Computers in Human Behavior Reports 19 (2025) 100734

Contents lists available at ScienceDirect

Computers in Human Behavior Reports

journal homepage: www.sciencedirect.com/journal/computers-in-human-behavior-reports

Military stress management training using virtual reality: A systematic review and delphi method

Ana Vianez [a,*], António Marques [a,**], Javier Loureiro [b], Paulo Veloso Gomes [a]

[a] Psychosocial Rehabilitation Laboratory, Center for Rehabilitation Research, School of Health, Polytechnic Institute of Porto, Portugal
[b] CITIC Research Center, University of A Coruña, A Coruña, Spain

ARTICLE INFO

Keywords:
Stress management training
Virtual reality
Military
Special operations forces
Resilience

ABSTRACT

Military personnel encounter stressors that uniquely affect their performance and well-being, necessitating effective stress management interventions. This study investigates the efficacy of VR-based Stress Management Training (SMT) programs in comparison to traditional approaches. Through a systematic review of 13 studies, VR-based SMT programs were found to significantly reduce stress and anxiety symptoms, with various physiological markers used to evaluate the stress response. Additionally, a Delphi study was conducted to gather expert insights on designing VR-based SMT programs specifically for special operations forces. The experts identified key stressors and coping strategies relevant to this group, emphasizing the potential of VR for stress monitoring, simulation training, and real-time feedback. The Delphi study reached a 97 % consensus on the design principles for these programs. Overall, the findings suggest that integrating VR into SMT programs offers a promising strategy for enhancing stress management and resilience in military personnel.

1. Introduction

Military personnel perform highly demanding operational tasks and are routinely exposed to stressors as an inherent part of their professional lives. War represents one of the most challenging situations for soldiers, as the multifaceted demands of combat (physical, cognitive, and psychological) place enormous stress on even the most prepared military personnel (Pols & Oak, 2007).

The concept of stress has been a subject of scientific debate ever since its first use in physiological and biomedical research in 1950. Originally defined as the non-specific response of the body to any noxious stimuli, stress is now understood as an external load or demand on a biological, social, or psychological system. In this review, stressors refer to external combat-related demands (e.g., sleep deprivation), while stress responses denote psychophysiological reactions (e.g., cortisol release) (Koolhaas et al., 2011; Lazarus, 1993).

In military contexts, stress factors can manifest in various forms, including physical (e.g., dehydration) and psychological (e.g., risk of injury or death) stressors. The military work environment is markedly different from civilian settings, leading to unique stressors such as deployment to missions, sleep deprivation, and high levels as noted by Jex et al. (2013). Moreover, factors such as family reintegration, physical injuries, and mental health conditions further exacerbate stress among military personnel (Bacon & Staudenmeier, 2003; Jex et al., 2013).

The presence of these stress factors highlights the importance of resilience-building programs and Stress Management Training (SMT), not only to reduce the risk of Post-Traumatic Stress Disorder (PTSD) and other psychological injuries but also to maintain optimal operational effectiveness. Elevated stress and anxiety levels have been shown to impair mission readiness and cognitive-behavioral performance (Castro, Adler, & Britt, 2006). Operational stress affects information processing and significantly impacts performance in situations requiring emotional, cognitive, and behavioral control, such as military operations (Thompson & McCreary, 2006).

Kavanagh (2005) notes that while few individuals are entirely immune to the effects of stress on performance, there are intervening variables known as “moderators” that can mitigate performance decline. Similarly, Dolan and Adler (2006) found that targeted training can enhance an individual’s capacity to regulate emotional and physiological responses to stress, ultimately improving performance in high-risk, high-demand situations.

* Corresponding author.
** Corresponding author.
E-mail addresses: acvianez@gmail.com (A. Vianez), ajmarques@ess.ipp.pt (A. Marques), javier.pereira@udc.es (J. Loureiro), pvg@ess.ipp.pt (P.V. Gomes).

https://doi.org/10.1016/j.chbr.2025.100734
Received 23 January 2025; Received in revised form 27 May 2025; Accepted 22 June 2025
Available online 2 July 2025

SMT encompasses a set of techniques to improve stress coping mechanisms, including approaches like Stress Inoculation Training (SIT) and Mental Resilience Training (MRT). SIT programs aim to increase stress tolerance through exposure, while MRT programs teach stress management techniques, integrating cognitive-behavioral and mindfulness-based approaches to strengthen resilience (Robertson et al., 2015).

In the military field, SIT attempts to improve stress resistance by training a set of skills that make an individual capable of responding more favorably when confronted with stressors in real life. Through training, individuals may also learn how to manage uncertainty and maintain high levels of performance despite its presence. It is described as "a flexible individually tailored multifaceted form of cognitive-behavioral therapy" (Popovic, Horvat, Kukolja, Dropuljic, & Cosic, 2009; Port et al., 2001).

Survival, Evasion, Resistance, and Escape (SERE) is a military program that incorporates SIT techniques. This program prepares soldiers for deployment to the Theater of Operations in the future. During the training, participants acquire skills to fully understand the concept of "Joint Personnel Recovery". Additionally, they are trained in survival techniques, evasion in case of capture, and the military code of conduct. SERE training is a requirement for all military personnel determined to be at risk of capture (Doran et al., 2012).

MRT integrates cognitive-behavioral and mindfulness-based principles and techniques that have been proven to improve stress regulation, emotional and cognitive control, and self-confidence. The aim is to maintain and improve psychological resilience. Its core objective is to provide a formal Psychological Skills Training (PST) curriculum within military training contexts, enabling soldiers to understand how their brain and body respond under stress and to develop skills to stay present, focused, and effective in high-pressure situations (Stanley et al., 2011).

Despite the effectiveness of MRT and SIT techniques, their implementation in real-world settings faces challenges such as limited accessibility, variability in delivery, and lack of contextual realism. VR offers a novel solution to these limitations by providing immersive, interactive, and controlled simulations that replicate operational stress scenarios with high fidelity (Carl et al., 2019; Gaggioli et al., 2014; Pallavicini et al., 2013).

Reger et al. (2008) emphasize that VR is highly effective in simulating complex, high-stress scenarios typical of military operations, such as combat situations and high-pressure decision-making environments. They argue that VR enables exposure to stress-inducing stimuli within a safe and controlled setting, thereby reducing the risks associated with real-life exposure. Complementarily, Gaggioli et al. (2014) demonstrated that immersive VR scenarios combined with real-time physiological monitoring constitute a promising approach to managing psychological stress. Both studies highlight that VR provides high levels of immersion and real-time feedback mechanisms, features that enhance the effectiveness of training under the demanding and unpredictable conditions often encountered in military contexts.

In 2003, the North Atlantic Treaty Organization (NATO) released a noteworthy report titled "Virtual Reality: State of Military Research and Applications in Member Countries." This report outlines several advantageous applications of virtual reality, including assistance in conceptualizing, evaluating, and testing weapon systems, aiding in mission planning and rehearsal, and real-world mission applications (Rizzo et al., n.d.; Pair et al., 2006).

This report emerged from a decades-long evolution in military stress management. The scientific understanding of operational stress progressed from reactive post-combat interventions in World War II to preventive frameworks like the U.S. Army's Battlemind program (1983) and standardized SERE training (Doran et al., 2012; NATO's 2003 study crystallized the transition to technology-enhanced solutions, positioning VR as a pivotal tool for modern SMT implementation. Contemporary programs now demonstrate 25–35 % reductions in PTSD symptoms through immersive, biometric-augmented protocols (Binnendijk, 2002; Castro, Adler, & Britt, 2006).

However, although several VR-based stress management programs have been developed in military contexts, the current literature lacks a systematic synthesis of their effectiveness, validated implementation protocols, and consensus-based design guidelines. To date, no review has combined systematic analysis with expert consensus (via the Delphi method) to guide future implementation, particularly in elite units such as special operations forces.

While prior reviews have addressed conventional approaches to SIT and MRT, such as the SERE program, none have systematically compared the efficacy of VR-SMT against traditional methods (e.g., field-based SERE) in critical metrics such as stress biomarker reduction or skill retention under simulated combat conditions. Additionally, although Gaggioli et al. (2014) reviewed general applications of VR in military contexts, their analysis did not examine specific physiological stress markers or their relevance in assessing training efficacy (Binnendijk, 2002). These limitations underscore the need for a comprehensive synthesis that integrates both psychological and physiological dimensions to evaluate the impact of immersive VR-based interventions in high-demand military settings.

To address these gaps, this study pursues four objectives:

1. Evaluate the effectiveness of VR-based SMT in reducing stress among military personnel; 2. identify key physiological and psychological markers of stress that can be measured in VR environments; 3. Develop consensus-driven design principles for special operations forces through the Delphi method; 4. Address implementation challenges such as scenario standardization, evaluation biases, and ethical concerns.

By integrating systematic evidence and expert perspectives, this review aims to inform future development and practical deployment of VR-based SMT strategies in high-stakes military environments.

2. Methods

2.1. Study 1

To provide a comprehensive overview of the benefits of VR-SMT programs developed for military contexts, a systematic search was conducted across the following databases: PsycINFO, Web of Science (Web of Knowledge), PubMed, B-on, Cochrane Library, and Scopus. These databases were selected for their extensive coverage of psychology, medicine, and military science, ensuring a multidisciplinary perspective.

The search strategy combined the following terms: ("Stress Management Training" OR "Stress Control" OR "Stress Reduction") AND ("Virtual Reality" OR "VR" OR "Immersive Environment") AND ("Military" OR "Armed Forces" OR "Defense Personnel") AND ("Special Operations Forces" OR "Special Forces" OR "Elite Military Units") AND ("Resilience" OR "Mental Toughness" OR "Emotional Strength") using the Boolean operator AND to connect each conceptual group.

The timeframe (2007–2023) was chosen to reflect the emergence and growing application of immersive VR technology in military training environments, particularly following increased international interest in the use of simulation tools for psychological preparedness post-2007.

All references were managed using EndNote, and the review process followed the PRISMA 2020 guidelines. Two independent reviewers screened titles and abstracts, followed by a full-text review. Disagreements were resolved through discussion or adjudication by a third reviewer.

Primary outcomes of interest included changes in psychological stress and anxiety symptoms. Secondary outcomes, when reported, included physiological measures (e.g., heart rate variability), mission readiness, and resilience-related scores.

Inclusion criteria were: Empirical studies published between 2007 and 2023. Studies investigating the use of VR in SMT programs specifically designed for military training. Use of VRET or cyber-SIT (Stress Inoculation Training using VR) as either a primary intervention or complementary tool within evidence-based SMT protocols. Assessment of stress, anxiety, or related resilience indicators among military personnel.

Exclusion criteria were: Studies published in languages other than English. Non-empirical articles such as opinion papers, reviews, or theoretical frameworks. Studies that did not adopt an experimental or quasi-experimental design (e.g., lacked randomised controlled trial methodology).

A PRISMA flow diagram (Fig. 1) illustrates the article selection process, from identification to final inclusion.

A comprehensive search across scientific databases initially identified a total of 312 articles. Following the removal of 12 duplicate records using automated tools, 300 unique articles remained and underwent title and abstract screening. This phase was conducted independently by two reviewers, who resolved any discrepancies through consensus or with the involvement of a third reviewer when necessary. During this stage, 247 records were excluded based on predefined eligibility criteria.

For the remaining 53 articles, full-text retrieval was attempted. Employing both automated protocols and manual search methods, 33 reports could not be retrieved due to access restrictions, institutional limitations, or database errors. The remaining 20 full-text articles were thoroughly assessed by two independent reviewers, with cross-verification procedures in place to ensure data accuracy and consistency.

Following detailed examination, 7 studies were excluded for reasons including absence of a VR intervention, lack of therapeutic focus, inconsistency with SMT protocols, non-military samples, and failure to apply any form of Cyber-SIT.

Ultimately, 13 studies fulfilled all inclusion criteria and were incorporated into the systematic review. To ensure methodological transparency and reproducibility, a standardized data extraction form was developed to collect and synthesize key information such as study design, type of VR technology used, participant demographics, therapeutic framework, and primary outcomes.

The study selection process adhered to PRISMA 2020 guidelines and is illustrated in Fig. 1, while Table 3 provides an overview of the bibliographic, methodological, and technological characteristics of the included studies.

2.1.1. Quality assessment

Every chosen article underwent a systematic quality assessment to evaluate both the reporting quality and the presence of methodological

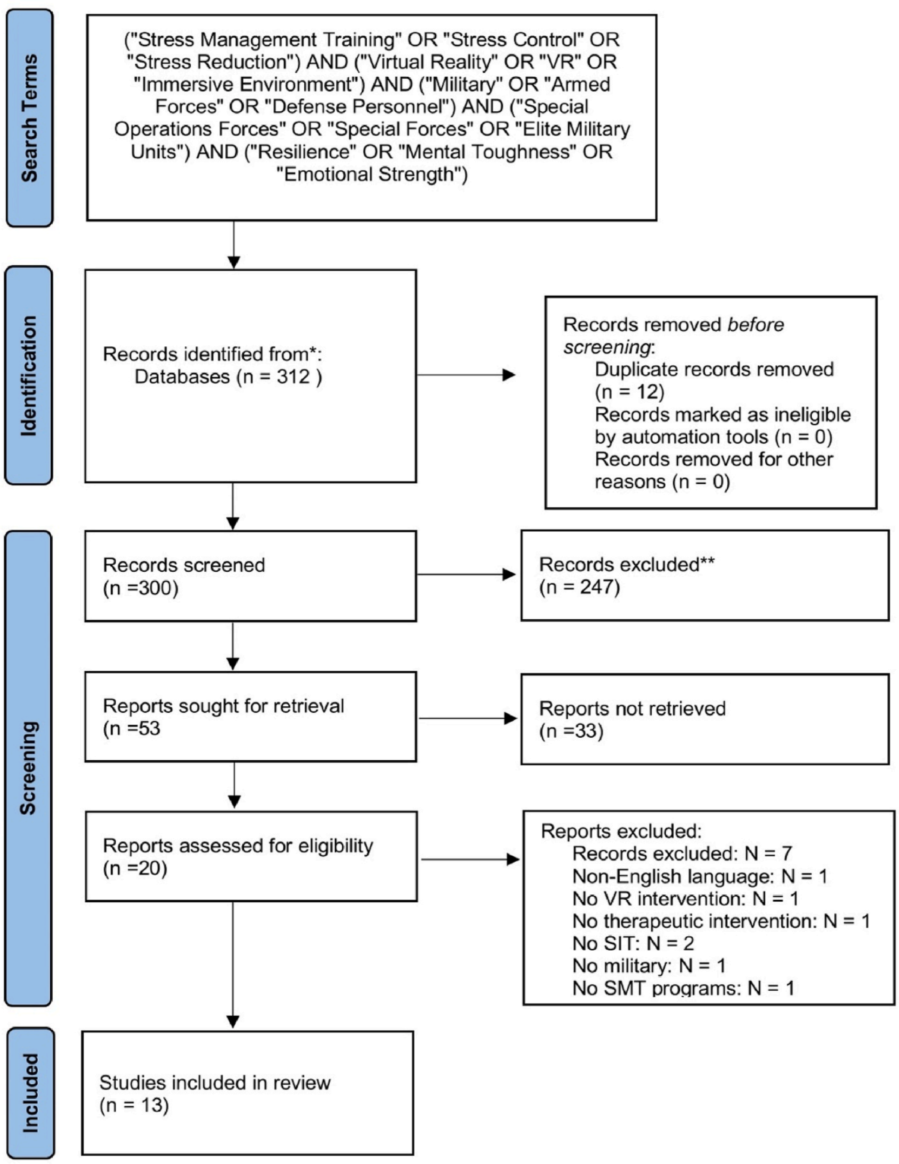

Fig. 1. PRISMA flowchart of screening, exclusion and inclusion criteria.

biases (check the Supplementary Materials). The assessment of study quality was conducted utilizing the Downs and Black checklist, which incorporates four distinct domains of evaluation: reporting quality, external validity, bias in internal validity, and confounding in internal validity (Downs & Black, 1998). Each study included in the review was methodically evaluated for quality employing this instrument as referenced in the source. A modification was applied to the scoring protocol of item 27, which probes the statistical power of the study to identify a clinically significant effect with a probability of a type I error below 5 %. The original scoring range proposed for this item, which consisted of five levels, was simplified to a dichotomous outcome. The revised scoring awarded a 1 for studies that performed a power calculation to detect a significant clinical effect with a power of at least 0.80 at an alpha level of 0.05 and a 0 for those that did not. Consequently, the scoring for each of the 28 items was binary, with responses being coded as affirmative (=1) or negative/unable to determine (=0), except item 5, which allowed for a gradation of responses: affirmative (=2), partial affirmation (=1), or negative (=0). The composite scores of the checklist were then categorized into four levels of methodological quality: excellent (26–28), good (20–25), fair (15–19), and poor (≤14).

According to the Downs and Black scores, in 4 of 13 studies (N = 4/13) 30,76 % had a result of excellent (26–28), in 5 of 13 studies (N = 5/13) 38,48 % had a result of good (20–25) and 4 of 13 studies (N = 4/13) 30,76 % had a result of fair (19 points) (Fig. 2).

2.2. Study 2

This study employed the Delphi technique—a structured, iterative, and anonymous communication methodology—to reach expert consensus on the design of a VR-based training program aimed at strengthening resilience and stress management competencies (Hsu & Sandford, 2007). This method was selected due to its suitability for complex decision-making in specialized contexts, such as military training, where hierarchical dynamics may compromise open dialogue.

Unlike focus groups or traditional surveys, the Delphi technique minimizes the influence of dominant participants and groupthink by preserving participant anonymity, thus fostering independent and candid input across rounds. This is particularly relevant in Special Operations Forces (SOF) environments, where operational structure and military rank can inhibit open exchange (Hsu & Sandford, 2007; Shang, 2023).

To accommodate the operational constraints of active-duty personnel, the Delphi study was conducted entirely online—using Welphi, a platform optimized for remote Delphi rounds. The platform enabled asynchronous participation and automated data management, including real-time analytics, round summaries, and controlled feedback loops. This approach reduced the burden of repeated in-person sessions while maintaining methodological rigor and enhancing participation rates.

The study design included three iterative rounds, during which feedback from previous rounds was integrated into successive questionnaires. This facilitated the refinement of proposed VR scenarios until expert consensus was achieved.

2.2.1. Participants

Participants were deemed eligible based on the following criteria: (1) completion of a Special Operations qualification more than three years prior; (2) operational deployment with National Force Deployed (NFD) units; (3) active service in a Special Operations Task Unit (SOTU); (4) proven experience in training or instruction of military personnel; (5) advanced knowledge of SOF tactics and procedures; and (6) familiarity with VR-based training methods and their comparison to traditional formats.

All participants provided both written informed consent and verbal assent, in line with ethical requirements. Full confidentiality was assured, and participants were informed of their right to withdraw at any stage without penalty.

The final sample included 30 male participants from army SOF units, with an average age of 33.43 years (range: 21–49). The all-male composition reflects current demographics within the Portuguese SOF, where female personnel are not yet integrated. Of the participants, 12 were single, 17 married or in relationships, and 1 divorced. The rank distribution included 10 Officers, 16 Sergeants, and 4 Corporals—reflecting the organizational structure of typical SOF units, where Sergeants are the most represented due to their central role in mission planning and execution.

This study was approved by the Chief of Staff of the Army (ref. GABCEME-RAG-2024-S-00707) and submitted to the Ethics Committee of the School of Health, Polytechnic of Porto (ref. CE-0083E).

2.2.2. Instruments

A 5-point Likert scale was used to capture expert ratings on statements presented during both the first and second rounds of the Delphi process. This scale ranged from 1 (strongly disagree) to 5 (strongly agree), with intermediate points representing agreement (2), neutral opinion (3), and disagreement (4). The efficacy and ease of use of the Likert scale have been well established in facilitating expert evaluation of items in Delphi studies (Giannarou & Zervas, 2014; Hsu & Sandford, 2007). This design allowed participants to express more nuanced opinions while avoiding forced polarization. The midpoint (3 = Neutral) was intentionally included to capture ambivalence or uncertainty, particularly in emerging or complex topics, such as VR realism requirements (Shang, 2023).

Additionally, three open-ended questions were included to provide specialists with an opportunity to contribute more freely to the development of the program. For example, one of the questions was: " *Could you please identify three situations/scenarios that, in your perspective, induce greater needs for stress management in special operations military personnel?"*

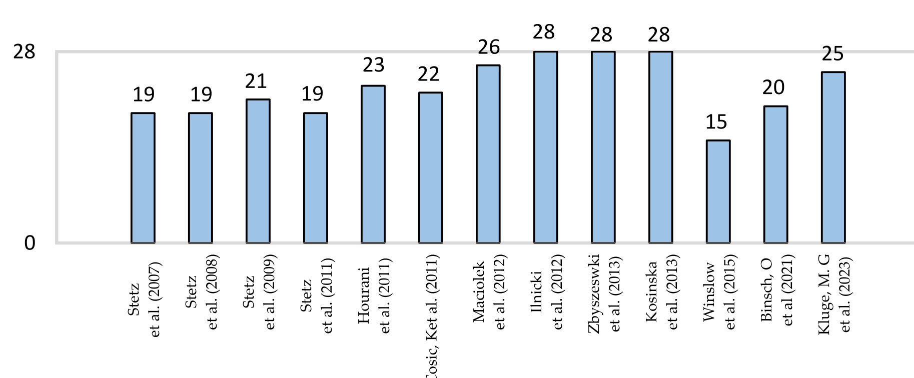

Fig. 2. Downs and Black (1998) —Checklist for assessment of the methodological quality.

2.2.3. Data collection

Round 1: Initial Data Collection – The first round of the study consisted of a questionnaire divided into two sections. The first section gathered sociodemographic information from the participants, such as rank, age, length of service, category, and marital status. The second section focused on assessing the characteristics of the proposed program for developing stress management skills through virtual reality. Experts were asked to evaluate various statements relevant to determining the program's requirements. For each item, experts were asked to express their opinion using a 5-point Likert scale.

The analysis of the responses focused on identifying items that achieved a 90 % consensus in the strong agreement categories (4 and 5). Items that met this criterion were considered consensual and did not advance to Round 2, allowing Round 2 to focus on items that still required further discussion.

Round 2: Review of Non-Consensual Items – In the second round, experts were sent the items that had not reached the 90 % consensus in the first round. Additionally, experts received a summary of the group's responses, allowing them to review their opinions in light of the collective consensus. The analysis followed the same criteria as Round 1, meaning that items that achieved 90 % or more agreement (in categories 4 and 5) were retained for Round 3, while items that did not meet the consensus criterion were discarded or did not progress.

Round 3: Consolidation and Final Confirmation – The third and final round involved presenting a consensus letter to the experts, containing all items that had met the 90 % or higher consensus criterion in Rounds 1 and 2. In the consensus letter, experts were invited to review their opinions based on the general consensus of the group and to agree or disagree with the items presented. The goal was to consolidate the final decisions on the program's characteristics, ensuring that the items included in the virtual reality scenarios for stress management were widely supported by the experts.

2.2.4. Data analysis

Descriptive statistics were calculated for both demographic characteristics and Likert scale responses. The distribution of responses for each question, including the number and percentage of participants ($N = 30$) selecting each option, was analyzed.

The open-ended questions from the initial phase were analyzed thematically by our research team, following the deductive and inductive approaches outlined by Braun and Clarke (2006). This rigorous approach involved a detailed examination of the text to identify recurring patterns (open coding), which were then further developed into themes through systematic inductive and deductive reasoning. Three team members independently reviewed all open-ended responses, formulating initial open and axial codes. Participant quotations that best exemplified each subtheme were selected based on their representativeness and the occurrence of diverse perspectives.

A concurrent parallel approach, guided by a data transformation model, facilitated data triangulation. This method enabled the research team to effectively juxtapose quantitative and qualitative data, thereby reinforcing subsequent phases of study design, data collection, and analysis. Moreover, consensus on each question was reached when 90 % or more of the experts agreed, strengthening the validity of our findings and ensuring a comprehensive understanding of participant perspectives. As suggested by Shang (2023), triangulating data from multiple sources helped ensure the robustness and reliability of the results, particularly in resolving any discrepancies between the quantitative and qualitative findings.

3. Results

3.1. Study 1 – Description of studies

The thirteen studies reviewed, encompassing 917 participants (80.5 % male; mean age range: 19–44 years), were predominantly conducted in the United States, with one exception in Australia (Kluge et al., 2023). These studies evaluated VR-based SIT for military personnel.

Table 3 summarizes the methodological characteristics of these studies, including references, country of origin, PTSD diagnostic instruments, study design, participant characteristics, scenario types, experimental conditions, and psychological outcomes. The PTSD Checklist – Military Version (PCL-M) served as the primary diagnostic tool for PTSD symptoms in eight studies (Ilnicki et al., 2012; Kosinska et al., 2012; Maciolek et al., 2014; Stetz et al., 2007, 2008, 2009, 2011; Zbyszewski et al., 2012), while three studies used the State-Trait Anxiety Inventory (STAI) to assess anxiety (Ilnicki et al., 2012; Stetz et al., 2011; Winslow et al., 2015). Other measures included the PCL-5 (Hourani et al., 2011), the Beck Depression Inventory-II (BDI-II), the Tellegen Absorption Scale (TAS) (Ilnicki et al., 2012), and assessments of temperament and coping via the Formal Characteristics of Behavior –

Table 1
Psychological indicators.

Psychological Indicators			
Indicator	Study Reference	Results	Statistical Data
MAACL-R (Anxiety)	Stetz et al. (2007)	35 % reduction in anxiety post-VR-SIT training.	Not reported.
MAACL-R (Hostility)	Stetz et al. (2007)	22 % reduction in hostility post-VR-SIT training.	Not reported.
MAACL-R (Depression)	Stetz et al. (2007)	15 % reduction in depression post-VR-SIT training.	Not reported.
MAACL-R (Anxiety, Hostility, Depression)	Stetz et al. (2008)	Significant reduction in anxiety, hostility, and depression (values unspecified).	Not reported.
Presence Questionnaire	Stetz et al. (2007, 2008)	VR-SIT group: Mean = 6.3/7 (SD = 0.4). Control group: Mean = 5.5/7. Higher immersion reported in VR-SIT.	Not reported.
STAI (State Anxiety)	Stetz et al. (2011); Maciolek et al. (2014); Winslow et al. (2015)	Significant reduction in state anxiety (VR-SIT: pre 42.3 ± 8.1 vs. post 38.5 ± 7.2). Large effect size (Cohen's d = 0.89; Winslow et al., 2015).	Wilcoxon Z = 2.06 (p = 0.04; Maciolek et al., 2014).
PTSD Symptoms (PCL-M)	Stetz et al. (2009); Houram et al. (2011); Maciolek et al. (2014)	13 % PTSD prevalence (9 % in VR-SIT vs. 18 % control). No significant differences post-training (Maciolek et al., 2014).	Not reported.
Depression (BDI)	Maciolek et al. (2014); Winslow et al. (2015)	Low baseline scores (mean < 10; Maciolek et al., 2014). Large reduction (Cohen's d = 0.87; Winslow et al., 2015).	Not reported.
Coping Strategies (CISS)	Maciolek et al. (2014)	Increased task-focused coping in VR-SIT group.	p = 0.03.
Engagement (UES-SF)	Kluge et al. (2023)	High engagement scores: Controlled breathing = 3.9/5; Muscle relaxation = 3.8/5.	Not reported.
State Mindfulness Scale	Kluge et al. (2023)	Significant increase post-grounding module (+6.14 points).	p < 0.001.
Threat/Challenge Appraisal (TCA)	Binsch et al. (2021)	3 participants classified as "threat response"; 26 as "challenge response."	Not reported.

Table 2
Physiological indicators.

Psychological Indicators			
Indicator	Study Reference	Results	Statistical Data
Heart Rate Variability	Stetz et al. (2007); Hourani et al. (2011); Kosinska et al. (2012)	18 % increase in HRV (VR-SIT group; Stetz et al., 2007). Improved recovery (↑ SDNN) in PRESIT group (Hourani et al., 2011).	Not reported.
Galvanic Skin Response	Stetz et al. (2007, 2008); Winslow et al. (2015); Binsch et al. (2021)	Reduced conductance peaks (VR-SIT group). Significant reduction (Cohen's d = 0.85; Winslow et al., 2015). Initial EDA increase (Binsch et al., 2021).	$p < 0.05$ (Stetz et al., 2007).
Cortisol Levels	Stetz et al. (2007, 2008); Hourani et al. (2011); Winslow et al. (2015)	Lower cortisol peaks in VR-SIT group. Large reduction (Cohen's d = 0.99; Winslow et al., 2015). Elevated anticipatory cortisol (Binsch et al., 2021).	Not reported.
Heart Rate	Maciolek et al. (2014); Winslow et al. (2015); Binsch et al. (2021)	Reduced from 85 ± 12 bpm to 72 ± 8 bpm (Maciolek et al., 2014). Moderate effect (Cohen's d = 0.64; Winslow et al., 2015).	$p < 0.05$ (Maciolek et al., 2014).
Respiratory Rate	Kluge et al. (2023)	Controlled breathing: 16.3 → 8.5 breaths/min. Muscle relaxation: 19 → 6.9 breaths/min.	$p < 0.001$.
Blood Pressure	Winslow et al. (2015)	Systolic: 136.7 → 128.3 mmHg (d = 0.83). Diastolic: 84.9 → 80.2 mmHg (d = 0.72).	$p < 0.05$.
Cardiac Output	Binsch et al. (2021)	Increased during initial VR exposure.	Not reported.
Total Peripheral Resistance	Binsch et al. (2021)	3 participants showed threat response (TPR >0).	Not reported.
Physiological Arousal	Kosinska et al. (2012)	Decreased arousal in 3/4 soldiers during final VR-SIT session.	Not reported.

Temperament Inventory (FCB-TI), Coping Inventory for Stressful Situations (CISS), and the NEO Personality Inventory – Revised (NEO-PI-R) (Ilnicki et al., 2012; Zbyszewski et al., 2012). Kluge et al. (2023) additionally used the State Mindfulness Scale (SMS) to assess mindfulness levels.

All studies applied SIT within VR environments, aiming to enhance resilience through cognitive-behavioral strategies such as progressive muscle relaxation, controlled breathing, and exposure to controlled stressors. Scenarios ranged from combat simulations (e.g., Combat Medic by Stetz et al., 2007, 2008; Virtual Battlespace by Winslow et al., 2015; Binsch et al., 2021) to relaxation-focused environments like Dream Island (Stetz et al., 2009, 2011) and mission-specific stressors such as Afghanistan's War (Ilnicki et al., 2012) and Physiological Control Exposure (Zbyszewski et al., 2012). Despite this diversity, all interventions emphasized exposure to adaptive stress challenges—including time pressure, auditory disruptions, and hostile virtual landscapes—to cultivate coping skills transferable to real-world operational settings (Hourani et al., 2011; Winslow et al., 2015).

Table 4 provides an overview of the therapeutic and technological dimensions, including duration of treatment, number of sessions, whether homework was assigned, psychological and physiological indicators measured, as well as the type of hardware and software used. Session durations ranged from 7 to 33.9 min (typically 20–45 min), with frequencies from two sessions (Hourani et al., 2011) to 327 sessions in longitudinal training (Kluge et al., 2023). Only Stetz et al. (2011) included homework, using portable VR devices for relaxation practice.

3.2. Psychological data

The reviewed studies employed a comprehensive array of psychological instruments to evaluate the impact of VR) interventions on stress, anxiety, PTSD, and coping mechanisms in military personnel. PTSD symptoms, assessed via the PCL-M across eight studies, showed significant post-intervention reductions in severity. For instance, Houram et al. (2011) reported a lower PTSD prevalence in VR-SIT groups (9 %) compared to controls (18 %), while longitudinal assessments highlighted contextual factors, such as mission completion, contributing to anxiety reductions in both intervention and control groups (Ilnicki et al., 2012; Maciolek et al., 2014).

Anxiety levels, measured using the State-Trait Anxiety Inventory (STAI) and Multiple Affect Adjective Checklist-Revised (MAACL-R), demonstrated consistent improvements. Stetz et al. (2007) found a 35 % reduction in anxiety following VR-SIT, with combined cognitive-VR training yielding lower post-treatment scores ($M = 48$, $SD = 5.5$) compared to VR-only groups ($M = 60$, $SD = 20$; $F = 4.16$, $p < 0.05$). These results were reinforced by Maciolek et al. (2014), who observed a decrease in state anxiety from 42.3 ± 8.1 to 38.5 ± 7.2 (Wilcoxon $Z = 2.06$, $p = 0.04$), and Winslow et al. (2015), who reported large effect sizes (Cohen's $d = 0.89$).

Immersion and engagement emerged as critical mediators of efficacy. The Presence Questionnaire (PQ) revealed higher immersion scores in VR-SIT groups ($M = 6.3/7$, $SD = 0.4$) versus controls ($M = 5.5/7$), suggesting enhanced realism and emotional regulation during simulations (Stetz et al., 2007, 2008). Modern VR platforms further underscored this relationship, with Kluge et al. (2023) linking mindfulness gains (State Mindfulness Scale, +6.14 points, $p < 0.001$) and high engagement scores (3.8–3.9/5) to improved resilience.

Beyond anxiety and PTSD, adaptive coping strategies and temperamental traits moderated outcomes. Soldiers with balanced temperament profiles (low emotional reactivity on the FCB-TI) exhibited superior stress adaptation (Ilnicki et al., 2012; Kosinska et al., 2012), while VR-SIT participants demonstrated increased task-focused coping (CISS, $p = 0.03$). Additionally, Binsch et al. (2021) identified distinct stress-response profiles: 26 of 63 participants exhibited "challenge" appraisals (perceived control under stress), compared to only 3 classified as "threat" responders.

Depression, assessed via the Beck BDI-II, showed low baseline scores ($M < 10$) and large post-intervention reductions (Cohen's $d = 0.87$; Winslow et al., 2015). However, mixed results in PTSD symptom persistence post-deployment (Maciolek et al., 2014) highlight the need for longitudinal designs to disentangle intervention effects from contextual confounders.

In summary, VR-based interventions—particularly VR-SIT—demonstrate robust efficacy in reducing psychological distress, enhancing immersion, and fostering adaptive coping. Key outcomes, including effect sizes and statistical significance, are synthesized in Table 1. Future research should prioritize individualized protocols and long-term follow-ups to optimize resilience training in high-stress occupations.

3.3. Physiological data

Physiological indicators were measured in nine of the thirteen studies, with heart rate (HR) emerging as the most frequently tracked metric. Heart rate variability (HRV), a key biomarker of autonomic regulation, showed significant improvements: Stetz et al. (2007) reported an 18 % increase in HRV post-VR-SIT, while Hourani et al. (2011)

Table 3
Descriptive characteristics = 13 included studies.

References	Country	Instrume for PTSD Diagnosis	Study Design	Sample and Participants	Scenario	Condition(s)	Outcomes
Stetz t al., (2007)	USA	PCL-M	RCT	Soldiers N = 25 Male: 100 % Age Range: 21–42 years;	Combat Medic Scenario Individual training VR-SIT	1 – VR group 2 – CT group 3 – CT + VR group 4 – Control	Psychological Measures: Post-Treatment Anxiety: Significant difference between groups (F(35,2) = 4.16F(35,2) = 4.16, p < 0.05p < 0.05). VR-only group: M = 60M = 60, SD = 20SD = 20. CT + VR: M = 48M = 48, SD = 5.5SD = 5.5. CT-only: M = 45M = 45, SD = 0SD = 0. Post-Treatment Dysphoria: Trending significance (F(35,2) = 3.53F(35,2) = 3.53, p < 0.10p < 0.10). Physiological Measures: Salivary Amylase Transformation: Increased stress in later sessions (F(18,2) = 4.78F(18,2) = 4.78, p < 0.05p < 0.05). First session: M = 250M = 250 s, SD = 106SD = 106. Second session: M = 362M = 362 s, SD = 185SD = 185. Breathing Patterns: Slight diaphragm expansion in CT groups (non-significant). Presence in VR (7-point scale): Involvement/Control: 3.73.7, SD = 0.78SD = 0.78. Natural Interaction: 3.23.2, SD = 1.3SD = 1.3. Interface Quality: 3.33.3, SD = 0.9SD = 0.9.
Stetz et al. (2008)	USA	PCL-M	RCT	Soldiers N = 63 Male: 47,75 % Age Range: 35 years	Combat Medic Individual training VR-SIT	1 – VR group 2 – CT group 3 – CT + VR group 4 – Control	Hostility: Higher in VR group vs. CT group (Δ = 6.1Δ = 6.1, p < 0.01p < 0.01). Sensation-seeking: Higher in CT + VR group vs. VR (Δ = 5.9Δ = 5.9, p < 0.01p < 0.01) and CT (Δ = 9.1Δ = 9.1, p < 0.001p < 0.001).Anxiety/Dysphoria: Higher in previously deployed participants in VR/CT groups (p < 0.05p < 0.05). Presence: Moderate realism (PQ scores: 3–5/7). MANOVA: Significant effect of condition on stress (F(12,242) = 3.3F(12,242) = 3.3, p < 0.001p < 0.001). Interaction Effects: Anxiety (F(2,126) = 3.3F(2,126) = 3.3, p < 0.05p < 0.05) and dysphoria (F(2,126) = 4.9F(2,126) = 4.9, p < 0.05p < 0.05) in deployed participants.
Stetz et al. (2009)	USA	PCL-M	RCT	Soldiers N = 39 Male: 65 % Age Range: 33 years;	"Dream Island", a virtual environment that embedded script explaining how to practice Progressive Muscle Relaxation and Controlled Breathing.	1 – EG 2 – CG	Participants reported enjoyment and willingness to continue practicing relaxation techniques through VR. Increased receptiveness toward stress management training. VR facilitated engagement and positive behavioral intent.
Stetz et al. (2011)	USA	PCL-M STAI	RCT	Soldiers N = 60 Male: 66 % Age Range: 35 years	Virtual calming environments. Dream Island	1 – EG 2 – CG	Anxiety Reduction: EG showed significant post-intervention STAI-State reduction vs. CG (F[1,40] = 5.77F[1,40] = 5.77, p < 0.021p < 0.021, η2 = 0.13η2 = 0.13). Stressful Tasks: No significant difference in anxiety between groups during simulations. Focus Group: 79 % of EG felt relaxed; 53 % preferred VR-guided relaxation; 50 % intended to continue SMTs post-study.
Hourani et al., (2011)	USA	PCL-5	RCT	Soldiers N = 77 Male: 100 % Age Range: n. r.	Infantry Immersion Training Iraqi village	1 – EG (PRESIT training with a "multimedia stress environment'');	HRV: SDNN showed increased arousal during MSE (p < 0.05) but no significant group differences. pNN50 indicated greater relaxation post-training in previously deployed PRESIT participants

(continued on next page)

Table 3 (*continued*)

References	Country	Instrume for PTSD Diagnosis	Study Design	Sample and Participants	Scenario	Condition(s)	Outcomes
						2 – CG ("current best practices'' training).	($p < 0.05$). RT: PRESIT group improved hit rate (91 % vs. 85 % in controls; $p < 0.04$). Qualitative: 86 % expressed interest in stress management techniques. Mixed-effects models accounting for squad-level clustering. Significant predictors of HRV improvement: prior deployment, baseline PTSD symptoms ($p < 0.05$), and prior stress management training.
Cosic et al., (2011)	USA	n.r.	RCT	Soldiers N = n.r. Age Range: n. r.	VRAS training system VRAS StimGen	1 – EG 2 – CG	fMRI Findings: Deployment-ready soldiers showed increased activation in temporo-parietal junction (TPJ) and superior parietal lobe vs. novices. Physiological: Lower reactivity in deployment-ready group (preliminary data).
Maciolek et al. (2014)	USA	PCL-M	RCT	Soldiers N = 118 Male: 112 (95 %) Female: 6 (5 %) Age Range: 21–44 years	Virtual Reality Medical Center of San Diego	1 – EG 2 – CG	Pre-training: No significant differences between E and C groups. Post-training: Significant reduction in state (X1) and trait (X2) anxiety in E group ($p = 0.04$). C group showed reduced X2 ($p = 0.01$) but no change in X1. Post-deployment: Both groups showed reduced X1 and X2 vs. pre-training ($p < 0.05$). X2 further decreased vs. post-training ($p < 0.01$). No long-term group differences: E and C groups did not differ significantly post-deployment.
Ilnicki et al. (2012)	USA	PCL-M STAI FCB-TI TAS CISS NEO-PI-R	RCT	Soldiers N = 118 Male: 112 (95 %) Female: 6 (5 %) Age Range: 21–44 years	Virtual Reality Medical Center of San Diego	1 – EG 2 – CG	Short-term: Significant reduction in state (X1) and trait (X2) anxiety post-SIT ($p < 0.05$). Long-term: No significant differences between E and C groups post-deployment. Both groups showed reduced anxiety, possibly due to non-SIT factors (e.g., mission completion). Temperament/personality: Higher emotional reactivity (FCB-TI) correlated with greater anxiety post-training. Balanced temperament predicted better SIT outcomes. PCL-M: Post-deployment PTSD symptoms linked to emotion-oriented coping(CISS) and high emotional reactivity.
Zbyszewski et al. (2012)	USA	PCL-M FCB-TI CISS NEO PI-R	RCT	Soldiers N = 120 Male: 100 % Age Range: n. r.	Physiological control exposure to virtual stressor	1 – EG (cyber-SIT); 2 – CG (military training).	Significant correlations found between emotional reactivity and stress levels (pre-, post-SIT, and post-deployment); Post-deployment correlations with emotion-based, avoidance coping and conscientiousness
Kosinska et al. (2012)	USA	PCL-M	Case Sudy	Soldiers N = 4 Male: 100 % Age Range: n. r.	Virtual Reality Medical Center of San Diego	1 – EG (cyber-SIT)	3/4 soldiers reduced arousal during final session 2/4 improved relaxation 3/4 reduced post-exposure arousal Harmonized temperament correlated with better outcomes
Winslow et al. (2015)	USA	STAI	RCT	Soldiers N = 40 Male: 33 (82,5 %) Female: 7 (17,7 %) Age Range: 21–35 years	Virtual Battlesplace 2	1 – VR-based training group) 2 – No training group.	Utility of adding stressors external to the virtual training environments. Between groups, the experimental group reported a significant increase in stress as compared to the control group in scenarios 1, 2, and 4 ($p \leq 0.05$) and during the TSST ($p \leq 0.001$). Differences between: Scenarios 3 and 4 ($p \leq 0.05$); Scenarios 1, 2, and 5 ($p \leq 0.001$); Scenarios 1 and 3 ($p \leq 0.001$); 4 and scenarios 2, 3, and 5 ($p \leq 0.05$).
Binsch et al. (2021)	USA	n.r.	RCT	Soldiers N = 64 Male: 61	Patrol scenario (military convoy) Virtual Battle Space	3 groups 5 (Condition) × 3 (Group)	5 Condition (C) × Group (G) HR – $F(2.36, 118) = 18.71$, $p < 0.001$, $\eta p2 = 0.27$;

(*continued on next page*)

Table 3 (*continued*)

References	Country	Instrume for PTSD Diagnosis	Study Design	Sample and Participants	Scenario	Condition(s)	Outcomes
				(95,31 %) Female: 3 (4,66 %) Age Range: 19–28 years			HRV – F(2.7135.8) = 7.52, p < 0.001, ηp2 = 0.13; S/DBP – no significant main effects; EDA – F(3.6, 177.7) = 7.54, p < 0.001, ηp2 = 0.13 CO – F(2.7, 133.0) = 6.31, p = 0.001, ηp2 = 0.11; TPR and challenge/threat index – no significant main effects; 4 Condition (C) × Group (G) Cortisol – no significant main effects; Performance scores – F (2.8, 135.6) = 7.53, p < 0.001, ηp2 = 0.13,
Kluge et al. (2023)	AUS	SMS	RCT	Soldiers N = 189 Male: 100 % Age Range: n.r.	Performance Edge	1 – VR-based training group)	Results showed positive impact of the program on stress management. The outcomes demonstrated the perceived usefulness, feasibility, usability, and positive training outcomes of the technology platform, training concept, and specific training modules within the intended real-world context and training population.

USA, United States of America; PCL-M, PTSD Checklist, Military Version; RCT, randomized controlled trial; VR, virtual reality; SIT, Stress Inoculation Training; CT, Coping Training; Post, post-treatment assessment; EG,Experimental group; CG, Control group; STAI, State-Trait Anxiety Inventory; PCL-5, PTSD checklist for DSM-5, measured via PCL-5 total sum score; PRESIT, Predeployment stress inoculation training; MSE, Multimedia Stressor Environment; PTSD, post-traumatic stress disorder, Pre, pre-treatment assessment; n.r., not reported; PTOAssoc, Parietal-Temporal-Occipital Association Cortex; OFC/VMPFC, Orbitofrontal/ventromedial prefrontal cortex; VRAS, Virtual reality adaptive stimulation; FCB-TI, Formal Characteristics of Behavior-Temperament Inventory; TAS, Tellegen Absorption Scale; CISS, Coping Inventory for Stressful Situations; NEO-PI-R, Revised NEO Personality Inventory; VLF, Very Low Frequency; LF, Low Frequency; TSST, Trier Social Stress Test; SMS, State Mindfulness Scale; HR, Heart rate; HRV, Heart rate variability; S/DBP, Systolic/Diastolic blood pressure; EDA, Electro dermal activity; CO, Cardiac output; TPR, Total peripheral resistance.

linked prior deployment experience to enhanced post-stress recovery (↑ SDNN, $p < 0.05$). These findings align with reductions in galvanic skin response (GSR) peaks, reflecting lower physiological arousal during simulations (Cohen's $d = 0.85$; Winslow et al., 2015).

Cortisol levels, a marker of hypothalamic-pituitary-adrenal (HPA) axis activation, demonstrated clinically meaningful changes. Winslow et al. (2015) documented large reductions (8.1 → 5.6 μg/dL; Cohen's $d = 0.99$), with VR-SIT participants exhibiting fewer cortisol spikes during high-stress scenarios—particularly when combined with tactical breathing techniques (Stetz et al., 2007, 2008).

Cardiovascular metrics also improved robustly. Resting heart rate decreased from 85 ± 12 bpm to 72 ± 8 bpm after 10 VR-SIT sessions ($p < 0.05$; Maciolek et al., 2014), while systolic and diastolic blood pressure dropped by 8.4 mmHg and 4.7 mmHg, respectively (Cohen's $d = 0.83$ and 0.72; Winslow et al., 2015). Respiratory regulation, a focus of VR biofeedback protocols, saw dramatic improvements: Kluge et al. (2023) reduced respiratory rates from 16.3 to 8.5 breaths/min during controlled breathing modules ($p < 0.001$).

Stress-response profiles varied significantly across participants. Binsch et al. (2021) classified 26/63 cadets as "challenge responders" (↑ cardiac output, ↓ peripheral resistance), while only 3 exhibited "threat responses" (↑ vasoconstriction), correlating with subjective anxiety. Neuroimaging data from Ćosić et al. (2011) revealed heightened activation in stress-processing neural regions (e.g., temporo-parietal junction, superior parietal lobe), particularly among deployment-ready soldiers, underscoring VR's role in modulating neurophysiological pathways.

Individual differences influenced outcomes: soldiers with prior combat exposure exhibited stronger physiological reactivity but greater responsiveness to combined cognitive-VR interventions (Stetz et al., 2008). Longitudinal programs, such as Kluge et al. (2023) 327-session protocol, confirmed the feasibility of sustained physiological monitoring in VR-SIT frameworks. However, hardware variability (e.g., Oculus Rift vs. 3D Visor Headsets) may have compromised immersion and data comparability across studies.

Limitations persist, including small sample sizes (e.g., $n = 4$ in Kosinska et al., 2012), technical challenges (e.g., lost blood pressure data in Binsch et al., 2021), and a lack of longitudinal tracking. Nevertheless, the aggregate evidence positions VR as a potent tool for enhancing physiological resilience, with biofeedback-driven modules offering targeted benefits for high-stress occupations.

For a detailed synthesis of physiological outcomes, including effect sizes and statistical thresholds, refer to Table 2.

3.4. Study 2

The authors devised a comprehensive questionnaire that included sociodemographic data and was structured into three phases: (1) evaluating stress perception through a series of five questions, (2) delineating the essential requirements for the VR program using 22 Likert scale questions, and (3) gathering valuable insights from military personnel regarding crucial aspects of the VR environment through three open-ended questions.

Descriptive statistics (i.e. % agreement) were computed for each survey item across all rounds, aggregating responses rated as "4 - Agree" and "5 - Strongly Agree" on the Likert scale. An a priori threshold was set, requiring a minimum of 90 % agreement for each requirement to achieve acceptance.

During Round 1, military personnel were asked about their perception of stress. In response to the question: "*To what extent do you consider stress a significant problem for you?*" 50 % indicated that stress was a significant issue for them. However, when asked: "*To what extent do you consider stress a significant problem for your comrades?*" the percentage increased, with 70 % responding that stress was a significant problem for

Table 4
Results of the qualitative analysis for each study.

References	Therapeutic	Period of Time	Number of Sessions	Homework	Psychological indicators	Physiological indicators	Hardware	Software
Stetz et al. (2007)	SIT	n.r.	2 to 4	n.r.	MAACL-R Anxiety Depression Hostility	PhysioLab MAACL-R HR PBT BR	Z800 3D Visor Head Mounted Display (HMD) Logitech Dual-Action Joypad Low Frequency Effects Speaker ScentPalette™	Combat Medic
Stetz et al. (2008)	SIT	n.r.	2 to 4	n.r.	MAACL-R Anxiety Depression Hostility Positive Affect Sensation Seeking	MAACL-R PMR CB	n.r.	Combat Medic Flight Medic
Stetz et al. (2009)	SIT	4 days Last day all participants answered the following questions as part of a Focus Group	n.r.	n.r.	n.r.	PMR CB	n.r.	Dream Island
Stetz et al. (2011)	SIT	7 min	3	Practicing technology assisted relaxation techniques by watching virtual reality videos on portable devices during the evenings following the training sessions.	STAI – Anxiety	PMR CB	2.5 Pro VR	Dream Island
Hourani et al., (2011)	SIT	Grup A – 1,5H Grup B – 30 min	2	n.r.	Stress in work or family (“a lot”) Depression Received stress man- agement for deploy/combat Practice relaxation techniques Practice martial arts Interested in learning stress reduction	HRV	LED pulse sensor (Cateye model PL-6000) Biocom Technologies model HRM02 SoundBlaster X-Fi Surround 5.1 USB adapter Audio amplifier with 5.1 surround capability Four Promonitor 1000 speakers ProSub 800 powered subwoofer Panasonic PTAE3000 high definition Projector	Infantry Immersion Training Iraqi village
Cosic et al., (2011)	SIT	n.r.	n.r.	n.r.	Stress	HR SC ST Respiration FMRI ECG	STIMDB SCNDB	VRAS training system; VRAS StimGen. EmoEst
Maciolek et al. (2014)	SIT	20–30 min	10	n.r.		Breathing HR SC	Two Dell Inspiron M1710 computers with Intel Core 2CPV 2 GHz processors, 2G RAM and the nVIDIA GeForce Go 7900 GS graphic card supporting the VR images andsounds). The Dell Inspiron MXC 061 computer with the Intel Core 2CPV 1.99 GHz processor and 2G RAM	Virtual Reality Medical Center Afghan Kabul Iraq Convoy Main PTSD Convoy PTSD Enchanted Forest

(continued on next page)

Table 4 (*continued*)

References	Therapeutic	Period of Time	Number of Sessions	Homework	Psychological indicators	Physiological indicators	Hardware	Software
Ilnicki et al. (2012)	SIT	5 days 90 min lecture on the nature	10	n.r.	STAI – Anxiety PCL-M – PTSD BDI-2 – Depression CISS – Stressful Situations FCB-TI – Behavior-Temperament NEO-PI-R	Breathing HR SC FT	n.r.	Afghanistan's war
Zbyszewski et al. (2012)	SIT	5 days	10		FCB-TI STAI – Anxiety Personality Temperamental structure Emotional Reactivity	n.r.	n.r.	Physiological control exposure to virtual stressor
Kosinska et al. (2012)	SIT	5 days 1 sessions – 13 min 2 to 10 – n.r.	10	n.r.	STAI NEO-PI-R FCB-TI	HR VLF LF	3 computers	Virtual Reality Medical Center of San Diego
Winslow et al. (2015)	SIT	n.r.	n.r.	n.r.	STAI – Anxiety	Cortisol HR SC ECG EMG	n.r.	Virtual Battlesplace 2
Binsch et al. (2021)	SIT	n.r.	n.r.	n.r.	Levels of Stress Stress Responses	ECG ABP EDA HR HVR CO S/DBP	Oculus Rift TENS- TDP, Medisana AcqKnowledge ECG (ECG100C-MRI) ABP (NIBP100D) EDA (100 C MRI) Torso (EL503) Finger (EL509) BIOPAC©	Virtual Battle Space
Kluge et al. (2023)	SIT	Module 1: 22.9 min Module 2: 17.4 min Module 3: 18.9 min Module 4: 33.9 min Module 5: 32.6 min	327	n.r.	Levels of Stress Physiological Stress Responses Relaxation Capacity Mindfulness and Emotional Awareness Motivation and Interest	Physical Assessment Scale HR RR GSR MT Cortisol	VR Headsets GoDirect respiratory belt (Vernier) Oculus Quest (Meta)	Performance Edge

SIT, Stress Inoculation Training; n.r., not reported; MAACL-R, Multiple Affect Adjective Check List-Revised; HR, Heart Rate; PBT, Peripheral Body Temperature; BR, Breathing Rate; Min, Minutes; PMR, Progressive Muscle Relaxation; CB, Controlled Breathing; STAI, State-Trait Anxiety Inventory; HVR, Heart rate variability; SC, Skin conductance; ST, Skin temperature; FMRI, Functional Magnetic Resonance Imaging; ECG, Electrocardiogram; PCL-M, PTSD Checklist, Military Version; BDI-2, Beck Depression Inventory; CISS, Coping Inventory for Stressful Situations; FCB-TI, Formal Characteristics of Behavior-Temperament Inventory; NEO-PI-R, NEO-Personality Inventory-Revised; FT, Finger temperature; VLF, Very Low Frequency; LF, Low Frequency; EMG, Electromyogram; ABP, Arterial blood pressure; EDA, electrodermal dermal activity; CO, Cardiac output; S/DBP, Systolic/Diastolic blood pressure; RR, Respiratory Rate; GSR, Galvanic skin Response; MT, Muscle tension.

both themselves and their comrades.

Furthermore, in Round 1, subject-matter experts (SMEs) evaluated various elements using twenty-two closed-ended questions with a Likert scale. They rated 72.27 % of the elements as "acceptable," while 27.72 % were rated as "unacceptable" because they did not meet the initial requirement of having an acceptance rate above 90 % (See Fig. 3 for Delphi flow results.).

Phase 3 of Round 1 consisted of three open-ended questions: (1) identify three situations/scenarios that, in your perspective, induce the greatest stress management needs in special operations military personnel; (2) identify three strategies you consider most effective for dealing with stressful situations; and (3) specify three requirements or characteristics that the program we are developing should consider.

The results indicated that decision-making under pressure scenarios caused the greatest stress for military personnel. The most important strategy for mitigating this stress was identified as training and experience. Studies have confirmed that decision-making under pressure is a significant source of stress for military personnel, and that comprehensive training and accumulated experience are critical strategies for effective stress management (Bartone et al., 2008).

Regarding the third question on the program requirements, the most frequently mentioned were: realism and immersion, mentioned 29 times (96.67 %) as the primary requirement; versatility and adaptability, mentioned 29 times (96.67 %); and evaluation and feedback, also mentioned 29 times (96.67 %).

Notably, all three requirements were mentioned with equal frequency, suggesting that an ideal program should prioritize realism and immersion, adaptability to diverse scenarios, and robust evaluation and feedback mechanisms.

The second Delphi questionnaire was conducted over a condensed two-day period. This focused approach addressed areas where the initial survey didn't achieve a strong consensus, meaning at least 90 % of experts agreed. Questions from the first round that sparked debate were presented again. Participants used a 5-point Likert scale (ranging from 1 for "strongly disagree" to 5 for "strongly agree") to provide nuanced evaluations on these revised questions. Additionally, three new yes-or-no questions were introduced to directly target the key themes identified from the open-ended questions in Round 1.

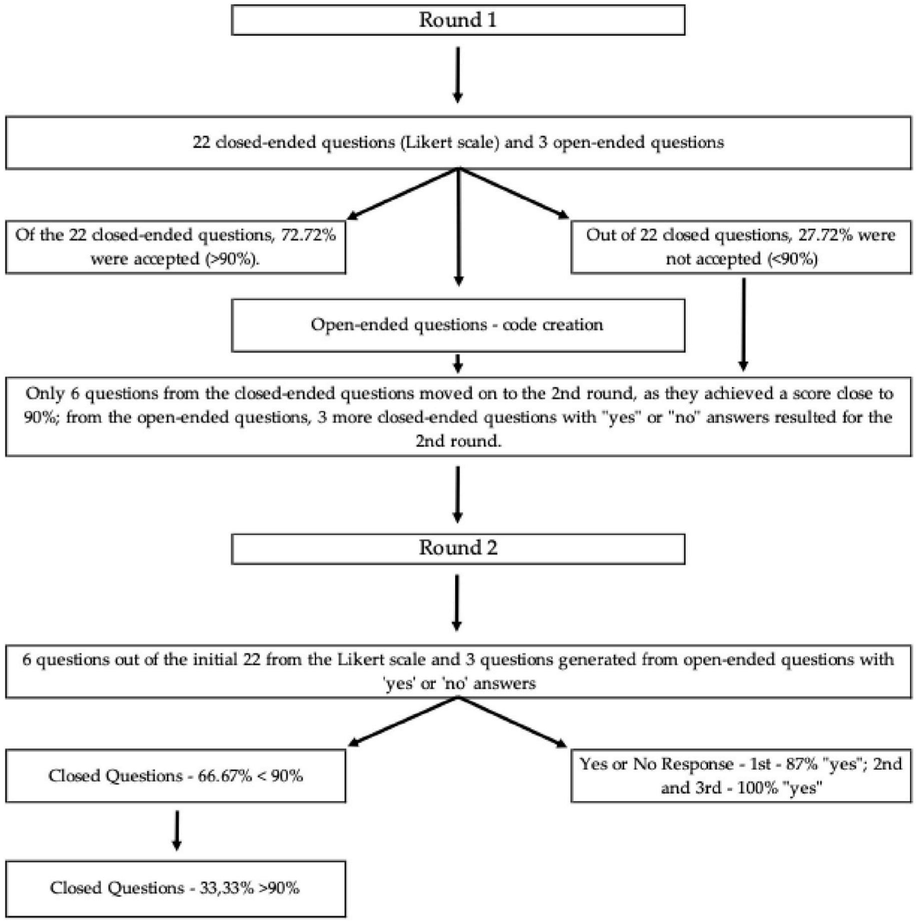

Fig. 3. Delphi study process and summary results.

To promote informed decision-making, participants received access to the aggregated results from the second survey. This transparency included all experts' individual responses and statistical summaries or items nearing consensus. By considering the collective insights and the distribution of opinions from other experts, participants could re-evaluate the importance of each item with a broader perspective.

This revised approach aimed to achieve consensus efficiently by focusing on areas of disagreement and leveraging both quantitative and qualitative data.

The third round culminated in a consensus letter, encompassing all points that garnered expert agreement exceeding 90 % on the Likert scale. This remarkable 97 % consensus signifies a strong alignment among the experts, solidifying the validity and uniformity of the points included in the final letter. The exceptionally high level of agreement not only highlights the robustness of the findings but also ensures that the recommendations are widely endorsed by the expert panel (see Table 5 for details).

This remarkable level of agreement underscores the strength and validity of the findings. The consensus letter, reflecting this shared perspective, provides a robust foundation for the further development and implementation of strategies based on the collective expertise of the specialists involved. It paves the way for significant advancements in training and preparation for special operations personnel.

The Delphi survey achieved near-unanimous agreement (97 %) on the key stressors faced by special operations personnel, including decisions under pressure, aiding the wounded, and isolation or capture. It also identified crucial stress management strategies, such as comprehensive training, meticulous planning, and effective teamwork. Additionally, the survey underscored the value of VR training in monitoring stress, simulating realistic environments, and providing actionable

Table 5
Final competencies – round 3.

Consensus Letter		
Category	*Final Competencies*	Percentage Agreement
Stressors in Special Operations	Decisions under pressure; Aid to wounded in combat; Isolation and capture; Urban combat.	97 %
Stress Management Strategies	Training and experience; Planning and preparation; Group work and camaraderie.	97 %
VR for training (Effectiveness)	Monitoring of biological signals; Simulation of extreme conditions; Adaptation of scenarios; Unpredictable elements; Detailed feedback on performance.	97 %
Operational Training	Team interaction; Immediate feedback; Sensory elements (vests, smells); Safe environment to learn from mistakes; Realism in training scenarios; Interactive elements; Varied environments; Repetition of scenarios for improvement; Simulation of extreme stress situations.	97 %

feedback. Emphasis was placed on fostering team interaction, delivering immediate feedback, and incorporating sensory elements into operational training to better prepare personnel for complex and high-pressure situations.

These findings underscore the importance of a multifaceted approach to stress management and training, ensuring that special operations personnel are equipped to handle the demanding conditions they encounter.

4. Discussion

4.1. Principal findings

This review provides important information about the efficacy of Virtual Reality-based Stress Management Training (VR-SMT) protocols for military populations. The 13 included studies (Stetz et al., 2007, 2008, 2009, 2011; Cosic et al., 2011; Hourani et al., 2011; Ilnicki et al., 2012; Kosinska et al., 2012; Maciolek et al., 2014; Winslow et al., 2015; Zbyszewski et al., 2012) demonstrate that VR-SMT—particularly those aligned with SIT and RT models—can lead to significant reductions in stress-related symptoms such as anxiety and PTSD, while also improving emotional regulation and physiological responses under pressure.

A recurrent finding across studies was the positive impact of VR-SMT on both self-reported and physiological indicators of stress. Standardized measures such as the STAI and the PCL-M were employed in a majority of studies to evaluate outcomes related to anxiety and trauma. The widespread use of these instruments (e.g., PCL-M in 8 out of 13 studies, 61.53 %) supports the validity and comparability of results across diverse samples and deployment contexts. Findings consistently suggest that VR-SMT helps military personnel maintain performance and cognitive clarity when faced with high-stress stimuli, both in training and post-deployment (Stetz et al., 2007, 2008; Hourani et al., 2011).

Physiological measures such as HRV, SC, HR, S/DBP, salivary biomarkers, and even functional MRI provided further validation of the interventions' effectiveness. For example, HRV—a widely recognized psychophysiological stress marker—was assessed in 9 out of 13 studies (69.23 %) (Stetz et al., 2007, 2008; Hourani et al., 2011). Some studies also observed increased salivary alpha-amylase activity post-intervention, reflecting acute stress induction and regulatory engagement (Hourani et al., 2011). These data reinforce the notion that VR-based SMT not only targets psychological coping but also modulates autonomic nervous system function (Stetz et al., 2009; Maciolek et al., 2014).

Moreover, a significant contribution of this review lies in the emphasis on personalization and individual variability. Several studies included personality assessments (e.g., NEO-PI-R, CISS, FCB-TI) to understand how different psychological profiles respond to VR-based stress training. This multidimensional approach enables a better understanding of who benefits most from these interventions and under which conditions (Zbyszewski et al., 2012; Kosinska et al., 2012).

The cyber-SIT protocols reviewed here demonstrate significant progress in merging traditional clinical principles with immersive VR environments.

A critical insight from Study 2, a Delphi study included in this review, emphasized the need to incorporate expert consensus and user feedback into the design of VR scenarios. Experts highlighted scenario realism, progressive exposure to stressors, and physiological feedback as essential components for effective intervention, all of which were echoed in several empirical studies (Stetz et al., 2008; Hourani et al., 2011).

Importantly, these interventions have proven adaptable across diverse military operational contexts, with some implemented pre-deployment (e.g., Hourani et al., 2011) and others showing sustained effects post-deployment (e.g., Maciolek et al., 2014). This adaptability reinforces the potential of VR-SMT to serve both preventive and rehabilitative purposes (Stetz et al., 2007, 2008).

Despite promising outcomes, the field is still in early stages, with several limitations outlined in the next section. Nonetheless, this review substantiates cyber-SIT and other VR-SMT modalities as innovative, empirically supported tools for enhancing resilience and psychological readiness in military personnel. The integration of rigorous clinical protocols, advanced technology, and interdisciplinary expertise holds promise for transforming stress management interventions in defense settings.

4.2. Limitations

Although the observed results are promising, it is important to acknowledge several limitations in the analysis of the 13 studies included in this review, which may influence the interpretation of the data and the overall conclusions. The primary limitation is the heterogeneity among the studies, which was evident in the variety of VR hardware used (e.g., Oculus Rift vs. 3D visors) and differences in scenario designs and intervention protocols. The variability in devices may have influenced the immersive experience and the effectiveness of the training. For example, studies by Kosinska et al. (2012) and Zbyszewski et al. (2012) utilized different types of VR technology, which may have led to divergent outcomes.

Furthermore, the lack of standardized intervention protocols made it difficult to draw direct comparisons between the studies. Specifically, Ilinski (2012) indicated that the intensity and type of stress induced varied significantly across scenarios. In some instances, the intensity of stress was not clearly specified, which complicates the interpretation of the effects of the training.

Another important limitation is the lack of strict control over participant characteristics, such as psychological history or pre-existing mental health conditions. Winslow (2015), for instance, did not explicitly control for participants' chronic stress levels or trauma history, which could have influenced their response to the training. This limits the generalizability of the findings, as different psychological profiles may result in divergent reactions to SMT. The majority of studies focused on male soldiers from SOF, such as those in Binsch (2021), with limited representation from other groups, such as women or soldiers from different units. This restricts the generalizability of the results, as stress responses may vary based on factors such as gender, age, and unit type.

Notably, the Delphi study 2, which aimed to establish expert consensus on the design of VR-based SMT, involved only male military personnel from SOF. This reflects the current reality in many military units, where women have historically been underrepresented or excluded from combat roles in certain countries, including Portugal. Although this is consistent with the structure of SOF in the Portuguese military, the lack of female representation in the Delphi study limits the generalizability of the findings to female military personnel. As female soldiers increasingly join diverse military roles, including combat units, future studies should strive for greater inclusivity to ensure that VR-SMT programs are relevant and effective for all military personnel.

Moreover, the studies employed diverse VR systems (e.g., Oculus Rift in Stetz et al. (2007, 2008, 2009, 2011) vs. 3D Visor Headsets in Zbyszewski et al. (2012) and Kosinska et al. (2012), leading to variability in immersion levels and training outcomes. Kosinska et al. (2012) reported reduced physiological arousal using low-fidelity systems, while Kluge et al. (2023) achieved higher engagement with advanced haptic feedback. This variability complicates cross-study comparisons and highlights the need for standardized VR hardware protocols.

In terms of evaluation methodology, many studies relied heavily on subjective instruments, such as the STAI and the PCL-M, as seen in Maciolek et al. (2014) and Murielle (2023). While these instruments are widely used, they are susceptible to response bias, such as social desirability bias or inaccuracies in participants' recall of past experiences. The integration of physiological biomarkers (such as HRV and cortisol) helped mitigate this limitation, but reliance on self-reports still affects the internal validity of some studies.

Additionally, the intensity and type of stressors varied significantly across the studies. Ilnicki (2012) and Maciolek et al. (2014) used combat scenarios with undefined stress thresholds, while Hourani (2011) focused on pre-deployment stressors without quantifying their psychological impact. This variability in stress induction complicates the interpretation of training effects.

Another major limitation concerns the duration and frequency of the training programs. Many studies, such as those by Kosinska et al. (2012) and Binsch et al. (2021), focused on short-term interventions, with sessions limited to only a few interactions. The lack of longitudinal studies prevents the assessment of the long-term effects of the training, particularly regarding post-deployment resilience or during the return to field operations. Additionally, as pointed out in studies like Ilinski (2012), the absence of long-term follow-up with participants hinders a complete understanding of the benefits and limitations of virtual reality training.

Moreover, the representativeness of the simulated scenarios in the studies was limited, with a predominant focus on combat or medical emergency situations, as seen in Winslow (2015) and Maciolek et al. (2014). While these scenarios are crucial, the absence of other types of stress faced by military personnel, such as psychological pressures in long-duration missions or non-combat environments, may affect the universal applicability of the results. Hourani (2011) highlighted this gap, noting that 72 % of military stressors occur in non-combat settings, yet none of the 13 studies addressed these contexts. The inclusion of other stressors, such as stress related to resource management in the field or the psychological adaptation to prolonged missions, could further enrich the training protocols and make them more adaptable to different situations.

Lastly, the lack of sample diversity (in terms of gender, age, and prior experience) in studies like Murielle (2023) and Binsch (2021) compromises the generalizability of the results to the broader military population. To ensure that SMT programs are truly effective for all military personnel, future research should include a broader range of participants, including women and soldiers from various military branches.

4.3. Future directions

Despite promising developments in VR-SMT for military personnel, this review has identified several critical areas requiring further investigation. One of the most pressing issues concerns the lack of standardization across studies, particularly regarding the types of VR hardware used, the design and intensity of scenarios, and the absence of consistent metrics for assessing stress exposure and immersion. Future research would benefit from the establishment of consensus-based guidelines, including standardized parameters for physiological indicators such as heart rate variability and validated immersion thresholds. Moreover, the limited ecological and longitudinal validity of current research designs restricts our understanding of the long-term impact of VR-SMT. With only one study (Maciolek et al., 2014) providing data beyond the immediate post-intervention phase, future work should prioritize longitudinal designs incorporating follow-ups extending across operational deployment cycles. Such studies would be strengthened by integrating real-world performance metrics to assess the transfer of training to operational effectiveness.

Another crucial direction involves the personalization of training, as evidence suggests that stress responses are deeply influenced by individual characteristics such as temperament and trauma history. Integrating machine learning approaches and biometric feedback could enable real-time adaptations of scenarios to better reflect the needs of individual users, thereby enhancing training efficacy. Additionally, generalizability remains a concern, as the majority of studies have been conducted with male SOF personnel from Western countries. This narrow participant pool limits the applicability of findings to broader military populations. Future research should strive for greater diversity by including women, non-SOF personnel, and multinational cohorts, as well as by developing culturally sensitive training scenarios tailored to diverse operational environments.

The incorporation of emerging technologies, including AR and neuroimaging methods such as fMRI and EEG, also represents a promising frontier. These tools could help validate the neural mechanisms underpinning stress management and expand the ecological relevance of training through hybrid AR-VR applications. Equally important is the development of ethical and safety frameworks. Despite the high-stress nature of VR-SMT, few studies systematically assess adverse effects, and the potential for retraumatization in individuals with a history of PTSD remains understudied. Establishing safety thresholds and monitoring protocols will be essential to ensure the well-being of participants. Another key consideration is the cost and scalability of VR-SMT systems. The reliance on high-end technology, such as haptic feedback suits, poses significant challenges to widespread adoption. Future studies should explore the feasibility of low-cost and mobile VR platforms and conduct cost-benefit analyses to inform implementation at scale within military structures.

Finally, the field would greatly benefit from enhanced interdisciplinary collaboration. Current research often reveals a disconnect between clinical psychology, military training, and VR engineering. Moving forward, collaborative efforts involving clinicians, veterans, engineers, and military decision-makers will be essential for developing training programs that are both scientifically grounded and operationally relevant. In sum, future research on VR-SMT should focus on standardization, longitudinal efficacy, personalization, inclusivity, and technological integration, while ensuring ethical safeguards and fostering interdisciplinary innovation. Together, these directions will contribute to the development of effective, scalable, and evidence-based interventions for enhancing stress resilience in military populations.

5. Conclusions

This systematic review and Delphi study sought to explore the current landscape of stress management training in military populations using VR, highlighting both empirical evidence and expert consensus. Findings revealed that although the integration of VR technologies into stress management programs has shown promising results—particularly in enhancing realism, engagement, and transferability of skills—the current body of research remains limited in methodological consistency, theoretical grounding, and longitudinal assessment of outcomes. The Delphi study further underscored the need for standardized protocols, ethical frameworks, and interdisciplinary collaboration to ensure the ecological validity and clinical utility of such interventions.

Importantly, this review identified gaps in the literature regarding the specific mechanisms by which VR-based training affects stress resilience and emotional regulation in military settings. The absence of long-term follow-up data and diverse sample representation calls for future research with more rigorous designs, including randomized controlled trials and implementation studies within operational environments. Furthermore, the lack of inclusion of military personnel in co-design processes suggests a critical area for improvement to ensure contextual fit and relevance.

Overall, this work contributes to a growing recognition of the potential of immersive technologies to transform mental health training in military contexts. By consolidating existing evidence and expert perspectives, it provides a foundation for the development of more comprehensive, theory-driven, and ethically informed VR-based stress management programs. Continued investment in research, cross-sector collaboration, and user-centered design will be essential for advancing this field and supporting the psychological readiness of military personnel.

As military institutions continue to prioritize psychological preparedness in increasingly complex and high-stress environments, VR-

SMT offers a compelling supplement to traditional training models. However, realizing its full potential requires sustained investment in research, cross-disciplinary collaboration, and inclusive design practices that reflect the diversity and evolving needs of contemporary armed forces. Through such efforts, immersive stress management may become a cornerstone of proactive mental health strategies in military settings—helping to prepare personnel not only to survive under pressure but to thrive in the face of it.

CRediT authorship contribution statement

Ana Vianez: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **António Marques:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Javier Loureiro:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Paulo Veloso Gomes:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Informed consent statement

Informed consent was obtained from all subjects involved in the study.

Institutional review board statement

The study was approved by Chefe De Estado Maior do Exército with the number GABCEME-RAG-2024-S-00707 and submitted to the Ethics Committee of the School of Health, Polytechnic of Porto with the number CE-0083E.

Funding

This work was supported by Fundação para a Ciência e Tecnologia (FCT) UIDB/05210/2020, and by CITIC, which is subsidized by the Department of Education, Science, Universities, and Vocational Training of the Xunta de Galicia. Additionally, it is co-financed by the European Union through the FEDER Galicia 2021-2027 operational programme (Ref. ED431G 2023/01).

Declaration of competing interest

The authors disclose the sources of funding that supported this work: Funding: This research was funded by Fundação para a Ciência e a Tecnologia (FCT) through R&D unit funding (UIDB/05210/2020). The publication of this article was funded by CITIC, a center accredited for excellence within the Galician University System and a member of the CIGUS Network (Ref. ED431G 2023/01).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chbr.2025.100734.

Data availability

Data will be made available on request.

References

Bacon, B. L., & Staudenmeier, J. J. (2003). A historical overview of combat stress control units of the U.S. army. *Military Medicine, 168*(9), 689–693.

Bartone, P. T., Roland, R. R., Picano, J. J., & Williams, T. J. (2008). Psychological hardiness predicts success in U.S. army special forces candidates. *International Journal of Selection and Assessment, 16*(1), 78–81. https://doi.org/10.1111/j.1468-2389.2008.00412.x

Binnendijk, H. (2002). *Transforming America's military*. National Defense University Press. https://digitalcommons.ndu.edu/ndupress-books/18.

Binsch, O., Bottenheft, C., Landman, A., Roijendijk, L., & Vermetten, E. H. G. J. M. (2021). Testing the applicability of a virtual reality simulation platform for stress training of first responders. *Military Psychology, 33*(3), 182–196. https://doi.org/10.1080/08995605.2021.1897494

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Carl, E., Stein, A. T., Levihn-Coon, A., Pogue, J. R., Rothbaum, B., Emmelkamp, P., Asmundson, G. J. G., Carlbring, P., & Powers, M. B. (2019). Virtual reality exposure therapy for anxiety and related disorders: A meta-analysis of randomized controlled trials. *Journal of Anxiety Disorders, 61*, 27–36. https://doi.org/10.1016/j.janxdis.2018.08.003

Castro, C. A., Adler, A. B., & Britt, T. W. (Eds.). (2006), *The military family: 3. Military life: The psychology of serving in peace and combat*. Security International: Praeger.

Ćosić, K., Popović, S., Horvat, M., Kukolja, D., Dropuljić, B., Kostović, I., Judaš, M., Radoš, M., Radoš, M., & Vasung, L. (2011). Virtual reality adaptive stimulation in stress resistance training. In *Proceedings NATO RTO-MP-HFM-205 on "mental health and well-being across the military spectrum*, 4-1–4-17.

Dolan, C. A., & Adler, A. B. (2006). Military hardiness as a buffer of psychological health on return from deployment. *Military Medicine, 171*(2), 93–98. https://doi.org/10.7205/MILMED.171.2.93

Doran, A. P., Hoyt, G. A. R. Y., & Morgan, C. A. (2012). Survival, evasion, resistance, and escape (SERE) training. In C. H. Kennedy, & E. A. Zillmer (Eds.), *Military psychology: Clinical and operational applications* (2nd ed., p. 306).

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology & Community Health, 52*(6), 377–384. https://doi.org/10.1136/jech.52.6.377

Gaggioli, A., Pallavicini, F., Morganti, L., Serino, S., Scaratti, C., Briguglio, M., Crifaci, G., Vetrano, N., Giulintano, A., Bernava, G., Tartarisco, G., Pioggia, G., Raspelli, S., Cipresso, P., Vigna, C., Grassi, A., Baruffi, M., Wiederhold, B., & Riva, G. (2014). Experiential virtual scenarios with real-time monitoring (interreality) for the management of psychological stress: A block randomized controlled trial. *Journal of Medical Internet Research, 16*(7), Article e167. https://doi.org/10.2196/jmir.3235

Giannarou, L., & Zervas, E. (2014). Using Delphi technique to build consensus in practice. *International Journal of Business Science and Applied Management, 9*(2), 65–82. https://hdl.handle.net/10419/190657.

Houram, L. L., Kizakevich, P. N., Hubal, R., Spira, J., Strange, L. B., Holiday, D. B., Bryant, S., & McLean, A. N. (2011). Predeployment stress inoculation training for primary prevention of combat-related stress disorders. *Cyberpsychology, Behavior, and Social Networking, 4*(1), 101–116. https://doi.org/10.1089/cyber.2010.0084

Hsu, C., & Sandford, B. A. (2007). The Delphi technique: Making sense of consensus. *Practical Assessment, Research and Evaluation, 12*(1), 10. https://doi.org/10.7275/pdz9-th90

Ilnicki, S., Wiederhold, B. K., Maciolek, J., Kosinska, L., Szymanska, S., Zbyszewski, M., Siatkowska, A., Opalko-Piotrkiewicz, E., Ilnicki, P., Filarowska, M., Glibowska, A., Borzetka, D., Pleskacz, K., & Murawski, P. (2012). Effectiveness evaluation for short-term group pre-deployment VR computer-assisted stress inoculation training provided to Polish ISAF soldiers. *Studies in Health Technology and Informatics, 181*, 113–117.

Jex, S. M., Kain, J., & Park, Y. (2013). Situational factors and resilience: Facilitating adaptation to military stressors. In R. R. Sinclair, & T. W. Britt (Eds.), *Building psychological resilience in military personnel: Theory and practice* (pp. 67–84). American Psychological Association. https://doi.org/10.1037/14190-004.

Kavanagh, J. (2005). *Stress and performance in the military: A literature review (technical report)*. RAND Corporation. https://coilink.org/20.500.12592/gps2q4.

Kluge, M. G., Maltby, S., Kuhne, C., Walker, N., Bennett, N., Aidman, E., Nalivaiko, E., & Walker, F. R. (2023). Correction: Evaluation of a virtual reality platform to train stress management skills for a defense workforce: Multisite, mixed methods feasibility study. *Journal of Medical Internet Research, 25*, Article e54504. https://doi.org/10.2196/54504

Koolhaas, J. M., Bartolomucci, A., Buwalda, B., de Boer, S. F., Flügge, G., Korte, S. M., Meerlo, P., Murison, R., Olivier, B., Palanza, P., Richter-Levin, G., Sgoifo, A., Steimer, T., Stiedl, O., van Dijk, G., Wöhr, M., & Fuchs, E. (2011). Stress revisited: A critical evaluation of the stress concept. *Neuroscience & Biobehavioral Reviews, 35*(5), 1291–1301. https://doi.org/10.1016/j.neubiorev.2011.02.003

Kosinska, L., Ilnicki, S., Wiederhold, B. K., Maciolek, J., Szymanska, S., Opalko-Piotrkiewicz, E., Siatkowska, A., Ilnicki, P., Zbyszewski, M., & Glibowska, A. (2012). VR stress inoculation training results for Polish ISAF soldiers – A study of 4 cases. In *NATO science for peace and security series - e: Human and societal dynamics, volume 108: New tools to enhance posttraumatic stress disorder diagnosis and treatment** (pp. 149–160). https://doi.org/10.3233/978-1-61499-189-2-149

Lazarus, R. S. (1993). From psychological stress to the emotions: A history of changing outlooks. *Annual Review of Psychology, 44*, 1–21. https://doi.org/10.1146/annurev.ps.44.020193.000245

Maciolek, J., Ilnicki, S., Wiederhold, B. K., Kosinska, L., Szymanska, S., Zbyszewski, M., Opalko-Piotrkiewicz, E., Siatkowska, A., Glibowska, A., Borzetka, D., Ilnicki, P., Filarowska, M., Pleskacz, K., & Murawski, P. (2014). The influence of pre-deployment VR computer-assisted stress inoculation training on the anxiety level in the Polish ISAF soldiers. *New Tools to Enhance Posttraumatic Stress Disorder Diagnosis and Treatment, 108*, 161–171. https://doi.org/10.3233/978-1-61499-189-2-161. NATO Science for Peace and Security Series - E: Human and Societal Dynamics.

Pair, J., et al. (2006). A virtual reality exposure therapy application for Iraq war post-traumatic stress disorder. *IEEE Virtual Reality Conference (VR 2006)*, 67–72. https://doi.org/10.1109/VR.2006.23

Pallavicini, F., Cipresso, P., Raspelli, S., Grassi, A., Serino, S., Vigna, C., Triberti, S., Villamira, M., Gaggioli, A., & Riva, G. (2013). Is virtual reality always an effective stressor for exposure treatments? Some insights from a controlled trial. *BMC Psychiatry, 13*, 52. https://doi.org/10.1186/1471-244X-13-52

Pols, H., & Oak, S. (2007). War & military mental health: The US psychiatric response in the 20th century. *American Journal of Public Health, 97*(12), 2132–2142. https://doi.org/10.2105/AJPH.2006.090910

Popović, S., Horvat, M., Kukolja, D., Dropuljić, B., & Cosić, K. (2009). Stress inoculation training supported by physiology-driven adaptive virtual reality stimulation. *Studies in Health Technology and Informatics, 144*, 50–54.

Port, C. L., Engdahl, B., & Frazier, P. (2001). A longitudinal and retrospective study of PTSD among older prisoners of war. *American Journal of Psychiatry, 158*(9), 1474–1479. https://doi.org/10.1176/appi.ajp.158.9.1474

Reger, G. M., Gahm, G. A., Rizzo, A. A., Swanson, R., & Duma, S. (2008). *Soldier evaluation of the virtual reality Iraq*. Telemedicine and e-Health. https://doi.org/10.1089/tmj.2008.0050

Rizzo, A. A., Graap, K., Pair, J., Reger, G., Treskunov, A., & Parsons, T. (n.d.). User-centered design driven development of a virtual reality therapy application for Iraq war combat-related post traumatic stress disorder. Institute for creative technologies, University of Southern California.

Robertson, I. T., Cooper, C. L., Sarkar, M., & Curran, T. (2015). Resilience training in the workplace from 2003 to 2014: A systematic review. *Journal of Occupational and Organizational Psychology, 88*(3), 533–562. https://doi.org/10.1111/joop.12120

Shang, Z. (2023). Use of Delphi in health sciences research: A narrative review. *Medicine, 102*(7), Article e32829. https://doi.org/10.1097/MD.0000000000032829

Stanley, E. A., Schaldach, J. M., Kiyonaga, A., & Jha, A. P. (2011). Mindfulness-based mind fitness training: A case study of a high-stress predeployment military cohort. *Cognitive and Behavioral Practice, 18*(4), 566–576. https://doi.org/10.1016/j.cbpra.2010.08.002

Stetz, M. C., Bouchard, S., Wiederhold, B. K., Riva, G., & Folen, R. A. (2009). The receptiveness of stress management techniques by military personnel. *Studies in Health Technology and Informatics, 144*, 125–127.

Stetz, M. C., Kaloi-Chen, J. Y., Turner, D. D., Bouchard, S., Riva, G., & Wiederhold, B. K. (2011). The effectiveness of technology-enhanced relaxation techniques for military medical warriors. *Military Medicine, 176*(9), 1065–1070. https://doi.org/10.7205/milmed-d-10-00393

Stetz, M. C., Long, C. P., Schober, W. V., Jr., Cardillo, C. G., & Wildzunas, R. M. (2007). Stress assessment and management while medics take care of the VR wounded. *Annual Review of CyberTherapy and Telemedicine, 5*, 165–171.

Stetz, M. C., Long, C. P., Wiederhold, B. K., & Turner, D. D. (2008). Combat scenarios and relaxation training to harden medics against stress. *Journal of CyberTherapy and Rehabilitation, 1*(3), Article 239+. https://link.gale.com/apps/doc/A225437106/HRCA.

Thompson, M. M., & McCreary, D. R. (2006). Enhancing mental readiness in military personnel. In A. B. Adler, C. A. Castro, & T. W. Britt (Eds.), *Operational stress: 2. Military life: The psychology of serving in peace and combat* (pp. 54–79). Security International: Praeger.

Winslow, B. D., Carroll, M. B., Martin, J. W., Surpris, G., & Chadderdon, G. L. (2015). Identification of resilient individuals and those at risk for performance deficits under stress. *Frontiers in Neuroscience, 9*, 328. https://doi.org/10.3389/fnins.2015.00328

Zbyszewski, M., Ilnicki, S., Wiederhold, B. K., Maciolek, J., Kosinska, L., Szymanska, S., Opalko-Piotrkiewicz, E., Swierczynska, A., Filarowska, M., Glibowska, A., Borzetka, D., Pleskacz, K., & Murawski, P. (2012). Personality and stress-coping factors in VR computer-assisted stress inoculation training of Polish ISAF soldiers. In **NATO science for peace and security series - e: Human and societal dynamics, volume 108: New tools to enhance posttraumatic stress disorder diagnosis and treatment** (pp. 172–180). https://doi.org/10.3233/978-1-61499-189-2-172