 2007).

VR exposure therapy (VRET) allows clinicians to control stimulus intensity, ensuring a gradual and personalized exposure process (Rothbaum et al., 1999, 2001). Additionally, VR ensures standardized therapeutic stimuli, improving replicability and consistency in treatment (Maples-Keller et al., 2017).

VRET has already been widely applied to specific phobias, such as fear of heights, flying, and public speaking, showing strong efficacy in reducing fear responses (Parsons & Rizzo, 2008). In PTSD treatment, VRET has been particularly successful in combat-related trauma, helping veterans and active-duty military personnel process traumatic memories in a controlled and immersive environment (Rizzo et al., 2009). Studies have demonstrated that VR-based exposure therapy leads to significant reductions in PTSD symptoms, comparable to or even exceeding traditional exposure therapy (Gonçalves et al., 2012; Reger et al., 2016; McLay et al., 2017).

Despite its promising applications for PTSD treatment, VRET has not been systematically applied to earthquake-related trauma, a unique and highly impactful form of trauma which imposes significant psychological burden. Given that earthquakes differ from other trauma types in their widespread community impact, a tailored approach is necessary to develop effective VR-based exposure treatments. This gap highlights a need for research into how VR can be adapted for seismic PTSD treatment. The rationale for its use for this specific type of trauma rests on two key therapeutic aims: reducing trauma-related hyperarousal through controlled desensitization, and targeting trauma-related cues that are not tied to specific locations, but to sensory and bodily experiences such as tremors, loud noises, and disorientation, which are core elements of seismic trauma that are often difficult to reproduce safely in vivo. These aims align with several theoretical frameworks. According to Emotional Processing Theory (Foa & Kozak, 1986), effective exposure requires activation of the trauma memory so that new, non-threatening information can be integrated; VR provides a safe means of re-engaging with these core features of the traumatic event. Inhibitory Learning Theory (Craske et al., 2008, 2014) reframes exposure not as the erasure of fear, but as the creation of new safety associations that inhibit, rather than overwrite, the original fear memory. In this context, the goal is not to challenge the belief that seismic events can be

dangerous, but rather to disconfirm the learned association that all physical sensations linked to the original trauma, such as ground vibrations or bodily arousal, necessarily signal current threat. Finally, interoceptive and predictive processing models (e.g., Seth, 2013; Barrett, 2017) emphasize how PTSD disrupts the brain's interpretation of bodily signals, leading to heightened threat perception and hypervigilance. By safely reintroducing internal and sensory cues associated with the traumatic event, VR exposure may help recalibrate threat predictions and support the re-establishment of bodily and emotional regulation. Together, these frameworks provide a strong foundation for exploring VR-based exposure to seismic trauma.

The present study aims to evaluate the feasibility of using VR as a tool for exposure to a simulated seismic event, focusing on a non-clinical population as a preliminary step toward PTSD treatment applications. Specifically, the study seeks to: a) assess the VR scenario's realism, representativeness, and participants' feedback to ensure its validity as an exposure tool; b) examine the emotional, cognitive, and physiological responses of individuals exposed to the VR earthquake scenario; c) evaluate the scenario's usability, immersion, and engagement to refine its application in future clinical settings; d) investigate whether individual characteristics, such as trait anxiety, personality traits, and cognitive emotion regulation strategies, affect participants' responses to the VR scenario. The decision to test a non-clinical population is driven by both ethical and methodological considerations. Ethically, it minimizes potential distress among participants, ensuring a safe first step before clinical applications. Methodologically, it provides a baseline for determining whether the VR scenario effectively elicits emotional and cognitive engagement. These baseline responses are essential for contextualizing and interpreting data from future studies involving clinical populations; in fact, establishing normative patterns of response enables researchers to identify which reactions are typical and which might be indicative of trauma-related distress in clinical populations.

2. Materials & Method

2.1 Expert Feedback

The initial phase of this research involved contacting a group of psychotherapists (n = 10; M_{Age} = 44.80, SD_{Age} = 6.94) with expertise in treating trauma-related disorders and conducting exposure therapy. The purpose was to gather their feedback on the VR scenario and the experimental procedure. These experts were invited to participate and asked to complete an online survey hosted on Qualtrics. The survey collected the following information: i) sociodemographic data; ii) clinical experience details, including years of practice and therapeutic orientation; iii) knowledge and prior use of VR in clinical contexts, as well as their attitudes toward VR in therapy. The latter was assessed using items derived from the Unified Theory of Acceptance and Use of Technology (Di Natale et al., 2024). These items evaluated performance expectancy and effort expectancy of VR tools in clinical practice on a 5-point Likert scale (1 = "*Completely Disagree*," 5 = "*Completely Agree*"); iv) therapist expertise, measured with the 32-item Psychotherapy Expertise Questionnaire (Gori et al., 2023), which includes four subjective factors subscales (performance, cognitive functioning, personal and relational qualities of the therapist, therapist self-assessment) and four objective factors subscales (experience, reputation with clients and colleagues, training and professional updating, deontological ethics and setting rules).

Following the survey, the psychotherapists viewed a 360° interactive video, embedded in the survey, simulating a seismic event, identical to the scenario participants would later experience in VR (see "Experimental Procedure"). They were then asked to evaluate the scenario on a 4-point Likert scale (1 = "*Not at All*", 4 = "*Very Much*") based on its realism, representativeness, and potential usefulness in clinical practice.

Overall, the experts found the VR scenario highly representative of a seismic event, realistic, and useful. Although none of the experts had prior experience using VR technologies in clinical practice, the majority (60%) expressed a favourable attitude toward potentially integrating such tools in their work. The detailed results of the expert survey are presented in Table 9.

Table 9

Summary of Expert Feedback on the VR Scenario and its Potential Clinical Applications

Variable	M (SD) or n (%)
Years of Registration as a Psychotherapist	12.20 (5.92)
Therapeutic Orientation	
Cognitive-Behavioral	9 (90%)
Transactional Analysis	1 (10%)
Knowledge of VR Clinical Applications	
None	1 (10%)
A Little	6 (6%)
Moderate	3 (%)
Very High	0 (0%)
VR Performance Expectancy (range 3-15)	10.67 (0.87)
VR Effort Expectancy (range 3-15)	7.44 (2.19)
Representativeness (range 1-4)	3.70 (0.48)
Realism (range 1-4)	3.30 (0.48)
Usefulness (range 1-4)	3.43 (0.54)
PEQ Performance (range 4-20)	16.50 (1.18)
PEQ Cognitive Functioning (range 4-20)	16.20 (2.20)
PEQ Personal and Relational Qualities (range 4-20)	17.10 (2.42)
PEQ Therapist Self-Assessment (range 4-20)	17.00 (2.26)
PEQ Experience (range 4-20)	15.40 (3.20)
PEQ Reputation with Clients and Colleagues (range 4-20)	15.90 (2.08)
PEQ Training and Professional Updating (range 4-20)	14.70 (2.31)
PEQ Deontological Ethics and Setting Rules (range 4-20)	19.00 (1.25)

Note. VR: Virtual Reality; PEQ: Psychotherapy Expertise Questionnaire.

2.2 Participants

A total of 82 individuals volunteered for the experiment. Eligibility criteria required participants to: a) be at least 18 years old; b) be fluent in Italian; and c) prior experience of a seismic event of any intensity. Exclusion criteria were: a) uncorrected (with glasses or contact lenses) visual impairments; b) conditions or medical devices that contraindicate VR headset use (i.e., pregnancy, being over 65 years old, heart conditions, psychiatric disorders, motion sickness, balance

problems, seizures, pacemakers, defibrillators, or hearing aids); c) prior experience of a seismic event that resulted in the loss of a loved one or the destruction of one's home or workplace; d) a score exceeding 33 on the Impact of Event Scale-Revised (IES-R; see "Measures") for the most intense seismic event they had experienced. Based on these criteria, 10 participants were excluded: 8 for having uncorrected visual impairments, 1 due to an IES-R score higher than 33, and 1 due to having experienced a seismic event causing the loss of a loved one or the destruction of their home or workplace. This resulted in a final sample of 72 participants (M = 30.11, SD = 6.29; 61.1% female). Participants received no compensation for their participation.

2.3 Experimental Procedure

Data collection took place between May and October 2024. Experimental sessions were conducted during working hours (9:00-17:00) in a quiet room at the Department of Human Neuroscience, Sapienza University of Rome. The study adhered to the principles of the Declaration of Helsinki and was approved by the Institutional Review Board of Psychology (Department of Psychological, Health, and Territorial Sciences, University "G. d'Annunzio" of Chieti-Pescara, ethics approval reference 24008). The experimental procedure lasted approximately 30 minutes and comprised three stages.

Stage 1. After providing informed consent, participants completed a pre-exposure questionnaire via Qualtrics, which collected information on: i) sociodemographic data; ii) inclusion and exclusion criteria; iii) state and trait anxiety; iv) subjective units of distress (SUDs) and the intensity of various emotional states at that moment; v) emotional regulation strategies used to manage negative or stressful events; vi) personality traits.

Stage 2. Participants were fitted with a Scosche Rhythm 24 armband sensor (Ventura County, CA, USA) to record baseline physiological data. Then they were equipped with the Meta Quest 2 VR head-mounted display (HMD) to experience a simulated seismic event. In more detail, this VR scenario depicted a seismic event taking place inside a classroom and lasted 1 minute and 30 seconds (see Figure 8); it was developed by Prof. Tomoki Itamiya's laboratory at Kanagawa Dental

University (Yokosuka, Japan). Following the scenario, physiological data were measured again to evaluate recovery. The sensor continuously recorded data throughout the entire process, from baseline to recovery. Participants were sitting throughout the recording.

Stage 3. Participants completed a post-exposure questionnaire via Qualtrics. This questionnaire assessed: i) state anxiety; ii) SUDs and intensity of various emotional states at that moment; and iii) their experience of the VR scenario. Finally, participants underwent a psychological debriefing conducted by a psychologist to address any questions, ensure their well-being, and mitigate potential distress.

Figure 8

A Frame of the 360° Video Shown to Participants



2.4 Measures

2.4.1 Impact of Event Scale-Revised

The pre-exposure assessment included the Italian version (Craparo et al., 2013) of the Impact of Event Scale-Revised (IES-R; Weiss, 1997), a 22-item self-report screening tool designed to assess individuals' subjective responses to a traumatic event. The scale evaluates reactions to a specific traumatic event, which in this study was defined as the "most intense seismic event experienced". Each item is rated on a 5-point Likert scale, ranging from 0 ("*Not at all*") to 4 ("*Extremely*"). Total scores range from 0 to 88, with higher scores indicating greater symptom severity. A score above the established cut-off of 33, which has been shown to provide optimal diagnostic accuracy for post-traumatic symptoms (Creamer et al., 2003), served as the exclusion criterion for potential post-traumatic symptoms related to a seismic event. The Cronbach's alpha for the scale in this study was .89.

2.4.2 State-Trait Anxiety Inventory - Form Y

The Italian version (Pedrabissi & Santinello, 1989) of the State-Trait Anxiety Inventory Y Form (STAI-Y; Spielberger et al., 1983) is a self-administered questionnaire comprising two 20-item scales that assess state and trait anxiety separately. State anxiety is defined as "a transitory response to an event perceived as adverse, characterized by feelings of tension, apprehension, nervousness, and worry" (Ilardi et al., 2021, p.2760) while trait anxiety refers to "a stable predisposition to perceive stressful situations as dangerous or threatening" (Ilardi et al., 2021, p.2760). The trait anxiety scale was administered only during the pre-exposure assessment, while the state anxiety scale was administered both before and after exposure. Each item is rated on a 4-point Likert scale, from 1 ("*Almost Never*") to 4 ("*Almost Always*"). Scores on each scale range from 20 to 80, with higher scores indicating higher anxiety levels. The Cronbach's alpha for the trait anxiety scale was .92, for the pre-exposure state anxiety scale was .91, and for the post-exposure state anxiety scale was .93.

2.4.3 Subjective Units of Distress

Subjective Units of Distress (SUDs; Wolpe, 1990) is a self-report measure used to assess the level of distress and discomfort. Participants rated their current distress on a scale from 0 (no distress) to 100 (the highest level of distress ever experienced),

before and after exposure. SUD values were determined based on previous literature (Šalkevičius et al., 2019).

2.4.4 Emotions

Participants rated the intensity of primary emotions (joy, sadness, fear, disgust, anger, surprise; Ekman, 1992) before and after exposure, using a scale from 0 (complete absence of the emotion) to 10 (maximum intensity of the emotion).

2.4.5 Cognitive Emotion Regulation Questionnaire-Short

The pre-exposure assessment included the Italian version (Cerolini et al., 2022) of the Cognitive Emotion Regulation Questionnaire-Short (CERQ-S; Garnefski & Kraaij, 2006), an 18-item self-report tool designed to measure the cognitive strategies individuals use to regulate their emotions in response to stressful or threatening situations. Each item is rated on a 5-point scale, from 1 ("*Never*") to 5 ("*Always*"). The questionnaire is divided into nine subscales, each consisting of two items: Self-blame, Other-blame, Rumination, Catastrophizing, Putting into Perspective, Positive Refocusing, Positive Reappraisal, Acceptance, and Planning. Higher scores on each subscale (ranging from 4 to 20) reflect a greater use of that specific cognitive strategy. For the analyses, two total scores were computed: one for adaptive strategies (including putting into perspective, positive refocusing, positive reappraisal, acceptance, and planning) and another for maladaptive strategies (including self-blame, other-blame, rumination, and catastrophizing). These scores were obtained by summing the values of their respective subscales, consistent with previous research (e.g., Vanderhasselt et al., 2014). In this study, Cronbach's alpha for the subscales ranged from .76 to .97.

2.4.6 Big Five Inventory-10

The pre-exposure assessment included the Italian version (Guido et al., 2015) of the Big Five Inventory-10 (BFI-10; Rammstedt & John, 2007), a 10-item self-report questionnaire designed to assess participants' personality traits. Each item is rated on a 5-point scale, ranging from 1 ("*Strongly Disagree*") to 5 ("*Strongly Agree*"). The questionnaire consists of five subscales which measure the following

personality dimensions: agreeableness, openness, emotional stability, extraversion, and conscientiousness. In this study, Cronbach's alpha for the subscales ranged from .57 to .79.

2.4.7 Virtual Reality Experience Questions

Following the VR exposure, participants responded to a series of ad hoc questions designed to assess their experience with the scenario. These questions were rated on a 4-point scale, ranging from 1 ("*Not at All*") to 4 ("*Very Much*"). The questions addressed several aspects, including: realism ("*How realistic the scenario seemed*"), representativeness ("*How representative the scenario was of a seismic event*"), experienced realism ("*I felt as though I was actually living the situation*"), spatial presence ("*I felt physically present in the virtual environment*"), engagement ("*I was fully engaged by the virtual scenario*"). The experienced realism, spatial presence, and engagement were combined to calculate an overall presence score (ranging from 3 to 12) Cronbach's alpha for the computed presence score was .81.

2.4.8 Virtual Reality Neuroscience Questionnaire

The Virtual Reality Neuroscience Questionnaire (VRNQ; Kourtesis et al., 2019) is a 20-item tool designed to evaluate the overall quality of VR applications. After exposure, participants completed two 5-item subscales: "user experience" and "virtual reality-induced symptoms and effects (VRISE)". The user experience subscale uses a 7-point scale, ranging from 1 ("*Extremely Low*") to 7 ("*Extremely High*"), with higher scores reflecting a better experience based on factors like immersion, enjoyment, and the quality of the graphics, sound, and VR hardware. The VRISE subscale is also rated on a 7-point scale, from 1 ("*Extremely Intense Feeling*") to 7 ("*Absent*"), with higher scores indicating more severe VR-related symptoms, such as nausea, disorientation, dizziness, fatigue, and instability. In this study, Cronbach's alpha for the user experience subscale was .78 and for the VRISE subscale was .76.

2.4.9 Physiological Measure

The Scosche Rhythm 24 (Ventura County, CA, USA), a wearable armband photoplethysmography (PPG) sensor, was used to collect the raw RR intervals, defined as the time between two successive R-wave peaks in a traditional electrocardiogram (ECG). PPG is a non-invasive optical technique that continuously measures changes in blood volume, respiration, and oxygen saturation by detecting variations in light absorption caused by fluctuations in blood volume within the tissues (Chiang et al., 2018). The accuracy of the Scosche Rhythm 24 has been evaluated in various studies, with generally positive results (e.g., Reece et al., 2021), although some research has raised concerns about the overall accuracy of PPG sensors for clinical use (e.g., Vinatzer et al., 2022) when compared to traditional ECG.

2.5 Data Processing

The data processing for heart rate variability (HRV) analysis was conducted using Kubios Scientific software version 4.1.1 (Kubios Oy, Kuopio, Finland), a widely recognized tool for advanced HRV analysis. Raw R-R interval data were exported from the recording device and imported into Kubios for analysis. The software's built-in algorithms were utilized for noise detection and artifact correction, employing an automatic threshold-based approach to ensure the quality and reliability of the data. Beat correction was performed using the automatic beat correction algorithm, which is specifically designed to address missed, extra, and misaligned beats in the R-R interval time series. All HRV parameters were computed in accordance with the standard guidelines provided within the software (Tarvainen & Niskanen, 2008).

HRV-related variables were categorized into general, time-domain, and frequency-domain. General variables included the sympathetic nervous system index (SNS index), parasympathetic nervous system index (PNS index), and Stress Index. The SNS index reflects the relative contribution of sympathetic modulation, while the PNS index indicates parasympathetic modulation, providing measures of their activation compared to normal resting values. For both indexes, positive and negative values reflect above- and below-average activity, respectively. The Stress Index, calculated as the square root of Baevsky's stress index (Baevsky &

Chernikova, 2017), estimates the level of physiological stress by analysing R-R interval distribution and density.

Time-domain variables assessed the variability of R-R intervals over time and included the mean R-R interval, mean heart rate (HR), the standard deviation of normal-to-normal intervals (SDNN), reflecting overall HRV, and the root mean square of successive differences (RMSSD), which is widely considered an indicator of parasympathetic nervous system activity (Shaffer & Ginsberg, 2017), and the triangular interpolation of NN intervals (TINN), which represents the width of the histogram of these intervals. Additionally, the percentage of successive R-R intervals differing by more than 50ms (pNN50) was analysed, offering further insight into parasympathetic activity.

Frequency-domain parameters included respiratory rate (RESP) and a calculation of the power spectrum via the Fast Fourier transform (FFT) applied to the detrended RR series. Very Low (VLF, band value between 0-0.04 Hz), Low (LF, band value between 0.04-0.15 Hz), and High (HF, band value between 0.15-0.4 Hz) frequencies and the ratio between LF and HF were measured in ms^2 and log-transformed. Research suggested that HF reflects parasympathetic activity, whereas LF and VLF reflect sympathetic activity (for a review, Pham et al., 2021).

2.6 Statistical Analyses

Statistical analyses were conducted using SPSS version 28 (IBM, 2021). First, descriptive statistics were calculated to summarize the sample characteristics and participants' VR experience. Next, repeated measures one-way ANOVAs were performed to examine changes in self-report and physiological measures, with time as the independent factor (i.e., pre and post for self-report measures, and pre, during, and post for physiological measures). When significant differences were identified, linear regression analyses were conducted, using the changes (delta) between time points as dependent variables and psychological variables as independent variables. Due to the low reliability of the BFI-10 subscales, personality traits were excluded, and psychological variables were limited to participants reported adaptive or maladaptive coping strategies and trait anxiety.

An a priori power analysis was conducted using G*Power version 3.1.9.7 (Faul et al., 2007) to estimate the required sample size for the repeated measures ANOVA. With an effect size of $f = 0.25$, a significance level of $\alpha = 0.05$, and a power of 0.80, the minimum required sample size was 34 participants. The obtained sample size of 72 was therefore sufficient to test the study hypotheses.

3. Results

3.1 VR Experience

Participants' perceptions of the VR scenario were assessed using ad hoc questions and validated scales. Overall, participants rated the scenario as moderately realistic and representative. The presence total score, calculated from the experienced realism, spatial presence, and engagement items indicated modest immersion in the virtual environment. On the VRNQ, the user experience subscale reflected a generally positive evaluation of the VR scenario. Additionally, participants reported low VR-related symptoms, as indicated by the VRISE subscale. These findings underscore the scenario's perceived realism and representativeness, alongside moderate immersion and favourable user experiences with minimal adverse effects. Descriptive statistics are presented in Table 10.

Table 10

Descriptive Statistics for VR Experience and VRNQ subscales.

Variable	Mean	Median	*SD*	Range
Realism	2.97	3.00	0.58	1-4
Representativeness	3.28	3.00	0.59	1-4
VRNQ User Experience	22.97	23.50	3.68	5-35
VRNQ VRISE	31.97	33.00	3.82	5-35
Presence	4.57	5.00	1.97	3-12

Note. VRNQ: Virtual Reality Neuroscience Questionnaire; VRISE: Virtual Reality Induced Symptoms and Effects.

3.2 Repeated Measures ANOVAs

A series of repeated measures ANOVAs were conducted to examine the effect of exposure on self-reported and physiological measures. Bonferroni corrections were

applied to all post-hoc pairwise comparisons within each ANOVA. In addition, to account for multiple ANOVAs conducted on conceptually related outcomes within the same sample, Bonferroni corrections were also applied at the ANOVA level. Specifically, for the self-reported emotional outcomes (including state anxiety), the adjusted alpha threshold was .006; for general HRV variables, the corrected threshold was .017; for time-domain HRV variables, the threshold was .010; and for frequency-domain HRV variables, the corrected threshold was .010. Partial eta-squared (η^2) was calculated to determine effect sizes, with values of approximately .01, .06, and .14 indicating small, medium, and large effects, respectively (Richardson, 2011).

The effect of exposure (pre and post) on self-reported state anxiety was significant, $F(1,71) = 16.36$, $p < .001$, partial $\eta^2 = .19$, indicating a large effect size. Similarly, significant effects of exposure (pre and post) were observed for self-reported levels of experienced joy, $F(1,71) = 10.77$, $p = .002$, partial $\eta^2 = .13$, indicating a medium effect size.

For the SNS index, Mauchly's test confirmed that the assumption of sphericity was met, $\chi^2(2) = 0.90$, $p = .638$. The effect of exposure (pre, during, and post) was significant, $F(2,142) = 4.75$, $p = .010$, partial $\eta^2 = .08$, indicating a medium effect size. Post-hoc pairwise comparisons with a Bonferroni adjustment showed that SNS index scores were significantly lower pre-exposure than during exposure ($p = .018$). No significant differences were observed between pre-exposure and post-exposure scores ($p = 1.00$) or during exposure and post-exposure scores ($p = .073$). For the mean HR, Mauchly's test indicated that the assumption of sphericity was violated, $\chi^2(2) = 8.31$, $p = .016$, and therefore degrees of freedom were corrected using Huynh-Feldt estimates of sphericity ($\varepsilon = 0.90$). The effect of exposure (pre, during, and post) was significant, $F(1.81,101.15) = 18.04$, $p < .001$, partial $\eta^2 = .24$, indicating a large effect size. Post-hoc pairwise comparisons with a Bonferroni adjustment showed that mean HR scores were significantly lower pre-exposure than during exposure ($p < .001$), significantly lower pre-exposure than post-exposure ($p = .021$), and significantly higher during exposure than post-exposure ($p = .017$).

The effects of exposure (pre and post) on sadness ($p = .389$), fear ($p = .933$), disgust ($p = .888$), anger ($p = .206$), and SUDs ($p = .456$) were not significant,

The effects of exposure (pre, during, and post) on the PNS index (p = .074), the Stress Index (p = .331), SDNN (p = .549), RMSSD (p = .865), pNN50 (p = .635), TINN (p = .075), RESP (p = .134), LF/HF ratio (p = .447), LF power log (p = .110), and HF power log (p = .569) were not significant.

Surprise ($F(1,71) = 5.78$, p = .019) and VLF ($F(2,142) = 3.41$, p = .036) were significant prior to applying the Bonferroni correction but were no longer significant after adjusting for multiple comparisons. Although these results do not meet the more stringent thresholds after correction, they still demonstrate moderate effect sizes (partial η^2 = .07 for surprise, partial η^2 = .06 for VLF), which suggests that these findings could be meaningful. Therefore, while these results should be interpreted with caution, they warrant further exploration, as discussed in the Discussion section.

Means and standard deviations for all analysed variables are presented in Table 11.

Table 11

Summary of Descriptive Statistics for Self-Reported and Physiological Measures Across Exposure Phases

Variable	**Pre-Exposure** *M (SD)*	**During Exposure** *M (SD)*	**Post-Exposure** *M (SD)*
State Anxiety	**33.99 (7.12)**	/	**37.61 (8.42)**
SUDs	31.22 (17.53)	/	30.42 (16.89)
Joy	**4.26 (2.90)**	/	**3.67 (3.01)**
Sadness	1.14 (1.80)	/	1.24 (1.78)
Fear	1.07 (2.00)	/	1.06 (1.72)
Disgust	0.21 (0.87)	/	0.22 (0.68)
Anger	0.58 (1.55)	/	0.44 (1.17)
Surprise	1.21 (2.01)	/	1.63 (2.18)
SNS Index	**0.70 (1.29)**	**0.90 (1.39)**	0.73 (1.28)
PNS Index	-0.11 (1.31)	-0.28 (1.17)	-0.22 (1.11)
Stress Index	10.80 (3.63)	10.88 (4.05)	10.38 (3.78)
Mean HR	**73.87 (12.46)**	**76.68 (12.65)**	**75.10 (11.83)**
SDNN	49.92 (19.11)	50.06 (19.22)	51.83 (17.21)
RMSSD	51.10 (28.03)	50.19 (23.84)	50.27 (22.32)
pNN50	27.19 (20.10)	26.09 (18.22)	26.01 (17.42)

TINN	209.63 (72.66)	228.00 (89.39)	227.79 (76.01)
RESP	0.25 (0.05)	0.26 (0.05)	0.25 (0.05)
VLF power (log)	4.00 (1.00)	4.44 (1.01)	4.33 (1.13)
LF power (log)	6.84 (0.77)	6.71 (0.83)	6.96 (0.91)
HF power (log)	6.60 (1.00)	6.49 (0.91)	6.56 (0.99)
LF/HF ratio (ms^2)	1.89 (2.07)	1.79 (1.53)	2.09 (2.11)

Note. Significant differences in bold. VLF, LF, and HF bands are their respective absolute powers' natural logarithm transformed values. SUDs: Subjective Units of Distress; SNS: Sympathetic Nervous System; PNS: Parasympathetic Nervous System; HR: Heart Rate; SDNN: Standard Deviation of the Normal-to-Normal interbeat intervals; RMSSD: Root Mean Square of Successive Differences between normal heartbeats; TINN: Triangular Interpolation of Normal-to-Normal interval histogram; RESP: respiratory rate; VLF: Very Low Frequency; LF: Low Frequency; HF: High Frequency.

3.3 Regression Analyses

Linear regression analyses were conducted to evaluate the extent to which adaptive or maladaptive cognitive coping strategies and trait anxiety could predict the changes between time points in the physiological measures identified as significant in the ANOVAs (i.e., SNS index, mean HR, VLF power). No significant regression was found for mean HR (delta between post-during) $F(3,68) = 0.67$, $p = .575$, mean HR (delta between during-pre) $F(3,68) = 0.86$, $p = .469$, mean HR (delta between post-pre) $F(3,68) = 0.49$, $p = .689$, SNS index (delta between during-pre) $F(3,68) = 0.69$, $p = .687$, and VLF power (delta between during-pre) $F(3,68) = 1.03$, $p = .387$.

4. Discussion

The present study aimed to evaluate the feasibility of using VR as a tool for exposure to a simulated seismic event, serving as a preliminary step toward its potential application in PTSD treatment. Specifically, the study examined the scenario's realism and representativeness, participants' emotional, cognitive, and physiological responses, its usability and engagement, and the role of individual differences—such as trait anxiety, personality traits, and cognitive emotion regulation strategies—in shaping responses to the VR exposure. The findings provide critical insights into the potential of VR for earthquake-related PTSD interventions and offer a foundation for future research in clinical populations.

Findings suggested that participants and experts perceived the VR scenario as representative of a seismic event and moderately realistic. The presence score revealed modest immersion, while the user experience subscale of the VRNQ highlighted generally favourable evaluations user-experience-wise, albeit below the suggested cut-off of 25 (Kourtesis et al., 2019). Importantly, participants reported minimal adverse effects, as indicated by the VRISE subscale scoring well above the conservative cut-off of 30. The interplay between presence and the absence of significant VRISE provides an interesting perspective. Despite the relatively low presence, the scenario could still evoke meaningful emotional and physiological responses, emphasizing that high presence may not always be necessary to elicit a significant autonomic or emotional response. Moreover, the absence of adverse VRISE suggests that the scenario successfully balanced emotional activation without overwhelming participants, a key factor for effective VR applications in controlled research and therapeutic settings.

Repeated measures ANOVAs revealed significant effects of exposure (pre and post) on self-reported state anxiety, experienced joy, and surprise, with partial η^2 values indicating medium to large effect sizes. These results underscore the emotional impact of the VR scenario and its capacity to modulate participants' affective states. Similar findings have been documented in studies where VR exposure elicited significant emotional responses, supporting its efficacy as a tool for emotion induction (Riva et al., 2021). In more detail, the increase in state anxiety suggests that participants consciously perceived their physiological activation. The decrease in joy is consistent with the notion that stress and anxiety narrow attentional focus toward threat-related cues, reducing the availability of positive affective states (Fredrickson, 1998). This shift in emotional processing is well-documented in stress research, where heightened vigilance in response to a perceived threat is accompanied by a suppression of positive emotions (e.g., Richardson, 2017). The increase in surprise is in line with findings from VR studies (e.g., Pallavicini et al., 2019); a possible explanation is that while participants were aware they were taking part in a VR study, they had not previously experienced an earthquake in this format. The intensity of immersion, coupled with the realism of the simulation, may have exceeded their expectations, eliciting a heightened

surprise response. While this effect should be interpreted with caution due to the statistical threshold applied, its consistency with prior literature and the observed effect size suggest it may still reflect a meaningful emotional reaction to the VR scenario.

Physiological measures further confirmed the effects of exposure, with the SNS index and mean HR showing significant changes across phases. Elevated mean HR and SNS index during exposure reflect heightened sympathetic nervous system activation in response to the VR scenario. This autonomic response aligns with a stress-induced arousal pattern (Kreibig, 2010). The sustained elevation in HR post-exposure compared to pre-exposure suggests a delayed recovery, possibly due to residual arousal or cognitive processing of the VR experience. Elevated HR in stress-inducing VR environments has been widely reported. For instance, VR simulations designed to mimic phobic or stressful situations consistently show significant HR increases, indicating effective activation of the autonomic stress response (e.g., Freitas et al., 2021). However, most time-domain measures, such as SDNN and pNN50, did not show significant changes. This discrepancy might stem from the short measurement intervals used in this study, which may not fully capture the full spectrum of autonomic activity. While short readings are practical for capturing immediate physiological shifts, they can reduce sensitivity to gradual or sustained autonomic changes, a limitation noted in prior research (Laborde et al., 2017). Another possible explanation lies in the notion that different physiological measures capture distinct facets of autonomic regulation and stress, with HRV components uniquely reflecting the dynamic interaction between sympathetic and parasympathetic activity that includes activation modes operating on different time scales (Agorastos et al., 2023). Lastly, the VR scenario might not have been sufficiently arousing to alter measures that might require more pronounced stressors to show detectable changes (Kim et al., 2018). Regarding frequency-domain measures, the significant effect on VLF power also underscores autonomic adjustments during the VR scenario, potentially reflecting stress-related processes. VLF power is, in fact, often associated with stress-related mechanisms, including thermoregulation and hormonal influences, and is particularly sensitive to arousal states (Laborde et al., 2017; Shaffer & Ginsberg, 2017). However, given the

multiple physiological processes contributing to VLF, such as hormonal, metabolic, and non-autonomic factors (Shaffer & Ginsberg, 2017), its exact role remains incompletely understood. The observed increase in VLF power during exposure compared to baseline, along with the absence of significant differences between baseline and recovery, contrasts with previous findings suggesting that VLF decreases during mental stress tasks and recovers slowly post-stress (Usui & Nishida, 2017). This outcome, while falling short of corrected significance levels, may still be informative, particularly in light of the moderate effect size and the methodological constraints, such as the relatively short recording period. Indeed, Kuusela (2013) recommends a recording duration of at least 5 minutes to maximize the interpretability of this index. Future research should extend both the exposure duration and physiological monitoring to explore these dynamics in greater depth. In summary, the VR scenario effectively elicited significant emotional and physiological responses while maintaining minimal VR-induced symptoms, highlighting its potential value in clinical and therapeutic applications. This is particularly important for exposure therapy, as emotional engagement "*is theorized to be a key component of emotional processing, and successful emotional processing should translate to the reduction of PTSD symptoms*" (Reger et al., 2019, p.75). Furthermore, the inability to achieve emotional engagement has been identified as a predictor of poor treatment outcomes (Jaycox et al., 1998). Previous research has emphasized that cognitive emotional regulation strategies influence the development of post-traumatic stress disorder symptoms (Puechlong et al., 2020) and the modulation of arousal (e.g., Pizzie et al., 2021). However, despite significant changes in physiological measures observed in the present study, regression analyses revealed that these strategies did not significantly contribute to modulating such changes. This result might be attributable to the variability in the role of cognitive emotional regulation strategies depending on emotion intensity (Langeslag & Surti, 2017). Another plausible explanation is that being a simulated seismic event may have influenced the salience and relevance of coping strategies and trait anxiety in this context. In other words, the physiological arousal observed could primarily reflect an autonomic response to the immersive and engaging VR scenario, but may not fully translate into the psychological processes associated

with coping in a real threat situation. In actual seismic events, the unpredictability, immediacy of danger, and genuine risk of harm activate both physiological stress responses and complex psychological coping mechanisms. These mechanisms often involve conscious appraisals of threat and the deployment of adaptive or maladaptive strategies to manage fear, uncertainty, and perceived vulnerability (Lazarus & Folkman, 1984). In contrast, the VR scenario, while moderately engaging, was appropriately not perceived as life-threatening, as participants were aware it was a controlled and simulated environment. This likely resulted in less activation of these psychological coping processes typically associated with real-world threats. The observed discrepancy between heightened physiological arousal and non-significant regression findings may reflect a distinction between automatic physiological responses and the more deliberate cognitive processes underlying coping strategies. VR-induced arousal might rely more heavily on automatic systems, such as the sympathetic nervous system, which are not as directly influenced by individual differences in trait anxiety or cognitive coping strategies when a genuine threat is absent. Furthermore, the brief duration of the VR exposure could have amplified the immediate physiological responses while limiting the time available for participants to consciously engage their coping mechanisms. Future studies could investigate whether longer VR scenarios might strengthen the relationship between trait anxiety, coping strategies, and physiological responses. Additionally, scenarios designed to evoke a stronger perception of real-world stakes might help bridge the gap between autonomic arousal and psychological processes. Despite its contributions, this study has several limitations. First, the sample consisted of a non-clinical population, limiting the generalizability of findings to individuals with PTSD or those who have experienced severe earthquake-related trauma. Future studies should include clinical populations to better assess VR's effectiveness in therapeutic settings. Second, the duration of VR exposure was relatively short, which may not fully capture the long-term emotional and physiological effects of earthquake-related trauma. Prolonged exposure scenarios could provide a more comprehensive understanding of VR's potential for PTSD treatment and investigating the long-term effects of repeated VR exposure sessions would help determine the durability of therapeutic benefits. Third, while

physiological measures such as HRV and subjective distress ratings were analysed, additional neurobiological markers (e.g., cortisol levels, electroencephalography activity) could offer deeper insights into the neurophysiological underpinnings of VR exposure. Fourth, the study relied on self-report measures, which are susceptible to response biases; incorporating behavioural assessments and clinician-rated measures would enhance validity. Fifth, another potential limitation of this study is the use of the IES-R, which is based on DSM-IV criteria. While the IES-R remains a reliable tool for screening PTSD (Chang et al., 2024) and allows comparability with past studies, future research should employ DSM-5-based instruments, such as the PTSD Checklist for DSM-5 (PCL-5), to ensure alignment with current diagnostic criteria. Finally, the VR scenario, while deemed realistic and representative, lacked interactive elements. Increased interactivity may further enhance immersion and ecological validity, making VR exposure more comparable to real-world experiences; additionally, building on the promising results of Başoğlu and colleagues (2003), enhancing VR scenarios with multi-sensory feedback, especially haptic, may further increase realism and engagement.

This study highlights the feasibility of using virtual reality as a tool for earthquake-related trauma exposure, demonstrating its capacity to elicit significant emotional and physiological responses while minimizing adverse effects. The findings suggest that VR can serve as a promising method for controlled trauma exposure, with potential applications in PTSD treatment. Although further research is needed to validate its effectiveness in clinical settings, VR offers a novel and adaptable approach that may enhance traditional exposure therapy by increasing engagement and standardizing therapeutic conditions. By addressing the limitations and expanding future investigations, VR could become a valuable tool in the psychological treatment of trauma survivors.

General Results and Conclusions

This research set out to evaluate the feasibility of using VR as a tool for trauma-related interventions, particularly in the treatment of earthquake-related PTSD. To do so, the studies encompassed in this thesis explored two interconnected aspects: how memory processes operate in VR compared to real-life experiences, and how individuals respond emotionally, cognitively, and physiologically to a VR-simulated earthquake scenario. Findings from Chapters 1 and 2 revealed that memory performance in VR differs significantly from real life experiences, with VR failing to fully replicate real-world memory encoding and recall. Real-life exposure led to superior memory performance across free recall, visual recognition, and suggestibility tasks, while VR responses were more comparable to those observed in a traditional 2D format. Emotional valence and typicality influenced memory encoding differently depending on the medium, with positive items being better remembered in VR and 2D, while atypical objects were more salient in real-life settings. These results suggest that VR does not inherently enhance memory recall compared to 2D images and cannot be assumed to be an ecologically equivalent substitute for real-world experiences. Chapter 3 built upon these insights by examining how participants responded to a VR-simulated seismic event through a combination of self-reported emotional responses and physiological measures (heart rate variability). Results showed that the VR earthquake scenario elicited significant increases in state anxiety, heart rate, and sympathetic nervous system activation, demonstrating that VR can effectively induce emotional and physiological arousal. However, no significant associations were found between individual psychological traits (such as trait anxiety or emotion regulation strategies) and physiological responses, suggesting that VR-induced arousal operates differently from real-world trauma responses and may be less influenced by pre-existing coping tendencies.

The results of this research lead to two key conclusions. First, VR does not fully replicate real-life memory encoding and recall, which challenges the assumption that it provides a naturally ecologically valid alternative for memory-based interventions. However, this difference does not undermine its therapeutic value;

rather, it highlights the need to strategically tailor VR-based interventions to account for its unique cognitive and psychological properties.

Second, while VR may not recreate real-life memory processes, it still successfully elicits significant emotional and physiological responses, reinforcing its potential as an effective medium for trauma treatment. Crucially, the fact that VR-induced memory engagement is less intense than real-life recall can be seen as a therapeutic advantage rather than a limitation. Rather than triggering overwhelming re-experiencing of trauma, VR offers a controlled and gradual exposure setting, making it particularly useful for early-stage PTSD treatment, where lower-intensity exposures are beneficial. This ability to moderate emotional intensity may help patients engage with trauma-related material without becoming emotionally flooded, providing a safer and more adaptable pathway for therapeutic progress. It is, however, important to note that the studies in Chapters 1 and 2 investigated how individuals encode and recall everyday experiences in different environments, thus these insights may not necessarily extend to traumatic memories; indeed, future research should explicitly examine how VR interacts with trauma-specific memory distortions, intrusions, and reconsolidation processes to refine its application in exposure-based therapy.

In conclusion, this thesis reframes the role of VR in trauma therapy, arguing that its value lies not in its perfect replication of reality, but in its distinctiveness from it. By leveraging this controlled engagement, VR-based exposure therapy can be optimized to balance emotional activation and safety, paving the way for innovative and accessible PTSD interventions.

Software and Resources

https://www.kubios.com/

https://www.scosche.com/

https://www.meta.com/it/quest/products/quest-2/

https://www.labpano.com/en/products/PilotOne

https://www.ibm.com/it-it/products/spss-statistics

https://www.r-project.org/

https://www.qualtrics.com/

https://www.scopus.com/

https://www.ebsco.com/products/research-databases/apa-psycinfo

https://www.webofscience.com/wos

https://pubmed.ncbi.nlm.nih.gov/

https://scholar.google.com/

References

Aggleton, J. P., & Brown, M. W. (2006). Interleaving brain systems for episodic and recognition memory. *Trends in Cognitive Sciences*, *10*(10), 455–463. https://doi.org/10.1016/j.tics.2006.08.003

Agorastos, A., Mansueto, A. C., Hager, T., Pappi, E., Gardikioti, A., & Stiedl, O. (2023). Heart rate variability as a Translational dynamic biomarker of altered autonomic function in Health and Psychiatric disease. *Biomedicines*, *11*(6), 1591. https://doi.org/10.3390/biomedicines11061591

Anderson, M. C. (2020). *Episodic memory: Organizing and remembering*. In A. Baddeley, M. W. Eysenck, & M. C. Anderson (Eds.), Memory (pp. 163–206). Routledge.

American Psychiatric Association. (2022). *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, Text Revision. https://doi/book/10.1176/appi.books.9780890425787

Baddeley, A. D. (2004). *The psychology of memory*. In A. D. Baddeley, M. Kopelman, & B. A. Wilson (Eds.), The essential handbook of memory disorders for clinicians (pp. 1–13). Wiley.

Baddeley, A. D., Eysenck, M. W., & Anderson, M. C. (2020). *Memory*. Routledge.

Baevsky, R. M., & Chernikova, A. G. (2017). Heart rate variability analysis: physiological foundations and main methods. *Cardiometry*, 10. https://doi.org/10.12710/CARDIOMETRY.2017.10.6676

Barreda-Ángeles, M., Aleix-Guillaume, S., & Pereda-Baños, A. (2021). Virtual reality storytelling as a double-edged sword: Immersive presentation of nonfiction 360-video is associated with impaired cognitive information processing. *Communication Monographs*, *88*(2), 154–173. https://doi.org/10.1080/03637751.2020.1803496

Barrett, L. F. (2017). The theory of constructed emotion: an active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, *12*(1), 1-23. https://doi.org/10.1093/scan/nsw154

Başoğlu, M., Livanou, M., & Şalcıoğlu, E. (2003). A single session with an earthquake simulator for traumatic stress in earthquake survivors. *American Journal of Psychiatry*, *160*(4), 788-790. https://doi.org/10.1176/ajp.160.4.788

Bell, P.A., Greene, T.C., Fisher, J.D. and Baum, A. (2001). *Environmental Psychology* (5th Edition), Harcourt College Publishers, New York.

Beschin, N., Urbano, T., & Treccani, B. (2013). *RMBT-3 Adattamento Italiano*. OS Organizzazioni Speciali Giunti Editore.

Botella, C., Fernández-Álvarez, J., Guillén, V., García-Palacios, A., & Baños, R. (2017). Recent progress in virtual reality exposure therapy for phobias: a systematic review. *Current Psychiatry Reports*, *19*, 1-13. https://doi.org/10.1007/s11920-017-0788-4

Bylinskii, Z., Isola, P., Bainbridge, C., Torralba, A., & Oliva, A. (2015). Intrinsic and extrinsic effects on image memorability. *Vision Research*, *116*, 165–178. https://doi.org/10.1016/j.visres.2015.03.005

Cerolini, S., Zagaria, A., Vacca, M., Spinhoven, P., Violani, C., & Lombardo, C. (2022). Cognitive emotion Regulation Questionnaire—Short: reliability, validity, and Measurement Invariance of the Italian version. *Behavioral Sciences*, *12*(12), 474. https://doi.org/10.3390/bs12120474

Chang, S., Kim, W. H., Jung, Y. E., Roh, D., Kim, D., Chae, J. H., & Park, J. E. (2024). Clinical Utility of Impact of Event Scale–Revised for Diagnostic and Statistical Manual of Mental Disorders-Posttraumatic Stress Disorder. *Psychiatry Investigation*, *21*(8), 870. https://doi.org/10.30773/pi.2024.0147

Chiang, P. Y., Chao, P. C. P., Yang, C. Y., & Tarng, D. C. (2018). Theoretical developments and clinical experiments of measuring blood flow volume (BFV) at arteriovenous fistula (AVF) using a photoplethysmography (PPG) sensor.

Microsystem Technologies, *24*, 4587-4603. https://doi.org/10.1007/s00542-018-3892-4

Craparo, G., Faraci, P., Rotondo, G., & Gori, A. (2013). The Impact of Event Scale-Revised: Psychometric properties of the Italian version in a sample of flood victims. *Neuropsychiatric Disease and Treatment*, *9*, 1427-1432. https://doi.org/10.2147/NDT.S51793

Craske, M. G., Kircanski, K., Zelikowsky, M., Mystkowski, J., Chowdhury, N., & Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour Research and Therapy*, *46*(1), 5-27. https://doi.org/10.1016/j.brat.2007.10.003

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy*, *58*, 10-23. https://doi.org/10.1016/j.brat.2014.04.006

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd Edition), Routledge, New York. https://doi.org/10.4324/9780203771587

Creamer, M., Bell, R., & Failla, S. (2003). Psychometric properties of the Impact of Event Scale-Revised. *Behaviour Research and Therapy*, *41*(12), 1489-1496. https://doi.org/10.1016/j.brat.2003.07.010

Dai, W., Chen, L., Lai, Z., Li, Y., Wang, J., & Liu, A. (2016). The incidence of post-traumatic stress disorder among survivors after earthquakes: a systematic review and meta-analysis. *BMC Psychiatry*, *16*, 1-11. https://doi.org/10.1186/s12888-016-0891-9

De Renzi, E., & Nichelli, P. (1975). Verbal and non-verbal short-term memory impairment following hemispheric damage. *Cortex*, *11*(4), 341-354. https://doi.org/10.1016/s0010-9452(75)80026-8

Di Natale, A. F., Bartolotta, S., Gaggioli, A., Riva, G., & Villani, D. (2024). Exploring students' acceptance and continuance intention in using immersive virtual reality and metaverse integrated learning environments: The case of an

Italian university course. *Education and Information Technologies*, 1-20. https://doi.org/10.1007/s10639-023-12436-7

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: A three-component model. Trends in *Cognitive Sciences*, *11*(9), 379–386. https://doi.org/10.1016/j.tics.2007.08.001

Difede, J., Cukor, J., Jayasinghe, N., Patt, I., Jedel, S., Spielman, L., ... & Hoffman, H. G. (2007). Virtual reality exposure therapy for the treatment of posttraumatic stress disorder following September 11, 2001. *Journal of Clinical Psychiatry*, *68*(11), 1639.

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, *6*(3-4), 169-200. https://doi.org/10.1080/02699939208411068

Elmqaddem, N. (2019). Augmented reality and virtual reality in education. Myth or reality?. *International Journal of Emerging Technologies in Learning*, *14*(3), 234-242. https://doi.org/10.3991/ijet.v14i03.9289

Elsey, J. W., van Andel, K., Kater, R. B., Reints, I. M., & Spiering, M. (2019). The impact of virtual reality versus 2D pornography on sexual arousal and presence. *Computers in Human Behavior*, *97*, 35-43. https://doi.org/10.1016/j.chb.2019.02.031

Ernstsen, J., Mallam, S. C., & Nazir, S. (2019, November). Incidental memory recall in virtual reality: An empirical investigation. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 63, No. 1, pp. 2277-2281). Sage CA: Los Angeles, CA: SAGE Publications. https://doi.org/10.1177/1071181319631411

Fairfield, B., Di Domenico, A., Serricchio, S., Borella, E., & Mammarella, N. (2017). Emotional prosody effects on verbal memory in older and younger adults. *Aging, Neuropsychology, and Cognition*, *24*(4), 408–417. https://doi.org/10.1080/13825585.2016.1219690

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and

biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191. https://doi.org/10.3758/BF03193146

Flannery, K. A., & Walles, R. (2003). How does schema theory apply to real versus virtual memories? *CyberPsychology & Behavior*, *6*(2), 151–159. https://doi.org/10.1089/109493103321640347

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: exposure to corrective information. *Psychological Bulletin*, *99*(1), 20. https://doi.org/10.1037/0033-2909.99.1.20

Fredrickson, B. L. (1998). What good are positive emotions?. *Review of General Psychology*, *2*(3), 300-319. https://doi.org/10.1037/1089-2680.2.3.300

Freeman, J., Avons, S. E., Pearson, D. E., & IJsselsteijn, W. A. (1999). Effects of sensory information and prior experience on direct subjective ratings of presence. *Presence*, *8*(1), 1-13. https://doi.org/10.1162/105474699566017

Freeman, J., Avons, S. E., Meddis, R., Pearson, D. E., & IJsselsteijn, W. (2000). Using behavioral realism to estimate presence: A study of the utility of postural responses to motion stimuli. *Presence: Teleoperators & Virtual Environments*, *9*(2), 149-164. https://doi.org/10.1162/105474600566691

Freitas, J. R. S., Velosa, V. H. S., Abreu, L. T. N., Jardim, R. L., Santos, J. A. V., Peres, B., & Campos, P. F. (2021). Virtual reality exposure treatment in phobias: a systematic review. *Psychiatric Quarterly*, *92*(4), 1685-1710. https://doi.org/10.1007/s11126-021-09935-6

Galea, S., Nandi, A., & Vlahov, D. (2005). The epidemiology of post-traumatic stress disorder after disasters. *Epidemiologic Reviews*, 27(1), 78-91. https://doi.org/10.1093/epirev/mxi003

Garnefski, N., & Kraaij, V. (2006). Cognitive emotion regulation questionnaire–development of a short 18-item version (CERQ-short). *Personality and Individual Differences*, *41*(6), 1045-1053. https://doi.org/10.1016/j.paid.2006.04.010

Gill, J. C., & Malamud, B. D. (2017). Anthropogenic processes, natural hazards, and interactions in a multi-hazard framework. *Earth-Science Reviews*, 166, 246-269. https://doi.org/10.1016/j.earscirev.2017.01.002

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*, *66*(3), 325-331. https://doi.org/10.1111/j.2044-8295.1975.tb01468.x

Gonçalves, R., Pedrozo, A. L., Coutinho, E. S. F., Figueira, I., & Ventura, P. (2012). Efficacy of virtual reality exposure therapy in the treatment of PTSD: a systematic review. *PloS one*, *7*(12), e48469. https://doi.org/10.1371/journal.pone.0048469

Gori, A., Topino, E., Cacioppo, M., Schimmenti, A., & Caretti, V. (2023). Definition and Criteria for the Assessment of Expertise in Psychotherapy: Development of the Psychotherapy Expertise Questionnaire (PEQ). *European Journal of Investigation in Health, Psychology and Education*, *13*(11), 2478-2497. https://doi.org/10.3390/ejihpe13110173

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1, pp. 1969–2012). Wiley.

Gudjonsson, G. H. (1997). *The Gudjonsson Suggestibility Scales*. Hove: Psychology Press.

Gudjonsson, G. H., & Clark, N. K. (1986). Suggestibility in police interrogation: A social psychological model. *Social Behaviour*, *1*(2), 83-104.

Guido, G., Peluso, A. M., Capestro, M., & Miglietta, M. (2015). An Italian version of the 10-item Big Five Inventory: An application to hedonic and utilitarian shopping values. *Personality and Individual Differences*, *76*, 135-140. https://doi.org/10.1016/j.paid.2014.11.053

Hallegatte, S., & Przyluski, V. (2010). The economics of natural disasters: concepts and methods. *World Bank Policy Research Working Paper*, (5507).

Hensley, L., & Varela, R. E. (2008). PTSD symptoms and somatic complaints following Hurricane Katrina: The roles of trait anxiety and anxiety sensitivity.

Journal of Clinical Child & Adolescent Psychology, *37*(3), 542-552. https://doi.org/10.1080/15374410802148186

Hodges, L. F., Kooper, R., Meyer, T. C., De Graaff, J. J. H., Rothbaum, B. O., Opdyke, D., ... & North, M. M. (1994). Presence as the defining factor in a VR application. (Technical Report No. GIT-GVU-94-5). Georgia Institute of Technology. http://hdl.handle.net/1853/3584

Hong, C., & Efferth, T. (2016). Systematic review on post-traumatic stress disorder among survivors of the Wenchuan earthquake. *Trauma, Violence, & Abuse*, *17*(5), 542-561. https://doi.org/10.1177/1524838015585313

Hosseinnejad, M., Yazdi-Feyzabadi, V., Hajebi, A., Bahramnejad, A., Baneshi, R., Sarabi, R. E., ... & Zolala, F. (2022). Prevalence of posttraumatic stress disorder following the earthquake in Iran and Pakistan: a systematic review and meta-analysis. *Disaster Medicine and Public Health Preparedness*, *16*(2), 801-808. https://doi.org/10.1017/dmp.2020.411

Hu-Au, E., & Okita, S. (2021). Exploring differences in student learning and behavior between real-life and virtual reality chemistry laboratories. *Journal of Science Education and Technology*, *30*(6), 862-876. https://doi.org/10.1007/s10956-021-09925-0

Hudson, S., Matson-Barkat, S., Pallamin, N., & Jegou, G. (2019). With or without you? Interaction and immersion in a virtual reality experience. *Journal of Business Research*, *100*, 459-468. https://doi.org/10.1016/j.jbusres.2018.10.062

Iacus, S. M., King, G., & Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political Analysis*, *20*(1), 1-24. https://doi.org/10.1093/pan/mpr013

IBM Corp. (2021). IBM SPSS statistics for Windows, Version 28.0. IBM Corp.

Ilardi, C. R., Gamboz, N., Iavarone, A., Chieffi, S., & Brandimonte, M. A. (2021). Psychometric properties of the STAI-Y scales and normative data in an Italian elderly population. *Aging Clinical and Experimental Research*, 33, 2759-2766. https://doi.org/10.1007/s40520-021-01815-0

Imel, Z. E., Laska, K., Jakupcak, M., & Simpson, T. L. (2013). Meta-analysis of dropout in treatments for posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology*, *81*(3), 394. https://doi.org/10.1037/a0031474

Jaycox, L. H., Foa, E. B., & Morral, A. R. (1998). Influence of emotional engagement and habituation on exposure therapy for PTSD. *Journal of Consulting and Clinical Psychology*, *66*(1), 185.

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies*, *66*(9), 641-661. https://doi.org/10.1016/j.ijhcs.2008.04.004

Johnsdorf, M., Kisker, J., Gruber, T., & Schöne, B. (2023). Comparing encoding mechanisms in realistic virtual reality and conventional 2D laboratory settings: Event-related potentials in a repetition suppression paradigm. *Frontiers in Psychology*, *14*, 1051938. https://doi.org/10.3389/fpsyg.2023.1051938

Johnson-Glenberg, M.C. (2019). *The Necessary Nine: Design Principles for Embodied VR and Active Stem Education*. In: Díaz, P., Ioannou, A., Bhagat, K., Spector, J. (eds) Learning in a Digital World. Smart Computing and Intelligence. Springer, Singapore. https://doi.org/10.1007/978-981-13-8265-9_5

Johnson-Glenberg, M. C., Bartolomea, H., & Kalina, E. (2021). Platform is not destiny: Embodied learning effects comparing 2D desktop to 3D virtual reality STEM experiences. *Journal of Computer Assisted Learning*, *37*(5), 1263-1284. https://doi.org/10.1111/jcal.12567

Kensinger, E. A. (2004). Remembering emotional experiences: The contribution of valence and arousal. *Reviews in the Neurosciences*, *15*(4), 241–252. https://doi.org/10.1515/REVNEURO.2004.15.4.241

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition*, *31*(8), 1169–1180. https://doi.org/10.3758/BF03195800 ù

Kensinger, E. A., & Corkin, S. (2004). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences*, *101*(9), 3310–3315. https://doi.org/10.1073/pnas.0306408101

Kensinger, E. A., & Schacter, D. L. (2016). *Memory and emotion*. In L. F. Barrett, M. Lewis, & J. M. Haviland-Jones (Eds.), Handbook of emotions (pp. 564–578). Guilford Publications.

Kim, H. G., Cheon, E. J., Bai, D. S., Lee, Y. H., & Koo, B. H. (2018). Stress and heart rate variability: a meta-analysis and review of the literature. *Psychiatry Investigation*, *15*(3), 235. https://doi.org/10.30773/pi.2017.08.17

Kisker, J., Gruber, T., & Schöne, B. (2021a). Experiences in virtual reality entail different processes of retrieval as opposed to conventional laboratory settings: A study on human memory. *Current Psychology*, *40*, 3190–3197. https://doi.org/10.1007/s12144-019-00257-2

Kisker, J., Gruber, T., & Schöne, B. (2021b). Virtual reality experiences promote autobiographical retrieval mechanisms: Electrophysiological correlates of laboratory and virtual experiences. *Psychological Research*, *85*, 2485-2501. https://doi.org/10.1007/s00426-020-01417-x

Knowles, K. A., & Olatunji, B. O. (2020). Specificity of trait anxiety in anxiety and depression: Meta-analysis of the State-Trait Anxiety Inventory. *Clinical Psychology Review*, *82*, 101928. https://doi.org/10.1016/j.cpr.2020.101928

Kotov, R., Gamez, W., Schmidt, F., & Watson, D. (2010). Linking "big" personality traits to anxiety, depressive, and substance use disorders: A meta-analysis. *Psychological Bulletin*, *136*(5), 768-821. https://doi.org/10.1037/a0020327

Kourtesis, P., Collina, S., Doumas, L. A., & MacPherson, S. E. (2019). Validation of the virtual reality neuroscience questionnaire: maximum duration of immersive virtual reality sessions without the presence of pertinent adverse symptomatology. *Frontiers in Human Neuroscience*, *13*, 417. https://doi.org/10.3389/fnhum.2019.00417

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology*, *84*(3), 394-421. https://doi.org/10.1016/j.biopsycho.2010.03.010

Krokos, E., Plaisant, C., & Varshney, A. (2019). Virtual memory palaces: immersion aids recall. *Virtual Reality*, *23*(1), 1-15. ttps://doi.org/10.1007/s10055-018-0346-3

Kuusela, T. (2013). *Methodological aspects of heart rate variability analysis*. In: M. V. Kamath, M. A. Watanabe, & A. R. M. Upton (Eds.). Heart rate variability (HRV) signal analysis: Clinical applications, (pp.9-42). Boca Raton, FL: CRC Press

Laborde, S., Mosley, E., & Thayer, J. F. (2017). Heart rate variability and cardiac vagal tone in psychophysiological research–recommendations for experiment planning, data analysis, and data reporting. *Frontiers in Psychology*, *8*, 213. https://doi.org/10.3389/fpsyg.2017.00213

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping* (Vol. 464). New York: Springer.

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, *60*, 225-236. https://doi.org/10.1016/j.learninstruc.2017.12.007

Mammarella, N., Di Domenico, A., Palumbo, R., & Fairfield, B. (2016). When green is positive and red is negative: Aging and the influence of color on emotional memories. *Psychology and Aging*, *31*(8), 914. https://doi.org/10.1037/pag0000122

Mania, K., & Chalmers, A. (2001). The effects of levels of immersion on memory and presence in virtual environments: A reality centered approach. *Cyberpsychology & Behavior*, *4*(2), 247-264. https://doi.org/10.1089/109493101300117938

Mania, K., Troscianko, T., Hawkes, R., & Chalmers, A. (2003). Fidelity metrics for virtual environment simulations based on spatial memory awareness states. *Presence*, *12*(3), 296–310. https://doi.org/10.1162/105474603765879549

Maples-Keller, J. L., Bunnell, B. E., Kim, S. J., & Rothbaum, B. O. (2017). The use of virtual reality technology in the treatment of anxiety and other psychiatric disorders. *Harvard Review of Psychiatry*, *25*(3), 103-113. https://doi.org/10.1097/HRP.0000000000000138

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*(2), 175-215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

McLay, R. N., Baird, A., Webb-Murphy, J., Deal, W., Tran, L., Anson, H., ... & Johnston, S. (2017). A randomized, head-to-head study of virtual reality exposure therapy for posttraumatic stress disorder. *Cyberpsychology, Behavior, and Social Networking*, *20*(4), 218-224. https://doi.org/10.1089/cyber.2016.0554

Mourkoussis, N., Rivera, F. M., Troscianko, T., Dixon, T., Hawkes, R., & Mania, K. (2010). Quantifying fidelity for virtual environment simulations employing memory schema assumptions. *ACM Transactions on Applied Perception (TAP)*, *8*(1), 1–21. https://doi.org/10.1145/1857893.1857895

Neria, Y., Nandi, A., & Galea, S. (2008). Post-traumatic stress disorder following disasters: a systematic review. *Psychological Medicine*, *38*(4), 467-480. https://doi.org/10.1017/S0033291707001353

Norman, D. G., Wade, K. A., Williams, M. A., & Watson, D. G. (2020). Caught virtually lying—Crime scenes in virtual reality help to expose suspects' concealed recognition. *Journal of Applied Research in Memory and Cognition*, *9*(1), 118-127. https://doi.org/10.1016/j.jarmac.2019.12.008

North, M. M., North, S. M., & Coble, J. K. (1998). Virtual reality therapy: an effective treatment for phobias. *Virtual Environments in Clinical Psychology and Neuroscience*, 112-119. https://doi.org/10.3233/978-1-60750-902-8-112

Pallavicini, F., Pepe, A., & Minissi, M. E. (2019). Gaming in virtual reality: What changes in terms of usability, emotional response and sense of presence compared to non-immersive video games?. *Simulation & Gaming*, *50*(2), 136-159. https://doi.org/10.1177/1046878119831420

Palumbo, R., Di Domenico, A., & Mammarella, N. (2023). Aging and mixed emotions: A word-suffix approach in free recall. *Behavioral Sciences*, *13*(2), 160. https://doi.org/10.3390/bs13020160

Parsons, T. D., & Rizzo, A. A. (2008). Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry*, *39*(3), 250-261. https://doi.org/10.1016/j.jbtep.2007.07.007

Pedrabissi, L., & Santinello, M. (1989). Verifica della validità dello STAI forma Y di Spielberger [Verification of the validity of the STAI, Form Y, by Spielberger]. Giunti Organizzazioni Speciali, 191-192, 11-14

Penolazzi, B., Di Domenico, A., Marzoli, D., Mammarella, N., Fairfield, B., Franciotti, R. (2010). Effects of transcranial direct current stimulation on episodic memory related to emotional visual stimuli. *PloS one*, *5*(5), e10623. https://doi.org/10.1371/journal.pone.0010623

Perugini, M., Gallucci, M., & Costantini, G. (2018). A practical primer to power analysis for simple experimental designs. *International Review of Social Psychology*, *31*(1), 1-23. https://doi.org/10.5334/irsp.181

Pham, T., Lau, Z. J., Chen, S. A., & Makowski, D. (2021). Heart rate variability in psychology: A review of HRV indices and an analysis tutorial. *Sensors*, *21*(12), 3998. https://doi.org/10.3390/s21123998

Pizzie, R. G., & Kraemer, D. J. (2021). The association between emotion regulation, physiological arousal, and performance in math anxiety. *Frontiers in Psychology*, *12*, 639448. https://doi.org/10.3389/fpsyg.2021.639448

Puechlong, C., Weiss, K., Le Vigouroux, S., & Charbonnier, E. (2020). Role of personality traits and cognitive emotion regulation strategies in symptoms of post-

traumatic stress disorder among flood victims. *International Journal of Disaster Risk Reduction*, *50*, 101688. https://doi.org/10.1016/j.ijdrr.2020.101688

R Core Team (2021) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, *147*, 103778. https://doi.org/10.1016/j.compedu.2019.103778

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, *41*(1), 203-212. https://doi.org/10.1016/j.jrp.2006.02.001

Rauch, S. A., Eftekhari, A., & Ruzek, J. I. (2012). Review of exposure therapy: a gold standard for PTSD treatment. *Journal of Rehabilitation Research and Development*, *49*(5), 679-687. https://doi.org/10.1682/JRRD.2011.08.0152

Reece, J. D., Bunn, J. A., Choi, M., & Navalta, J. W. (2021). Assessing heart rate using consumer technology association standards. *Technologies*, *9*(3), 46. https://doi.org/10.3390/technologies9030046

Reger, G. M., Koenen-Woods, P., Zetocha, K., Smolenski, D. J., Holloway, K. M., Rothbaum, B. O., ... & Gahm, G. A. (2016). Randomized controlled trial of prolonged exposure using imaginal exposure vs. virtual reality exposure in active duty soldiers with deployment-related posttraumatic stress disorder (PTSD). *Journal of Consulting and Clinical Psychology*, *84*(11), 946. https://doi.org/10.1037/ccp0000134

Reger, G. M., Smolenski, D., Norr, A., Katz, A., Buck, B., & Rothbaum, B. O. (2019). Does virtual reality increase emotional engagement during exposure for PTSD? Subjective distress during prolonged and virtual reality exposure therapy. *Journal of Anxiety Disorders*, *61*, 75-81. https://doi.org/10.1016/j.janxdis.2018.06.001

Rhodes, S., Greene, N. R., & Naveh-Benjamin, M. (2019). Age-related differences in recall and recognition: A meta-analysis. *Psychonomic Bulletin & Review*, *26*, 1529–1547. https://doi.org/10.3758/s13423-019-01649-y

Richardson, J. T. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, *6*(2), 135-147. https://doi.org/10.1016/j.edurev.2010.12.001

Richardson, C. M. (2017). Emotion regulation in the context of daily stress: Impact on daily affect. *Personality and Individual Differences*, *112*, 150-156. https://doi.org/10.1016/j.paid.2017.02.058

Riva, G., Di Lernia, D., Sajno, E., Sansoni, M., Bartolotta, S., Serino, S., Gaggioli, A., & Wiederhold, B. K. (2021). Virtual reality therapy in the metaverse: Merging VR for the outside with VR for the inside. *Annual Review of Cybertherapy & Telemedicine*, *19*, 3-8.

Rizzo, A., Cukor, J., Gerardi, M., Alley, S., Reist, C., Roy, M., ... & Difede, J. (2015). Virtual reality exposure for PTSD due to military combat and terrorist attacks. *Journal of Contemporary Psychotherapy*, *45*, 255-264. https://doi.org/10.1007/s10879-015-9306-3

Rizzo, A., Reger, G., Gahm, G., Difede, J., & Rothbaum, B. O. (2009). *Virtual Reality Exposure Therapy for Combat-Related PTSD.* In: J. LeDoux, T. Keane, P. Shiromani. (eds) Post-Traumatic Stress Disorder, pp. 375-399, Humana Press. https://doi.org/10.1007/978-1-60327-329-9_18

Rothbaum, B. O., Hodges, L., Alarcon, R., Ready, D., Shahar, F., Graap, K., ... & Baltzell, D. (1999). Virtual reality exposure therapy for PTSD Vietnam veterans: A case study. *Journal of Traumatic Stress: Official Publication of The International Society for Traumatic Stress Studies*, *12*(2), 263-271. https://doi.org/10.1023/A:1024772308758

Rothbaum, B. O., Hodges, L. F., Ready, D., Graap, K., & Alarcon, R. D. (2001). Virtual reality exposure therapy for Vietnam veterans with posttraumatic stress disorder. *Journal of Clinical Psychiatry*, *62*(8), 617-622.

Šalkevičius, J., Miškinytė, A., & Navickas, L. (2019). Cloud based virtual reality exposure therapy service for public speaking anxiety. *Information*, *10*(2), 62. https://doi.org/10.3390/info10020062

Schöne, B., Kisker, J., Sylvester, R. S., Radtke, E. L., & Gruber, T. (2021). Library for universal virtual reality experiments (luVRe): A standardized immersive 3D/360 picture and video database for VR based research. *Current Psychology*, *42*(7), 5366-5384. https://doi.org/10.1007/s12144-021-01841-1

Schöne, B., Wessels, M., & Gruber, T. (2017). Differences between real-life and laboratory memory: Evidence from EEG and virtual reality. *Psychophysiology*, *54*, S128–S128.

Seligowski, A. V., Lee, D. J., Bardeen, J. R., & Orcutt, H. K. (2015). Emotion regulation and posttraumatic stress symptoms: A meta-analysis. *Cognitive Behaviour Therapy*, *44*(2), 87-102. https://doi.org/10.1080/16506073.2014.980753

Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends in Cognitive Sciences*, *17*(11), 565-573. https://doi.org/10.1016/j.tics.2013.09.007

Shaffer, J. P. (1995). Multiple hypothesis testing. *Annual Review of Psychology*, *46*(1), 561-584. https://doi.org/10.1146/annurev.ps.46.020195.003021

Shaffer, F., & Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Frontiers in Public Health*, *5*, 258. https://doi.org/10.3389/fpubh.2017.00258

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3549-3557. https://doi.org/10.1098/rstb.2009.0138

Slater, M., Linakis, V., Usoh, M., & Kooper, R. (1996, July). Immersion, presence and performance in virtual environments: An experiment with tri-dimensional

chess. In Proceedings of the ACM symposium on virtual reality software and technology (pp. 163-172). https://doi.org/10.1145/3304181.3304216

Smith, S. A. (2019). Virtual reality in episodic memory research: A review. *Psychonomic Bulletin & Review*, *26*(4), 1213–1237. https://doi.org/10.3758/s13423-019-01605-w

Smith, S. M. (1979). Remembering in and out of context. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(5), 460. https://doi.org/10.1037/0278-7393.5.5.460

Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review*, *8*(2), 203–220. https://doi.org/10.3758/BF03196157

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory (STAI)(Form Y1–Y2)*. Palo Alto, CA: Consulting Psychologists Press.

Spinnler, H. (1987). Standardizzazione e taratura italiana dei test neuropsicologici. *Italian Journal of Neurological Sciences*, *6*, 8-20.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149. https://doi.org/10.3758/BF03207704

Sutcliffe, A., Gault, B., & Shin, J. E. (2005). Presence, memory and interaction in virtual environments. *International Journal of Human-Computer Studies*, *62*(3), 307-327. https://doi.org/10.1016/j.ijhcs.2004.11.010

Tarrier, N., Sommerfield, C., Pilgrim, H., & Faragher, B. (2000). Factors associated with outcome of cognitive-behavioural treatment of chronic post-traumatic stress disorder. *Behaviour Research and Therapy*, *38*(2), 191-202. https://doi.org/10.1016/S0005-7967(99)00030-3

Tarvainen, M. P., & Niskanen, J. P. (2008). Kubios HRV user's guide. Biosignal Analysis and Medical Imaging Group (BSAMIG) (http://bsamig. uku. fi/), Department of Physics, University of Kuopio, Kuopio, Finland.

Taylor, D. A., & Dando, C. J. (2018). Eyewitness memory in face-to-face and immersive avatar-to-avatar contexts. *Frontiers in Psychology*, *9*, 507. https://doi.org/10.3389/fpsyg.2018.00507

Tian, F., & Zhang, W. (2021). The difference of emotional arousal between traditional 2D and VR movies: A comparative study based on EEG signals. In *2021 6th International Conference on Intelligent Computing and Signal Processing* (ICSP) (pp. 734–737). IEEE. https://doi.org/10.1109/ICSP51882.2021.9408991

United Nations Office for Disaster Risk Reduction. (2022, October 14) 2021 Global Disaster Assessment Report. https://www.preventionweb.net/publication/2021-global-disaster-assessment-report

Usui, H., & Nishida, Y. (2017). The very low-frequency band of heart rate variability represents the slow recovery component after a mental stress task. *PloS one*, *12*(8), e0182611. https://doi.org/10.1371/journal.pone.0182611

van Helvoort, D., Stobbe, E., Benning, R., Otgaar, H., & van de Ven, V. (2020). Physical exploration of a virtual reality environment: Effects on spatiotemporal associative recognition of episodic memory. *Memory & Cognition*, *48*(5), 691–703. https://doi.org/10.3758/s13421-020-01024-6

van Minnen, A., Arntz, A., & Keijsers, G. P. J. (2002). Prolonged exposure in patients with chronic PTSD: Predictors of treatment outcome and dropout. *Behaviour Research and Therapy*, *40*(4), 439-457. https://doi.org/10.1016/S0005-7967(01)00024-9

Vanderhasselt, M. A., Koster, E. H., Onraedt, T., Bruyneel, L., Goubert, L., & De Raedt, R. (2014). Adaptive cognitive emotion regulation moderates the relationship between dysfunctional attitudes and depressive symptoms during a stressful life period: A prospective study. *Journal of Behavior Therapy and Experimental Psychiatry*, *45*(2), 291-296. https://doi.org/10.1016/j.jbtep.2014.01.003

Vinatzer, H., Rzepka, A., Hayn, D., Ziegl, A., & Schreier, G. (2022, July). Investigation of the time shift between wearable photoplethysmography sensors used for continuous heart rate monitoring. In 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) (pp. 4308-4311). IEEE. https://doi.org/10.1109/EMBC48229.2022.9871629

Vogt, S., & Magnussen, S. (2007). Long-term memory for 400 pictures on a common theme. *Experimental Psychology*, *54*(4), 298. https://doi.org/10.1027/1618-3169.54.4.298

Wagler, A., & Hanus, M. D. (2018). Comparing virtual reality tourism to real-life experience: Effects of presence and engagement on attitude and enjoyment. *Communication Research Reports*, *35*(5), 456-464. https://doi.org/10.1080/08824096.2018.1525350

Watier, N., & Collin, C. (2012). The effects of distinctiveness on memory and metamemory for face–name associations. *Memory*, *20*(1), 73–88. https://doi.org/10.1080/09658211.2011.637935

Weinberg, M., Gil, S., Besser, A., & Bass, J. (2021). Personality traits and trauma exposure: The relationship between personality traits, PTSD symptoms, stress, and negative affect following exposure to traumatic cues. *Personality and Individual Differences*, *177*, 110802.https://doi.org/10.1016/j.paid.2021.110802

Weiss, D.S. (2007). The Impact of Event Scale: Revised. In: Wilson, J.P., Tang, C.Sk. (eds) *Cross-Cultural Assessment of Psychological Trauma and PTSD*. International and Cultural Psychology Series. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-70990-1_10

Wilson, B. A., Cockburn, J., & Baddeley, A. D. (1985). *The Rivermead behavioural memory test*. Thames Valley Test Company.

Wolpe, J. (1990). *The practice of behavior therapy*. Pergamon press.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*(3), 441–517. https://doi.org/10.1006/jmla.2002.2864

Yonelinas, A. P., Aly, M., Wang, W. C., & Koen, J. D. (2010). Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus*, *20*(11), 1178–1194. https://doi.org/10.1002/hipo.20864

Zlotnik, G., & Vansintjan, A. (2019). Memory: An extended definition. *Frontiers in Psychology*, *10*, 2523. https://doi.org/10.3389/fpsyg.2019.02523

Supplementary Materials

Table S1

Objects' List, Rating and Classification on the Three Dimensions Rated

		Valence		Arousal		Typicity	
code	**description**	**Mean**	**Category**	**Mean**	**Category**	**Mean**	**Category**
1	Crushed paper	3.54	LOW	5.64	HIGH	6.41	HIGH
2	Dirty tissue	3.37	LOW	5.33	HIGH	5.12	HIGH
3	Broken pencil	2.98	LOW	6.30	HIGH	5.06	HIGH
4	Ripped book	3.02	LOW	6.15	HIGH	5.00	HIGH
5	Dying plant	3.01	LOW	4.72	LOW	3.27	LOW
6	Full ashtray	2.26	LOW	6.01	HIGH	3.46	LOW
7	Dirty-water bottle	2.22	LOW	5.77	HIGH	1.81	LOW
8	Broken chair	2.15	LOW	6.33	HIGH	3.61	LOW
9	Tossed surgical mask	3.04	LOW	5.78	HIGH	3.93	LOW
10	Spilling waste basket	3.40	LOW	5.25	HIGH	7.26	HIGH
11	Terroristic attack painting	2.28	LOW	7.15	HIGH	1.93	LOW
12	Book	7.14	HIGH	3.54	LOW	8.48	HIGH
13	Coffee maker	7.25	HIGH	3.95	LOW	6.47	HIGH
14	Pen	6.80	HIGH	3.54	LOW	8.69	HIGH
15	Chair	6.72	HIGH	3.09	LOW	8.50	HIGH
16	Mini cooler	6.64	HIGH	3.59	LOW	3.78	LOW
17	Tissue box	5.62	HIGH	4.09	LOW	6.17	HIGH
18	Computer and mouse	6.97	HIGH	3.70	LOW	8.22	HIGH
19	Mug	6.37	HIGH	3.15	LOW	5.55	HIGH
20	Sunglasses	6.09	HIGH	3.74	LOW	3.11	LOW
21	Glass	5.54	HIGH	4.09	LOW	6.19	HIGH
22	City painting	5.02	HIGH	5.07	HIGH	2.93	LOW
23	Flowers in a vase	7.12	HIGH	2.86	LOW	4.93	LOW
24	Fruit juice	5.92	HIGH	3.89	LOW	3.81	LOW
25	Banknote	6.41	HIGH	4.81	LOW	2.38	LOW
26	Cookies' box	6.67	HIGH	3.53	LOW	3.94	LOW
27	Candies in a bowl	7.14	HIGH	3.30	LOW	5.17	HIGH
28	Star-shaped paperweight	6.28	HIGH	3.49	LOW	5.89	HIGH
29	Dreamcatcher	6.78	HIGH	3.01	LOW	2.51	LOW
30	Trumpet	5.49	HIGH	5.01	HIGH	2.09	LOW
31	Rubik's cube	5.93	HIGH	4.71	LOW	3.16	LOW
32	Bonbonnière	7.22	HIGH	3.92	LOW	4.12	LOW
33	Island picture	7.42	HIGH	2.18	LOW	3.85	LOW
34	Closet	5.73	HIGH	3.93	LOW	5.38	HIGH
35	Furniture with shelves	5.54	HIGH	4.09	LOW	6.35	HIGH
36	Window	6.12	HIGH	3.61	LOW	7.43	HIGH
37	Air conditioner	7.11	HIGH	3.17	LOW	6.70	HIGH

38	Shelves	6.20	HIGH	3.59	LOW	7.17	HIGH
39	Radiator	7.27	HIGH	2.68	LOW	7.74	HIGH
40	Printer	6.63	HIGH	3.52	LOW	7.93	HIGH
41	Dresser	6.23	HIGH	3.55	LOW	7.59	HIGH
42	Whiteboard	6.31	HIGH	4.31	LOW	6.74	HIGH
43	Telephone	6.48	HIGH	4.32	LOW	6.99	HIGH
44	Desk	6.80	HIGH	3.51	LOW	8.42	HIGH
45	Calendar	6.11	HIGH	4.16	LOW	7.55	HIGH
46	Coat hanger	6.20	HIGH	3.49	LOW	6.80	HIGH
47	Two-door furniture	5.50	HIGH	3.91	LOW	6.06	HIGH
48	Door	6.48	HIGH	3.97	LOW	8.31	HIGH
49	Emergency light	5.82	HIGH	4.60	LOW	6.42	HIGH
50	Socket	6.09	HIGH	4.18	LOW	8.15	HIGH
51	Writing “The Boss”	4.42	LOW	5.05	HIGH	2.07	LOW
52	Succulent plant	6.02	HIGH	3.42	LOW	5.30	HIGH
53	Graduation certificates (on the wall)	7.57	HIGH	3.99	LOW	7.36	HIGH

Figure S1

Pictures Used in the 2D Condition



Table S2

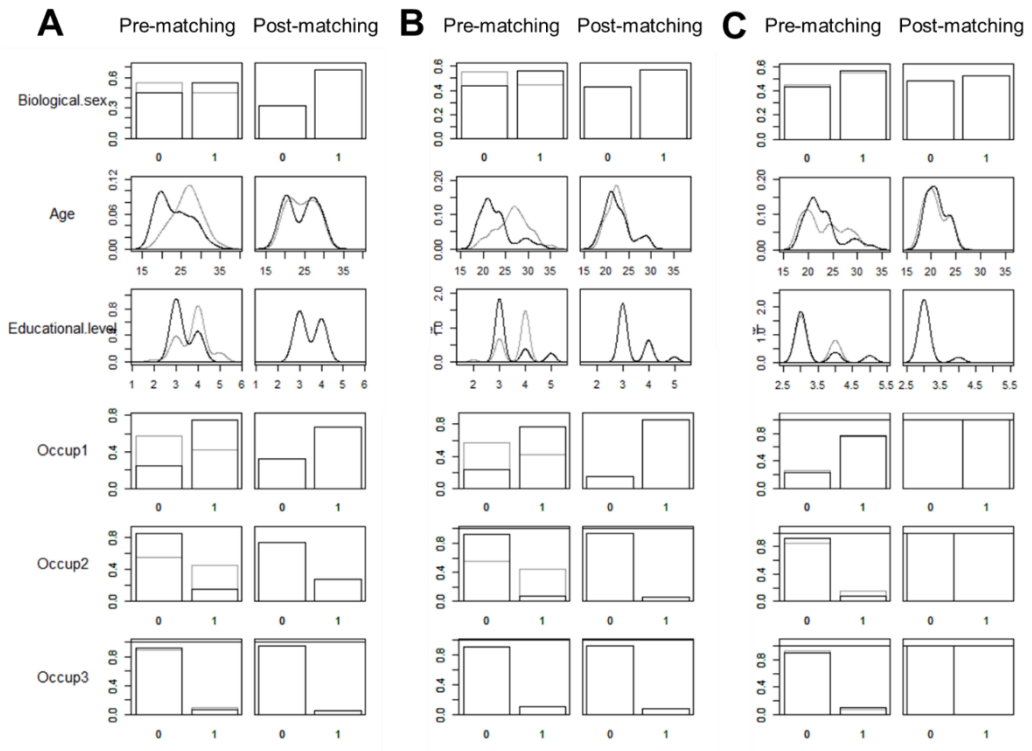
Descriptive Statistics for Pre- and Post-Matching Samples

Variables	Pre-matching			Post-matching		
	Means* Treated	Means* Control	Std. Mean Diff.	Means* Treated	Means* Control	Std. Mean Diff.
G2vsG1						
Biological sex	0.55	0.45	0.20	0.68	0.68	0.00
Age	23.25	26.53	-0.84	24.21	24.32	-0.03
Educational level	3.33	3.78	-0.78	3.46	3.46	0.00
Occupation (Unemployed)	0.75	0.43	0.70	0.68	0.68	0.00
Occupation (Student)	0.15	0.45	-0.69	0.27	0.27	0.00
Occupation (Employed)	0.08	0.10	-0.09	0.05	0.05	0.00
G3vsG1						
Biological sex	0.56	0.45	0.23	0.57	0.57	0.00
Age	22.79	26.53	-1.04	22.39	22.81	-0.12
Educational level	3.36	3.78	-0.63	3.37	3.37	0.00
Occupation (Unemployed)	0.77	0.43	0.75	0.86	0.86	0.00
Occupation (Student)	0.08	0.45	-0.93	0.06	0.06	0.00
Occupation (Employed)	0.10	0.10	0.01	0.09	0.09	0.00
G3vsG2						
Biological sex	0.56	0.55	0.03	0.52	0.52	0.00
Age	22.79	23.25	-0.12	20.99	20.77	0.06
Educational level	3.36	3.33	0.06	3.08	3.08	0.00
Occupation (Unemployed)	0.77	0.75	0.05	1.00	1.00	0.00
Occupation (Student)	0.08	0.15	-0.23	0.00	0.00	0.00
Occupation (Employed)	0.10	0.08	0.10	0.00	0.00	0.00

* Proportions were used when variables are categorical

Figure S2

Density plots of the covariates before (pre-) and after (post-) matching procedure



List of Ph.D. Student Papers

Journal Articles, Books and Book Chapters

Mazza, C., Ceccato, I., Cannito, L., Monaro, M., Ricci, E., Bartolini, E., Cardinale, A., Di Crosta, A., Cardaioli, M., La Malva, P., **Colasanti, M.**, Tambelli, R., Giromini, L., Palumbo, R., Di Domenico, A., Roma, P. (2024). A Step Forward in Identifying Socially Desirable Respondents: An Integrated Machine Learning Model Considering T-Scores, Response Time, Kinematic Indicators, and Eye Movements. Human Behavior and Emerging Technologies, 2024(1).

Marchitelli, S., Mazza, C., Ricci, E., Faia, V., Biondi, S., **Colasanti, M.**, Cardinale, A., Roma, P., & Tambelli, R. (2024). Identification of Psychological Treatment Dropout Predictors Using Machine Learning Models on Italian Patients Living with Overweight and Obesity Ineligible for Bariatric Surgery. *Nutrients*, *16*(16), 2605.

Monaro, M., Mazza, C., **Colasanti, M.**, Colicino, E., Bosco, F., Ricci, E., Biondi, S., Rossi, M., & Roma, P. (2024). Testing memory of a VR environment: comparison with the real environment and 2D pictures. *Virtual Reality*, *28*(2), 1-12.

Ceccato, I., Ricci, E., Mazza, C., Bartolini, E., Di Crosta, A., La Malva, P., Biondi, S., **Colasanti, M.**, Mammarella, N., Palumbo, R., Roma, P., & Di Domenico, A. (2024). Influence of stimuli emotional features and typicality on memory performance: insights from a virtual reality context. *Psychological Research*, *88*(1), 257-270.

Colasanti, M., Melis, G., Monaro, M., Ricci, E., Bosco, F., Rossi, M., Biondi, S., Verrocchio, M. C., Di Domenico, A., Mazza, C., & Roma, P. (2024). Did You Commit a Crime There? Investigating the Visual Exploration Patterns of Guilty, Innocent, Honest, and Dishonest Subjects When Viewing a Complex Mock Crime Scene. *Journal of Nonverbal Behavior*, 1-25.

Ricci, E., **Colasanti, M.**, Monaro, M., Mazza, C., Cardinale, A., Bosco, F., Mazzoni, G., Rossi Arnaud, C., Ferracuti, S., Di Domenico, A., & Roma, P. (2023). How to Distinguish Feigned from Genuine Depressive Symptoms: Response

Patterns and Content Analysis of the SIMS Affective Disorder Scale. *Psychological Injury and Law*, *16*(3), 237-248.

Roma, P., **Colasanti, M.**, Ricci, E., Biondi, S., & Mazza, C. (2023). Extreme Violence: Homicides in Europe from an Italian Perspective. In Martin, C.R., Preedy, V.R., Patel, V.B. (eds), *Handbook of Anger, Aggression, and Violence* (pp. 2457-2481). Springer, Cham.

Marchetti, D., Maiella, R., Palumbo, R., D'Ettorre, M., Ceccato, I., **Colasanti, M.**, Di Crosta, A., La Malva, P., Bartolini, E., Biasone, D., Mammarella, N., Porcelli, P., Di Domenico, A., & Verrocchio, M. C. (2023). Self-Reported Mental Health and Psychosocial Correlates during the COVID-19 Pandemic: Data from the General Population in Italy. *Data*, *8*(6), 111.

Colasanti, M., Ricci, E., Cardinale, A., Amati, F., Mazza, C., Biondi, S., Ferracuti, S., & Roma, P. (2023). Homicide-Suicide in Italy Between 2009-2018: An Epidemiological Update and Time Series Analysis. *European Journal on Criminal Policy and Research*, 1-16.

Marchetti, D., Fontanesi, L., Camisasca, E., **Colasanti, M.**, Covelli, V., Di Giandomenico, S., Miragoli, S., & Verrocchio, M. C. (2023). The Role of Pre-Pandemic Mental Health Status and Personality Traits on Psychological Distress During the COVID-19 Lockdown Among Italian Young Adults. *Behavioral Sciences*, *13*(2), 141.

Colasanti, M. (2023). La Realtà Virtuale può Essere Utile per il Trattamento del Disturbo Post Traumatico da Stress. In M. C. Verrocchio (Ed.), *Manuale di Psicologia Clinica* (p. 309). Piccin.

Roma, P., **Colasanti, M.**, Rossi, M., Biondi, S., & Mazza, C. (2023). Trauma e Disturbi Correlati. In M. C. Verrocchio (Ed.), *Manuale di Psicologia Clinica* (pp. 295-308). Piccin.

Mazza, C., Ricci, E., **Colasanti, M.**, Cardinale, A., Bosco, F., Biondi, S., Tambelli, R., Di Domenico, A., Verrocchio, M. C., & Roma, P. (2022). How Has COVID-19 Affected Mental Health and Lifestyle Behaviors After 2 Years? The Third Step of a Longitudinal Study of Italian Citizens. *International Journal of Environmental Research and Public Health*, *20*(1), 759.

Mazza, C., Formica, F., Ferracuti, S., Ricci, E., **Colasanti, M.**, Biondi, S., Di Domenico, A., & Roma, P. (2022). High expressed emotion (HEE), assessed using the five-minute speech sample (FMSS), as a predictor of psychiatric relapse in patients with schizophrenia and major depressive disorder: a meta-analysis and meta-regression. *Journal of Clinical Medicine*, *11*(21), 6533.

Conference Proceedings

Mazza, C., Tibermachine, I.E., Monaro, M., Bosco, F., **Colasanti, M.**, Biondi, S., Napoli, C., Roma, P. (2024). *On the Side of the Child: Impact of an Immersive Virtual Reality (VR) Experience on Parental Reflective Functioning and Empathy*. XXIV Congresso Nazionale Associazione Italiana di Psicologia della Sezione di Psicologia Clinica e Dinamica, Salerno (IT), 12-15/09/2024.

Bosco, F., **Colasanti, M.** (2024). *Beyond Self-Report Assessment: Evaluating Socially Desirable Responding at an Implicit Level*. Annual Conference of the European Association of Psychology and Law, Caparica (PT), 09-12/07/2024

Colasanti, M., de Roos, M., Biondi, S., van Dongen, J. (2024). *Cyber Dating Abuse among Dutch and Italian Young Adults*. Annual Conference of the European Association of Psychology and Law, Caparica (PT), 09-12/07/2024.

Ricci, E., Bosco, F., **Colasanti, M.**, Biondi, S., Mazza, C., Roma, P. (2024). *Socially Desirable Responding: New Insights for Psychological Assessment*. Society of Personality Assessment 2024 Convention, San Diego (US), 13-17/03/2024.

Mazza, C., Ceccato, I., Cannito, L., Monaro, M., Ricci, E., Bartolini, E., **Colasanti, M.**, Biondi, S., Bosco, F., Di Crosta, A., Cardaioli, M., Cardinale, A., La Malva, P., Palumbo, R., Giromini, L., Palumbo, R., Di Domenico, A., Roma, P. (2024). *An Integrated Machine Learning Model Considering T-Scores, Response Time, Kinematic Indicators, and Eye Movements in Identifying Socially Desirable Responding*. Society of Personality Assessment 2024 Convention, San Diego (US), 13-17/03/2024.

Mazza, C., **Colasanti, M.**, Ricci, E., Ferracuti, S., Roma, P. (2023). *Assessment of Criminal Responsibility Using the Shedler-Westen Assessment Procedure-200 (SWAP-200): A Clinical-Forensic Proposal*. 1° Congresso Internazionale su Decision Making in Medicine and Law Opportunità e Rischi delle Tecnologie Informatiche. Roma (IT), 14-15/12/2023

Mazza, C., Ceccato, I., Cannito, L., Monaro, M., Ricci, E., Bartolini, E., **Colasanti, M.**, Di Crosta, A., Cardaioli, M., Cardinale, A., La Malva, P., Palumbo, R., Giromini, L., Palumbo, R., Di Domenico, A., Roma, P. (2023). *An Integrated Machine Learning Model Considering T-Scores, Response Time, Kinematic Indicators, and Eye Movements in Identifying Socially Desirable Responding*. New Horizons in Psychological Assessment, European Federation of Psychologists' Associations, Padova (IT), Virtual Presentation, 22-23/09/2023.

Colasanti, M., Ricci, E., Mazza, C. (2023). *How to Assess the Credibility of a Patient With Depression? An Empirical Decision-Making Strategy in Interpreting Test Results*. XXIII Congresso Nazionale Associazione Italiana di Psicologia della Sezione di Psicologia Clinica e Dinamica, Firenze (IT), 14-17/09/2023.

Colasanti, M., Facchino, A.P., Verrocchio, M. C. (2023). *Can Serious Gaming Improve Psychologists' Professional Training? A Systematic Review*. XXIII Congresso Nazionale Associazione Italiana di Psicologia della Sezione di Psicologia Clinica e Dinamica, Firenze (IT), 14-17/09/2023.

Colasanti, M., Ricci, E., Mazza, C., Bosco, F., Monaro, M., Ferracuti, S., Verrocchio, M. C., Roma, P. (2023). *Does SIMS Affective Disorder (AF) Scale Actually Measure What it States it Measures? An Empirical Investigation To Distinguish Feigned Depression from Genuine Patients*. Annual Conference of the European Association of Psychology and Law, Cluj-Napoca (RO), 04-07/07/2023

Bosco, F., Biondi, S., Ricci, E., **Colasanti, M.**, Mazza, C., Melis, G., Monaro, M., Di Domenico, A., Roma, P. (2023). *Did you Commit a Crime There? Investigating the Visual Exploration Patterns When Viewing a Complex Mock Crime Scene: A Comparison Between Guilty Vs. Innocent and Honest Vs. Dishonest Participants*. Annual Conference of the European Association of Psychology and Law, Cluj-Napoca (RO), 04-07/07/2023

Mazza, C., Spinetti, E., Barchielli, B., Yared, S., Bosco, F., Ricci, E., **Colasanti, M.**, Biondi, S., Napoli, C., Roma, P. (2023). *On the Side of the Child: Impact of an Immersive Virtual Reality (VR) Experience on Parental Reflective Function*. International Congress on Mentalization Based Treatment for Children, Adolescents and Families, Barcelona (ES), 29-30/06/2023

Roma, P., **Colasanti, M.**, Biondi, S. (2022). *Forensic Psychological Assessment: Which Credibility?*. XXX Congresso Nazionale Associazione Italiana di Psicologia, Padova (IT), 26-30/09/2022.

Roma, P., **Colasanti, M.**, Biondi, S. (2022). *The Role of Virtual Reality in the Treatment of Sex Offenders and Perpetrators of Intimate Partner Violence (IPV): Possible Applications*. XXX Congresso Nazionale Associazione Italiana di Psicologia, Padova (IT), 26-30/09/2022.

Mazza, C., **Colasanti, M.**, Ricci, E., Di Domenico, A., Roma, P. (2022). *Behavioral Indicators to Detect Socially Desirable Responding*. Society of Personality Assessment 2022 Convention, Chicago (US), Virtual Presentation, 09-13/03/2022.