Article

Manipulating Stress Responses during Spaceflight Training with Virtual Stressors

Tor Finseth [1], Michael C. Dorneich [2,*], Nir Keren [3], Warren D. Franke [4] and Stephen B. Vardeman [2,5]

1   Department of Aerospace Engineering, Iowa State University, Ames, IA 50011, USA; tfinseth@iastate.edu
2   Department of Industrial and Manufacturing Systems Engineering, Iowa State University,
     Ames, IA 50011, USA; vardeman@iastate.edu
3   Department of Agricultural and Biosystems Engineering, Iowa State University, Ames, IA 50011, USA;
     nir@iastate.edu
4   Department of Kinesiology, Iowa State University, Ames, IA 50011, USA; wfranke@iastate.edu
5   Department of Statistics, Iowa State University, Ames, IA 50011, USA
*   Correspondence: dorneich@iastate.edu

Abstract: Virtual reality (VR) provides the ability to simulate stressors to replicated real-world situations. It allows for the creation and validation of training, therapy, and stress countermeasures in a safe and controlled setting. However, there is still much unknown about the cognitive appraisal of stressors and underlying elements. More research is needed on the creation of stressors and to verify that stress levels can be effectively manipulated by the virtual environment. The objective of this study was to investigate and validate different VR stressor levels from existing emergency spaceflight procedures. Experts in spaceflight procedures and the human stress response helped design a VR spaceflight environment and emergency fire task procedure. A within-subject experiment evaluated three stressor levels. Forty healthy participants each completed three trials (low, medium, high stressor levels) in VR to locate and extinguish a fire on the International Space Station (VR-ISS). Since stress is a complex construct, physiological data (heart rate, heart rate variability, blood pressure, electrodermal activity) and self-assessment (workload, stress, anxiety) were collected for each stressor level. The results suggest that the environmental-based stressors can induce significantly different, distinguishable levels of stress in individuals.

Keywords: virtual reality; virtual environment; stress; spaceflight; training

1. Introduction

For decades, astronauts have used virtual reality environments (VRE) to practice extra vehicular activity, mass handling, and robotic arm manipulation [1–3]. These VR training scenarios have focused on tasks external to the International Space Station (ISS), rather than tasks on the interior of the station. Full-scale interior mock-ups of the ISS modules and visiting vehicles (e.g., Soyuz capsule) are used to train emergency procedures for fire, depressurization, or contaminant leaks [4,5]. While these training models are effective for task acquisition, the simulation of stressors is constrained by the resources available and physical infrastructure. This constraint may be detrimental as training environments that lack the magnitude of stress felt in a real situation may inadequately prepare individuals for coping [6]. Inadequate training to handle specific stress in the operational environment can result in the diversion of cognitive resources for managing emotional states (e.g., fear, distress, and anxiety), leaving less resources available for task problem solving.

When an individual perceives a situation as stressful, stress will consistently influence physiology and human performance [7,8]. The general impact of exposure to acute high stress is the reduction of cognitive resources available for processing task relevant information as well as reduction in cognitive regulation [9,10]. A deficiency in cognitive resources can affect many performance attributes, including attention, decision time, and short-term

memory [7]. All these performance attributes may be critical for conducting an emergency procedure and avoiding harm. Developing strategies to train individuals for stress can reduce the impact on performance.

The ability to train with stressors in VREs may provide a safe, reliable, and operationally relevant environment for training resilience in hazardous situations. Stressors are defined as stimuli that are appraised as a threat/challenge and elicit a stress response in humans [11]. Past research on stressful operations in VR has included aerial firefighting [12], municipal firefighters [13], first-responders [14], surgeons [15], and military [16]. Emergency operations can be trained for specific simulated tasks or stressors to strengthen transfer to the real world. Some researchers have found that coping skills developed in simulated stressful task environments may be retained and transferred to environments with novel stressors and novel tasks [17]. However, the high rate of individual variability in perception of stress and lack of understanding the underpinning of stressors, make challenges in manipulating human stress responses via exposure to different stressor levels. While some research exists on virtual reality training for spaceflight emergency operations [18–20], more research is needed on the stressors present in the emergency and how manipulating stressors in VEs can approximate the physiological and subjective responses expected during a stress response.

The appraisal of a stressor cannot occur without experiencing a stressful event [21]. Therefore, task training that integrates stress management or graduated stress exposure often opt to experimentally induce stressful events [11]. This is advantageous with VR as the technology allows for experiencing a hazardous environment in a safe and controlled setting while measurements can be collected before the stress induction is divulged to the participant. However, the integration of VR with stress training has been limited by empirical research. As stated by [22], "Most of the studies, in fact, have VR-based stress management training programs only in theory, without providing data about trials conducted to test the effectiveness of the proposed approaches". The authors of [23] found that challenges exist in the selection of effective technologies for delivering stress training, methodological rigor of the training pedagogy, and multi-dimensional assessment of stress. Therefore, more research is needed to verify the efficacy for VR stress training, especially before integration with NASA skill training.

The objective of this study was to assess the extent to which operationally relevant VR stressor levels (low, medium, high) derived from existing emergency spaceflight procedures could evoke a reliable stress response. A panel of experts in spaceflight procedures and human stress response identified relevant stressors and created an experimental emergency procedure. The stressors and procedure were then used to design the VR International Space Station (VR-ISS) spaceflight environment. Participants were trained with the procedures and conducted a VR-ISS fire response for three stressor levels. Self-assessments of stress, workload, and anxiety were administered for each level. Physiological stress measures were used to distinguish varied stress responses. Reliable methods to induce different levels of stress in trainees in a virtual environment would enable training in operationally relevant scenarios such as spaceflight emergency procedures.

2. Background on Manipulating Stress in VR

The stress response induced by a VRE is a product of the stressors employed by the simulation and the cognitive appraisal process. A stressor refers to a specific stimulus within the environment that exert perceived demands on an individual with the outcome. Examples of stressors include noise, pain, task load, ambient temperature, and time pressure [7]. Stimuli are perceived as stressors through the cognitive appraisal process, which evaluates the meaning and significance to individual wellbeing. The appraisal process has two different categories: Primary appraisal, which evaluates the relevance and congruence of a situation with regards to the individual's goals and wellbeing, and secondary appraisal, which involves the evaluation of resources and options for coping with the perceived demand [21]. While the primary and secondary appraisals may be shaped by personal

and situational factors, some researchers have focused on communal stressors related to self-preservation that have a higher likelihood of being appraised as threats [24,25]. Further, these stressors related to self-preservation are thought to be composed of underlying "thematic" stressor elements which may play a pivotal role in how stimuli are cognitively apprised as a stressor.

Of the known stressor elements, much attention has been paid to uncontrollability and unpredictability [24,25]. Uncontrollability manifests with situations or consequences that are outside one's control, difficult to change, and impede progress from attaining one's desired goal [26]. Unpredictability refers to the absence of knowing the conditions under which an event will occur or what an event will be like [27]. The context of the situation (information about outcome, time, stimuli, etc.) and the individual interpretation of the stimuli determine the amount that uncontrollability and unpredictability experienced. Thus, the same stressor manipulated in different contexts may have different amounts of unpredictability and uncontrollability.

Aside from stressor elements, the appraisal process also needs simulations to be immersive enough to recognize a stimulus as an existential threat or challenge. One remarkable aspect of simulated VREs is that individuals tend to respond realistically to virtual simulations [28]. Furthermore, exposure to VR simulations results in a "response-as-if-real" even when the visual fidelity is low, and the representation of the physical reality is significantly reduced. Three concepts have been reported as strong generators of immersive simulations: place illusion, plausibility, and virtual body ownership. Place illusion is the strong sensation of being present in the space generated by the virtual reality environment, even though participants know they are not there [29–32]. Plausibility is the component of presence that is the illusion that the perceived events in the virtual environment are really happening [33]. In comparison to place illusion, which is a static characteristic, plausibility is more concerned with the dynamics of events and the situation portrayed. Lastly, virtual body ownership means utilization of multisensory correlations to provide people the illusion that foreign objects are part of their body [34]. For example, when individuals with virtual hands touch a simulated spider, results show that they find that the virtual hand increased fear [35].

Insight into how to induce stress can be gained from standardized stress tests which have been used by researchers to reliably manipulate stress levels precisely and consistently in a laboratory setting. Several stress tests include public speaking, mental arithmetic, and cold-pressor [36,37]. These stress tests aim to use the most potent stressors possible to elicit a general physiological stress response, often unrelated to the environment the individual is in. However, these lab-based stress manipulations do not involve any operationally-relevant tasks. Potential stressors used for task training should be stressors that could occur in the operational environment in which the individual will perform tasks [38]. These stressors could be related to the environment (e.g., distractions), the task (e.g., increased difficulty), or that state of the human (e.g., fatigue). To identify stressors that are relevant to the operational environment, where the environment is more dynamic and it is harder to anticipate stressors, some researchers have used expert opinion to identify prominent stressors and measured the manipulation using subjective stress ratings and physiological indices of stress [39,40]. However, more research is needed on stressors for operational tasks and how they can be manipulated in VR.

3. Environment Design

To inform how stressors in VR should be designed, a workshop was held with a panel of subject matter experts to identify stressors that contribute to the stress in a spaceflight environment. Subject matter experts have been used in previous research to identify environmental stressors (e.g., noise, radiation) in a disabled submarine scenario [41]. Attending experts included a retired U.S. NASA astronaut, a retired NASA ISS flight director, and three experts on psychological and physiological stress. The panel was presented with existing ISS emergency fire procedures and layouts of the ISS [42]. The panel identified a

series of potential stressors, how they are mapped onto the emergency response procedures, and what effects the stressors might have on astronauts. Each stressor's intensity was rated on how the subject was likely to feel, and what the effects on performance will be. In addition, it was assumed that certain stressors would increase stress levels without influencing task-based workload, so the stressors were categorized into task-related and environment-related stressors. In other words, the stressors during an emergency fire were categorized by their ability to change the environment without substantially changing how the individual would perform the procedure.

A list of stressors was compiled by the panel (Table 1). These stressors are related to an emergency fire procedure, and therefore may exclude peripheral stressors previously identified by subject-matter experts that may be associate with deep-space missions [43]. The stressors were categorized by the type of stress induction, manipulation type (environment or task manipulation), within or between experimental comparison. The process of categorizing environmental stressors was similar to the method proposed by [41]. Stressor levels and potential deviations from the selected emergency procedure were also listed. Three environmental stressors were selected for use in the study and the stressor intensity was varied among training scenarios: alarms, flickering lights, and degraded visibility from smoke. Priority was given to stressor manipulations that did not affect tasks requirements, to avoid later confounds in experimental design. The goal was to increase stress while keeping the task requirements the same to allow direct comparisons of stress training with and without stressor manipulation. By using only environmental stressors and not task stressors, any changes from such a comparison would not be confounded by differences in task requirements. For example, intensity of noise and visible smoke may be stressful, but will not change the task procedure for locating and extinguishing a fire. The selected stressors were then used to establish three stressor levels of low, medium, and high during the VR training scenarios. A few of the task stressors (e.g., rising atmospheric contaminants) were included in all the training scenarios to distribute task load equally.

**Table 1.** Stressors Identified from ISS Emergency Fire Procedure.

Stressor	How to Manipulate	Type of Stress Induction	Scale	Manipulation Type (Task/Envir.)	Within/Between Training Session
Fire Alarm	Magnitude	Divided Attention	Linear	Envir.	Within
	Noise type	Divided Attention	Binary	Envir.	Between
Warning Alarm	Magnitude	Divided Attention	Linear	Envir.	Within
	Noise type	Divided Attention	Binary	Envir.	Between
Caution Alarm	Magnitude	Divided Attention	Linear	Envir.	Within
	Noise type	Divided Attention	Binary	Envir.	Between
Lights	Flashing	Divided Attention	Linear	Envir.	Within
Visibility (smoke obscurity)	Magnitude	Task Difficulty, Time pressure	Linear	Task, Envir.	Within
	Rate of change	Task Difficulty, Time pressure	Linear	Task, Envir.	Within

Table 1. *Cont.*

Stressor	How to Manipulate	Type of Stress Induction	Scale	Manipulation Type (Task/Envir.)	Within/Between Training Session
Multiple Sources (fire, smoke, electrical trips)	Multiple sources in module(s)	Concurrent Task Mgmt.	Binary	Task	Between
Location of smoke/fire	Source in Soyuz	Task Difficulty	Binary	Task	Between
	Source separating crew	Task Difficulty, Empathy, Difficult Tradeoff	Binary	Task	Between
	Source in module(s)	Task Difficulty	Binary	Task	Between
Open Flame	Magnitude	Visual Threat, Task Difficulty, Time pressure	Linear	Task	Within
	Spread rate	Visual Threat, Task Difficulty, Time pressure	Linear	Task	Within
Oxygen/Emer. Mask	Limited oxygen supply	Time Pressure	Linear	Task	Within
	No oxygen, faulty	Task Difficulty	Binary	Task	Between
	Reduced peripheral vision	Task Difficulty	Linear	Task	Between
	Clarity of other crew's voices	Task Difficulty	Linear	Task	Within
Contaminants	Magnitude	Threat, Task Difficulty	Linear	Task	Within
	Rate	Time Pressure	Linear	Task	Within
Team Member	Language, accent	Task Difficulty	Linear	Task	Between
	Experience of crew	Task Difficulty	Linear	Task	Between
	Location of crew	Task Difficulty	Linear	Task	Between
	Sleeping crew	Task Difficulty	Binary	Task	Between
	Reaction of crew	Task Difficulty	Linear	Task	Within
Comm Operation	Comm delay	Task Difficulty	Linear	Task	Between
	No Comm	Task Difficulty	Binary	Task	Between
	No lights (power outage)	Concurrent Task Mgmt.	Binary	Task	Within
Power Outage	Compromised life support	Concurrent Task Mgmt., Time pressure	Binary	Task	Between
MCC Communication	Frequency of communication	Concurrent Task Mgmt.	Linear	Task	Between
	Inconsistent instructions	Task Difficulty	Linear	Task	Between
Fire Extinguishers	Limited PFEs available	Task Difficulty	Linear	Task	Between
	Limited PFE uses	Task Difficulty	Linear	Task	Between

The panel then developed a simplified fire procedure for trainees to follow. A flowchart of the selected emergency procedure that could be used for laboratory experiments is illustrated in Figure 1. This procedure was modified from the existing ISS Emergency Procedures [42] by shortening the duration of the procedure down to 5–10 min, eliminating communication with Mission Control Center (MCC) and crewmember induced stressors, having the smoke alarm indicate the presence of a fire (as opposed to being a false alarm), and minimizing tangential procedure steps that do not directly help the trainee locate the

fire source. This procedure is also like the ISS fire strategy created by [18] for the purpose of task training. Further, this procedure was designed to have limited decision nodes to avoid task branches and maintain experimental replicability.
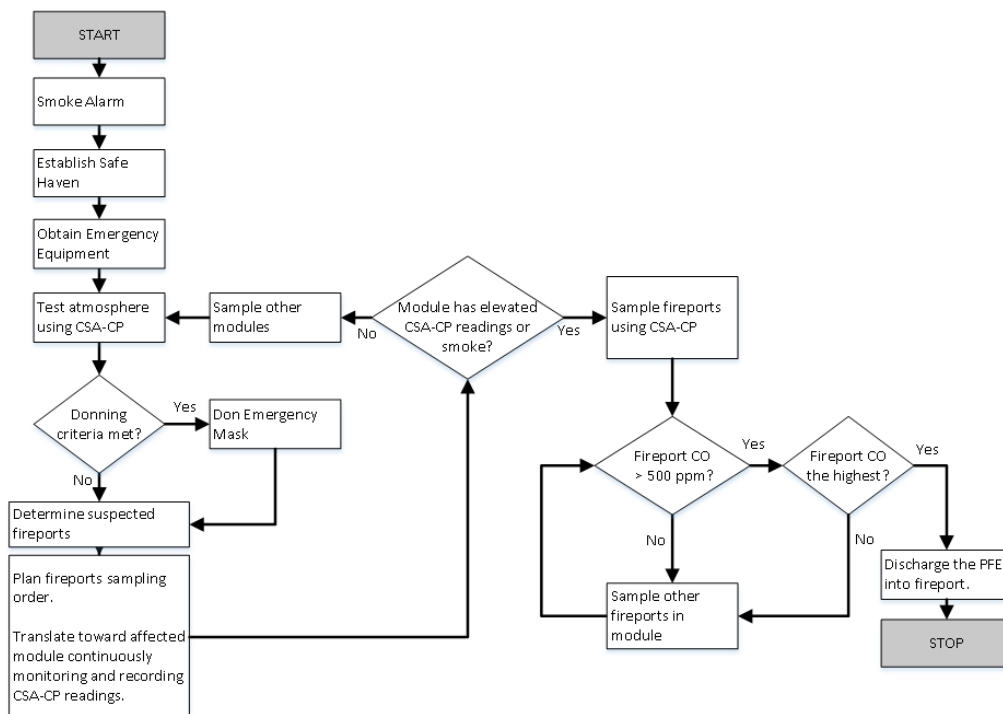

Figure 1. VR-ISS emergency fire procedure steps created by the workshop, modified from existing NASA ISS Emergency Fire Procedures.

This procedure begins with an auditory smoke alarm alerting crew to the presence of a fire. A safe haven is to be established in a location closest to the point of egress (e.g., Soyuz capsule) where crewmembers can formulate a response plan and assign duties. Emergency equipment vital to the fire response is obtained from storage compartments. Readings of contaminant levels are reviewed with a NASA Compound Specific Analyzer–Combustion Products (CSA-CP). The purpose of the CSA-CP onboard the ISS is to determine the level of atmospheric contaminants, including carbon monoxide (CO) in parts per million. If CO levels are rising, crewmembers should don oxygen masks. Based on CSA-CP readings, suspected fireports should be determined and crewmembers should translate from module to module, then narrow the search to sample local fireports as contaminant levels increase with proximity. If the sampled fireport CO > 500 and increasing, then discharge the portable fire extinguisher (PFE).

To integrate this procedure into VR, the stressors and procedure were then used to design the VR-ISS spaceflight environment. The VR-ISS environment is based on 3D models from [44] and [19] but has been largely modified from its former state to be used with VR head mounted display (HMD) and facilitate spaceflight procedure training. To simplify training participants, the VR-ISS consisted only of three of the existing U.S. Orbital Segment modules, Figure 2 illustrates the VR-ISS configuration used in the experiment: only Node 1, US Lab, and Node 2 were used.

An avatar was used to increase virtual body ownership, illustrated in Figure 3. The avatar's hands tracked VR hand controllers. Zero gravity mocking locomotion is integrated into the simulation to allow participants to float across the VE by grabbing and pushing against objects/surfaces.
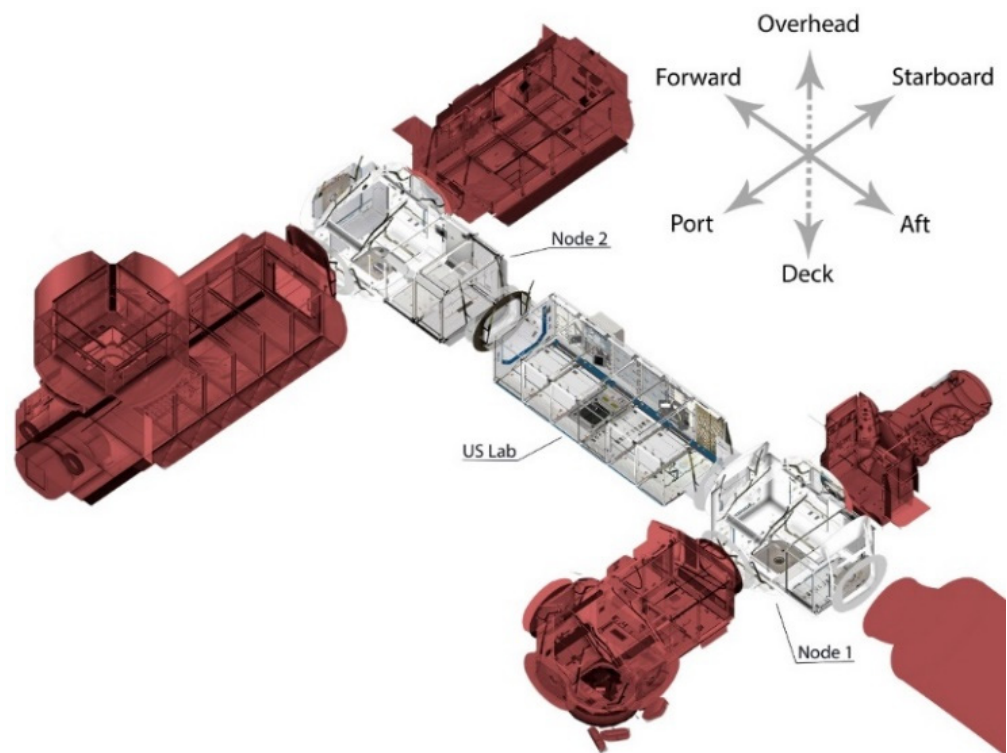
Figure 2. VR-ISS configuration. Some sections of the ISS U.S. and Russian segment were not included in the simulation reproduced with permission from [45], IEEE, 2020.


Figure 3. Third person view of participant's avatar in the VR-ISS.

Several dynamic interactions were included in the VR-ISS to aid detecting and locating the source of a fire. These dynamic functions were designed to enhance the immersion through increased place illusion and plausibility. Atmospheric contaminant levels rose as a function of time and distance from the virtual fire source. Since, the readings change as a function of time, the participant must not only remember previous readings at different

locations to compare with the current reading, but also account of the fact that readings become less reliable the further back in time it was taken. Virtual smoke changed in density as a function of time and spread in a uniform pattern, consistent with expected smoke behavior in a microgravity; therefore, participants could not rely solely on visual smoke patterns to detect the location of the source.

Emergency equipment was also simulated for plausibility and place illusion. Readings of contaminant levels could be collected with a simulated NASA CSA-CP. The purpose of the CSA-CP onboard the ISS is to determine the level of atmospheric contaminants. Virtual CSA-CP displayed levels of oxygen, carbon monoxide (CO), hydrogen chloride (HCl), and hydrogen cyanide (HCN) in parts per million (Figure 4). Using voice commands, a floating CSA-CP appeared in front of the participant with the contaminant concentration values visible. The window disappeared after three seconds. Participants are expected to identify the location of the source by following the invisible path mentally established from recalling highest levels of contaminants in each VR-ISS module. A recording at the start of the simulation gave instructions to retrieve a Portable Fire Extinguisher (PFE) and Portable Breathing Apparatus (PBA) when the contaminate levels are excessive (Figure 4). The PFE is used to extinguish a fire source behind a rack fireport. The PFE has the capability for two uses before the canister is empty. Five PFEs are available in cabinets in the VR-ISS. PBAs are available in the same cabinets and can be done on the participant avatar's head. A Caution and Warning (C&W) panel displayed flashing lights to alert participants to a potential fire (Figure 4). Once the participants identified where the fire source was located, they began sampling fireports within the module to locate the "rack" that caused the fire. The VR-ISS included approximately 150 fireport labels, accurately placed on the racks throughout the ISS.

(**A**)	(**B**)	(**C**)	(**D**)
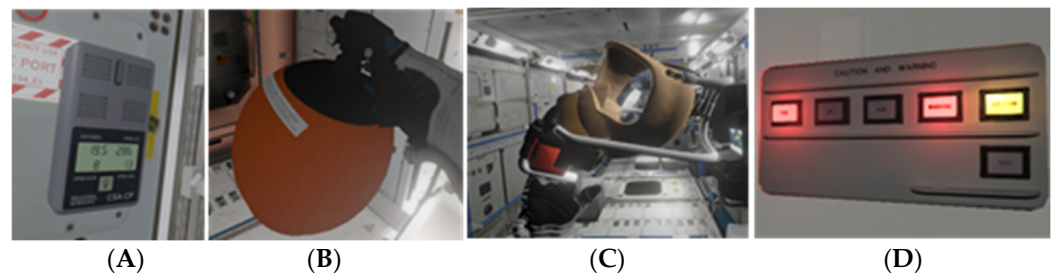
Figure 4. VR-ISS emergency fire equipment, (**A**) Compound Specific Analyzer–Combustion Products next to a fireport label, (**B**) Portable Fire Extinguisher, (**C**) Portable Breathing Apparatus, and (**D**) Caution & Warning Panel reproduced with permission from [45], IEEE, 2020.

4. Materials and Methods

4.1. Participants

Forty subjects participated (Male = 37, Female = 3). The study was reviewed and approved by Iowa State University Institutional Review Board. The participant mean age was 20.5 years (SD = 2.6). Thirty-eight of the participants in this study had a STEM background.

4.2. Task Environment

VR-ISS is the operational environment for this experiment. Participants were tasked with locating and extinguishing the location of a potential fire on the VR-ISS. Several dynamic interactions were included in the VR-ISS to aid detecting and locating the source of a fire. Atmospheric contaminant levels rose as a function of time and distance from the virtual fire source. Virtual smoke changed in density as a function of time and spread in a uniform pattern, consistent with expected smoke behavior in a microgravity; therefore, participants could not rely solely on visual smoke patterns to detect the location of the source.

Participants reviewed readings of contaminant levels with CSA-CP. When participants identified the fireport which had the highest level of contaminant and extinguished the fire with the PFE, the contaminate levels were reset and a new randomized fire source was created. The task ended five minutes after the beginning of the simulation.

4.3. Independent Variables

The VE was designed with three environmental stressors identified from the panel SMEs: alarms, flickering lights, and degraded visibility from smoke. The smoke filled the modules rapidly until the intensity level was achieved, but also moved dynamically through the module to increase plausibility. The lights for selected modules flickered randomly to simulate temporary power failure, presenting the participant with temporary near-complete darkness. A fire alarm and caution alarm were used for different levels.

The simulation had three different stress levels, each with a fire location randomized (Figure 5). The low stress level indicated a fire using increased CSA-CP contaminate values and C&W panel lights. The medium stress level indicated a fire using increased CSA-CP contaminate values, C&W panel lights, a continuous Caution alarm, and low levels of smoke (visibility of 6 ft). The high stress level indicated a fire using increased CSA-CP contaminate values, C&W panel lights, a continuous Caution alarm, a continuous Fire alarm, flickering lights, and high levels of smoke spread over time (visibility of 1 ft).
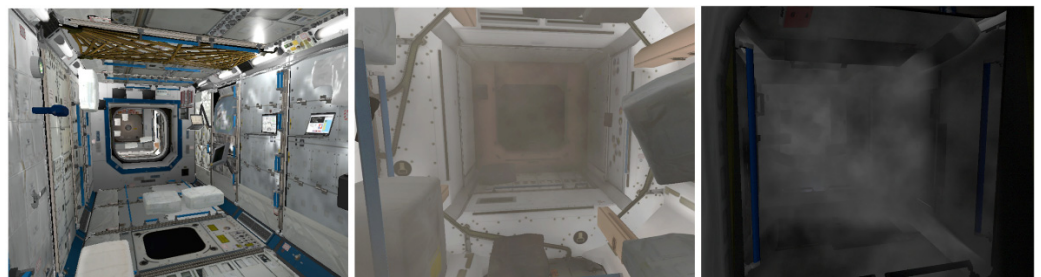
Figure 5. VR-ISS emergency fire (low, medium, high stressor scenarios). Reproduced with permission from [45], IEEE, 2020.

4.4. Dependent Variables

The study used both physiological and subjective indices of stress. The dependent variables are summarized in Table 2.

Table 2. Description of dependent variable sensor/metrics and frequencies.

Dependent Variable	Sensor/Metric	Features/Type	Description, Association	Frequency
Physiological stress response	Electrocardiogram (ECG) and Heart Rate Variability (HRV)	HR	Heart rate	Throughout trial
		RMSSD	The root mean of the sum of the squares of differences between beat intervals	
		pNN50	Proportion of the successive normal to normal beat intervals that differ more than 50 ms	
Physiological stress response	Continuous Non-Invasive Blood Pressure	SBP	Systolic Blood Pressure	Throughout trial
		DBP	Diastolic Blood Pressure	
Physiological stress response	Electrodermal Activity (EDA)	EDA	Tonic component (0–0.16 Hz)	Throughout trial
Subjective Stress	Post Stress Task Reaction Measure (PTSR)	9-point Likert scale		After trial
Subjective Stress	Free Stress Comparison	100-point scale		After experiment
Subjective Stress	Short State Stress Questionnaire (SSSQ)	24 items, 5-point Likert scale	Engagement, Distress, Worry	Before, after trial
Workload	NASA TLX	6 items, 21-point scale	Mental demand, Physical demand, Temporal demand, Performance, Effort, and Frustration	After trial
Anxiety	State-trait anxiety inventory (STAI)	20 items, 7-point Likert scale		After trial

4.4.1. Heart Rate and Heart Rate Variability

Heart rate increases with higher physiological stress response. Heart rate data were collected via electrocardiogram (ECG, modified CS5 lead configuration). The ECG was sampled at 2048 Hz using Biopac MP150 hardware and recorded using AcqKnowledge software (Version 3.8.2, Biopac Systems Inc.). Heart Rate is influenced by both sympathetic and parasympathetic branches of the autonomic nervous system [46]. Activity in these branches is indicative of the body's state of excitation vs. relaxation (vagal response).

Time domain analysis of the ECG was performed to quantify changes in Heart Rate Variability (HRV). The vagal response, which is indicative of relaxation, was assessed with two HRV metrics relating to the amount of variance in the inter-beat-interval through the root-mean-square difference of successive normal R-wave intervals (RMSSD) and the proportion of the number of pairs of successive normal R-waves that differ by more than 50 ms (pNN50). Increased pNN50 and RMSSD indicated a more relaxed state and greater ability to cope with increased stress. More specifically, both RMSSD and pNN50 represent vagal control within the time domain and are correlated to high frequency power [46]. The HRV time domain for each participant were calculated in 30-s intervals over the duration of each session.

4.4.2. Blood Pressure

Systolic blood pressure (SBP) and diastolic blood pressure (DBP) were collected as another measure of cardiovascular reactivity. Both SBP and DBP have been shown to increase in reaction to increased stress [47]. A finger cuff was placed on the participants' non-dominant hand over the middle phalanx of either the long or ring finger (CNAP Monitor 500, CNSystems Medizintechnik AG). The non-dominant arm was placed in an arm-sling to standardize the position of the hand relative to the heart between all participants. Data were recorded at 1024 Hz. To calibrate the finger cuff, an oscillometric non-invasive blood pressure cuff was placed on the participant's non-dominant upper arm. The CNAP blood pressure readings were calibrated before the start of each trial. The participant remained seated for the duration of the experiment to prevent orthostatic pressure changes.

4.4.3. Electrodermal Activity

The electrodermal activity (EDA) signal increases under higher acute stress [48]. EDA is sympathetically mediated, therefore can be used to verify autonomic influences on heart rate. The EDA signal was corrected with an IIR low pass 2nd-order Butterworth filter fixed at 5 Hz. The EDA tonic component, also called skin conductance level, was extracted by low pass filtering with a cut-off frequency of 0.16 Hz and used as an indicator of sympathetic activity [49]. Data were recorded in 30 s epochs for statistical analysis.

4.4.4. Subjective Stress

Two kinds of ratings were used to assess the perception of low, medium, and high stress: Post Task Stress Reaction (PTSR) [50], Free Stress Comparison of events. The PTSR and Free Stress Comparison questionnaires were modified from previous uses [40,47,51]. The PTSR asks participants to rate the ground truth simulations on a scale of "1" to "9" where a rating of "1" was used to represent experiencing "no stress", "5" was used to represent "medium stress", and "9" was used to represent experiencing "high stress". The PTSR is intended to measure the immediate retrospective stress after completing a trial.

In addition, the Free Stress Comparison has participants rate the relative stress level on a scale of 0 to 100 (least to most stressful) in comparison to other simulations. The stress appraisal process is continuous and relative, with reappraisals of the experience happening long after the stressful exposure [52]. The Free Stress Comparison was intended to measure the retrospective reappraisal after completing all the simulations, then relatively comparing their stressfulness.

4.4.5. Subjective Stress State

The Short Stress State Questionnaire (SSSQ) [53] assessed the subjective states pre- and post-trial to measure three state factors: task engagement, distress, and worry. Engagement refers to qualities of energetic arousal, motivation, and concentration. Distress is defined as feelings of tense arousal, hedonic tone, and confidence-control. Worry relates to self-focus, self-esteem, and cognitive interference [54]. The stress state acts as a mediator between the stressor and cognition or information processing, whereby the three aspects represent components of conscious experience during person–task–environment transactions [55].

4.4.6. Workload

The NASA Taskload Index (TLX) [56] was used to assess the subjective workload during exposure. The NASA TLX measures six dimensions of workload: mental demand, physical demand, temporal demand, performance, effort, and frustration level. NASA TLX was administered after the completion of a trial. Participant scores on the six numerical rating scales were computed in the 0 to 100 range and as an unweighted participant mean for each of the six-dimensional subscales [57].

4.4.7. Anxiety

To measure the anxiety during the experiment, State-Trait Anxiety Inventory (STAI) [58] Form Y-1 was given after the completion of a trial. The STAI-Y1 is a 20-item self-report scale for the assessment of state anxiety in adults. Based on the answers, the STAI score can be interpreted in the ranges: no stress <30, low <40, medium 40–55, high >55 [59].

4.5. Experiment Design

A within-subject 1×3 (trial) experiment was conducted. Each participant completed the same task of locating an onboard fire, but each trial had one of three different stressor levels (low, medium, and high). The order of stressor levels was assigned via Latin square to counterbalance and minimize the effect of training order's influence from differences among the treatment effects (i.e., each levels stressors) [60].

4.6. Hypothesis

It was hypothesized that manipulating different levels of VR environmental stressors (low, medium, and high) during an operational task would induce different levels of stress, with higher stressor levels resulting in greater increases in HR, decreases in RMSSD, decreases in pNN50, increases in SBP and DBP, increases in EDA, increases in PTSR, increase in Free Stress Comparison, no change in engagement, increases in distress, no change in worry, no change in workload, and increases in anxiety.

4.7. Procedure

The experiment was completed in a single laboratory visit, lasting approximately 60 min. At the beginning of the experiment, participants completed a series of pre-trial questionnaires, including demographic questions, a SSSQ to measure the stress in response immediately before the trials, and training on how to use the NASA TLX. To acclimate to VR before the data collection tasks, participants were trained on navigating, operating, and controlling the VR simulation (e.g., head-mounted display, hand controls, "play-area" boundaries represented but a visual blue-grid). Participants were asked to report cybersickness and were withdrawn if symptoms persisted.

For the VR-ISS, participants completed a VR interactive tutorial that included information about the ISS layout, how to navigate, fire equipment, and the appropriate emergency fire response. Participants practiced the procedure in the tutorial until memorized.

Participants then completed three trials: low, medium, and high stressor levels. After each trial, participants completed several questionnaires, including the post-trial SSSQ, NASA TLX, STAI, and PTSR. Participants were given 5–10 min between trials to complete

questionnaires. At the end of the experiment, participants completed the Free Stress Comparison to compare the trials in retrospect.

4.8. Experiment Materials

The apparatus consisted of two parts: an HTC VIVE (professional version; HTC, 2016) consumer VR headset. The Unity (5.4.0f3, Unity Technologies, 2014) 3D game engine was used to facilitate all aspects of the VR-ISS as a virtual environment. The HTC VIVE setup consists of the HMD and two Lighthouse sensors that are responsible for tracking the headset position and orientation. For this experiment, the lighthouse sensors were positioned facing each other at opposite ends of our lab space, 8 ft high with 12×12 ft detectable play area.

4.9. Data Analysis

Data analysis was performed using SPSS software (Version 28.0; IBM Corp.). Distributions were tested for normality using skewness and kurtosis divided by the standard error and concluded to be normal if less than 1.96 [61]. Each dependent variable was visually inspected for errors (e.g., signal artifacts, miscalibration, electrode disconnect) and participant trials were omitted if they displayed erratic patterns. Alterations were made to discontinue some questionnaires during the experiment, resulting in analysis on only a subset of the sample. Repeated measure analysis of variance (RM-ANOVA) was used to calculate the fixed effect of stressor level. Physiological data were standardized against the first 90-s of the low stressor trial to emulate change from baseline, to which the 90-s were then omitted from the analysis. The first 90-s was also omitted from the medium and high stressor trials to account for the physiological transition in response to a stressor. To account for correlation of physiological data, autoregressive (AR1) models were used for the covariance matrix with a participant random effect. AR1 model fit was assessed with Akaike's Information Criteria (AIC). Significant differences were located using pairwise comparisons, and acceptance level was adjusted to control for type I errors (Bonferroni adjustment). Results were considered significant for $p \leq 0.05$ and marginally significant for $p < 0.10$ [62]. The effect sizes, given in partial eta squared for the models, were transformed, and reported as Cohen's d effect size under the conservative assumption the stress level means are separated with maximum variability [63]. Cohen's d was used for assessing effect size, where $0.2 < |d| < 0.5$ is considered small effect size, medium effect size when $0.5 < |d| < 0.8$, and large effect size for $|d| > 0.8$ [63].

The three-factor SSSQ scale scores for pre- and post-trial were calculated for each participant. The factor scores from both pre- and post-trial are standardized against normative means and standard deviation values from a large sample of British participants [64] and standardized using methods in [55]. Change scores were calculated for each factor using the z-score Formula (1) which has been used in previous studies [55]. The z-score then represents the change between pre- and post-trial in units of the deviation from the population mean.

$$z = (\text{standardized post-score} - \text{standardized pre-score}). \tag{1}$$

5. Results

A summary of the results is provided in Table 3, with greater description of the results in the following subsections.

5.1. Heart Rate and Heart Rate Variability

The main effect of stressor level on the change in heart rate ($N = 35$) was significant, $F_{(2,63.7)} = 4.34$, $p = 0.017$, with a large effect size, $d = 0.90$ (Figure 6a). Pairwise comparison indicated the change in heart rate was significantly higher ($p = 0.017$, $d = 0.58$) for participants in high stressor ($M = 1.41$, $SD = 3.05$) compared to low stressor ($M = 0.37$, $SD = 1.42$), but not significantly different ($p = 0.16$) for high stressor compared to medium stressor

($M = 0.58$, $SD = 1.59$) and not significantly higher ($p = 0.99$) the for medium stressor compared to low stressor.

Table 3. Summary of Results.

| Metric	N	Low M (SD)	Medium M (SD)	High M (SD)	Main Effect F (p)	Low vs. Medium p (d)	Low vs. High p (d)	Medium vs. High p (d)
Physiological Measures								
ΔHR	35	0.37 (1.42)	0.58 (1.59)	1.41 (3.05)	4.34 (0.017)	0.99	0.017 (0.58)	0.16
ΔRMSSD	35	0.63 (3.07)	0.48 (2.78)	−0.25 (2.44)	3.29 (0.04)	0.99	0.048 (0.37)	0.18
ΔpNN50	35	−0.02 (1.42)	0.13 (1.81)	−0.076 (1.51)	0.079 (0.92)	NA	NA	NA
ΔSBP	32	0.95 (2.44)	1.94 (4.11)	2.04 (6.09)	2.84 (0.066)	0.26	0.08 (0.43)	0.99
ΔDBP	32	1.09 (2.91)	2.05 (6.20)	0.82 (3.70)	0.69 (0.51)	NA	NA	NA
ΔEDA	32	−0.82 (2.45)	1.94 (4.42)	1.29 (3.97)	7.05 (0.002)	0.002 (0.71)	0.046 (0.42)	0.71
Subjective Measures								
PSTR	39	4.67 (1.61)	5.26 (1.6)	6.34 (1.68)	27.9 (<0.001)	.11	<0.001 (1.0)	<0.001 (0.66)
Free Stress	39	30.4 (20.5)	45.9 (17.6)	78.7 (18.8)	80.6 (<0.001)	<0.001 (0.81)	<0.001 (2.46)	<0.001 (1.8)
ΔEngage	23	−0.024 (0.51)	0.089 (0.61)	0.012 (0.65)	0.47 (0.63)	NA	NA	NA
ΔDistress	23	0.5 (0.79)	0.39 (0.59)	0.82 (0.94)	7.05 (0.002)	0.73	0.072 (0.37)	0.01 (0.55)
ΔWorry	23	−0.28 (0.73)	−0.38 (0.59)	−0.22 (0.8)	3.13 (0.054)	0.3	0.99	0.096 (0.23)
Workload	39	45.1 (17.8)	46.9 (17.4)	57.5 (17.6)	12.5 (<0.001)	0.99	0.002 (0.7)	<0.001 (0.61)
Anxiety	23	37.5 (9.76)	36.2 (9.58)	40.6 (10.7)	5.25 (0.009)	0.92	0.24	0.005 (0.43)
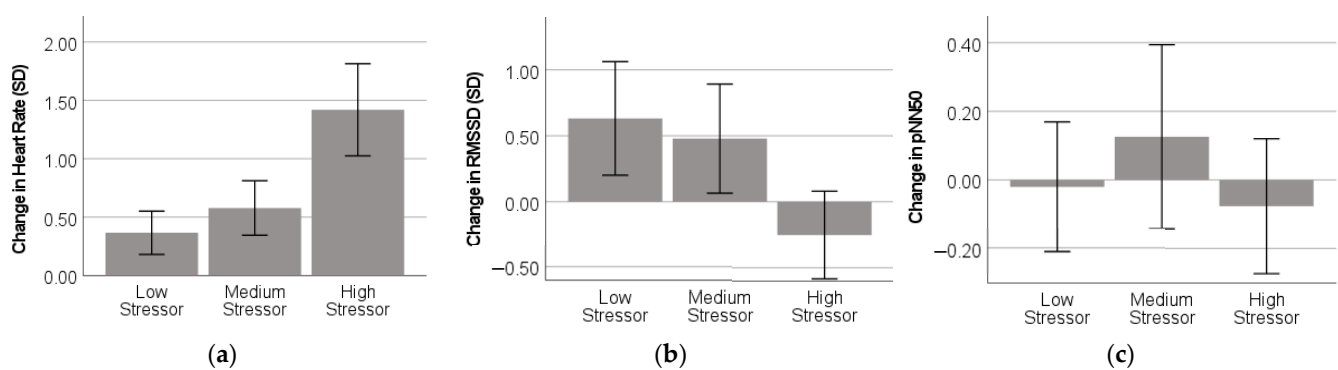
Figure 6. Heart rate metrics: (**a**) Heart Rate; (**b**) RMSSD; (**c**) pNN50. Error bars representing 95% confidence intervals.

The main effect of stressor level on the change in RMSSD ($N = 35$) was significant, $F(2,170) = 3.29$, $p = 0.04$, with a medium effect size, $d = 0.48$ (Figure 6b). Pairwise comparison indicated the change in RMSSD was significantly lower ($p = 0.048$, $d = 0.37$) for participants in high stressor ($M = −0.25$, $SD = 2.44$) compared to low stressor ($M = 0.63$ $SD = 3.07$), but not significantly different ($p = 0.18$) for high stressor compared to medium stressor ($M = 0.48$, $SD = 2.78$) and not significantly higher ($p = 0.99$) the for medium stressor compared to low stressor.
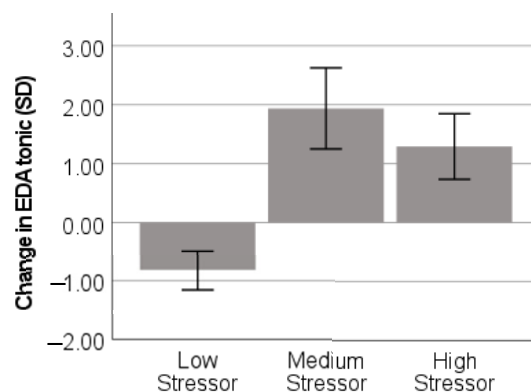
The main effect of stressor level on the change in pNN50 ($N = 35$) was not significant, $F(2,185) = 0.079$, $p = 0.92$ (Figure 6c) for the low stressor ($M = -0.02$, $SD = 1.42$), medium stressor ($M = 0.13$, $SD = 1.81$), or high stressor ($M = -0.076$, $SD = 1.51$).

5.2. Blood Pressure

The main effect of stressor level on the change in SBP ($N = 32$) was marginally significant, $F(2,60.5) = 2.84$, $p = 0.066$, with a medium effect size, $d = 0.61$ (Figure 7a). Pairwise comparison indicated the change in blood pressure was marginally higher ($p = 0.08$, $d = 0.43$) for participants in high stressor ($M = 2.04$, $SD = 6.09$) compared to low stressor ($M = 0.95$, $SD = 2.44$), but not significantly higher ($p = 0.99$) for high stressor compared to medium stressor ($M = 1.94$, $SD = 4.11$) and not significantly higher ($p = 0.26$) for medium stressor compared to low stressor.
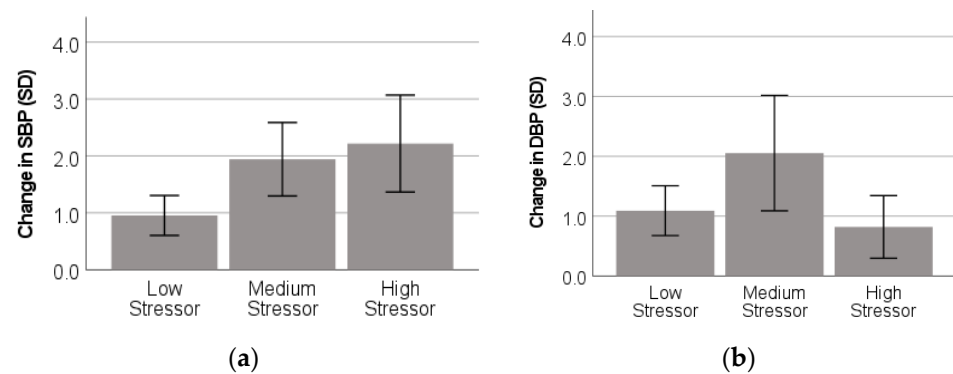
Figure 7. Blood pressure metrics: (**a**) systolic blood pressure (SBP); (**b**) diastolic blood pressure (DBP). Error bars representing 95% confidence intervals.

The main effect of stressor level on the change in DBP ($N = 32$) was not significant, $F(2,70.4) = 0.69$, $p = 0.51$ (Figure 7b) for the low stressor ($M = 1.09$, $SD = 2.91$), medium stressor ($M = 2.05$, $SD = 6.20$), or high stressor ($M = 0.82$, $SD = 3.70$).
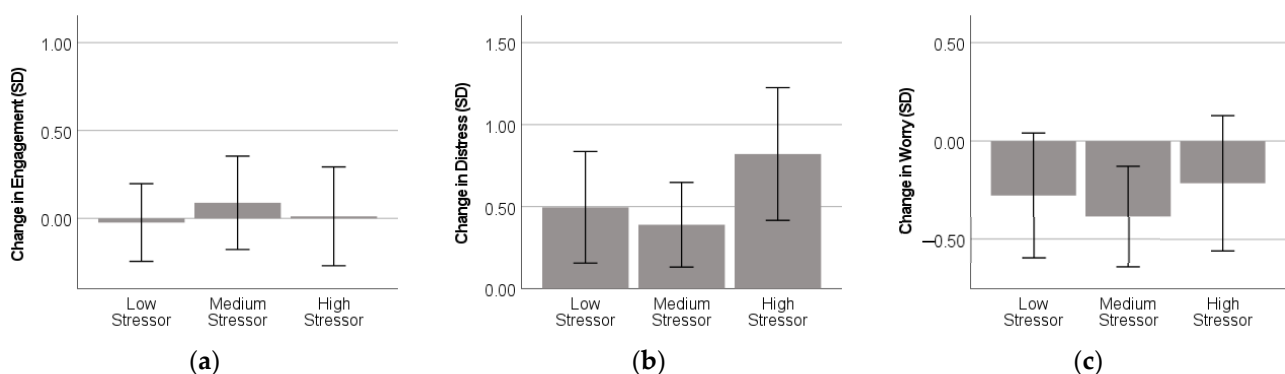
5.3. EDA Tonic

The main effect of stressor level on the change in EDA tonic ($N = 32$) was significant, $F(2,64.6) = 7.05$, $p = 0.002$, with a large effect size, $d = 0.93$ (Figure 8). Pairwise comparison indicated the change in EDA was significantly higher ($p = 0.046$, $d = 0.42$) for participants in high stressor ($M = 1.29$, $SD = 3.97$) compared to low stressor ($M = -0.82$, $SD = 2.45$), and significantly higher ($p = 0.002$, $d = 0.71$) for the medium stressor (M = 1.94, SD = 4.42) compared to low stressor, but not significantly different ($p = 0.71$) for high stressor compared to medium stressor.

Figure 8. Electrodermal Activity tonic component. Error bars representing 95% confidence intervals.

5.4. Subjective Stress

The main effect of stressor level on subjective stress measured by PTSR ($N = 39$) was significant, $F(2,76) = 27.9$, $p < 0.001$, $d = 1.71$ (Figure 9a). Pairwise comparison indicated the subjective stress was significantly higher ($p < 0.001$, $d = 1.0$) for participants in high stressor ($M = 6.34$, $SD = 1.68$) compared to low stressor ($M = 4.67$, $SD = 1.61$), significantly higher ($p < 0.001$, $d = 0.66$) for high stressor compared to medium stressor ($M = 5.26$, $SD = 1.6$), but not significantly different ($p = 0.11$) for medium stressor compared to low stressor.
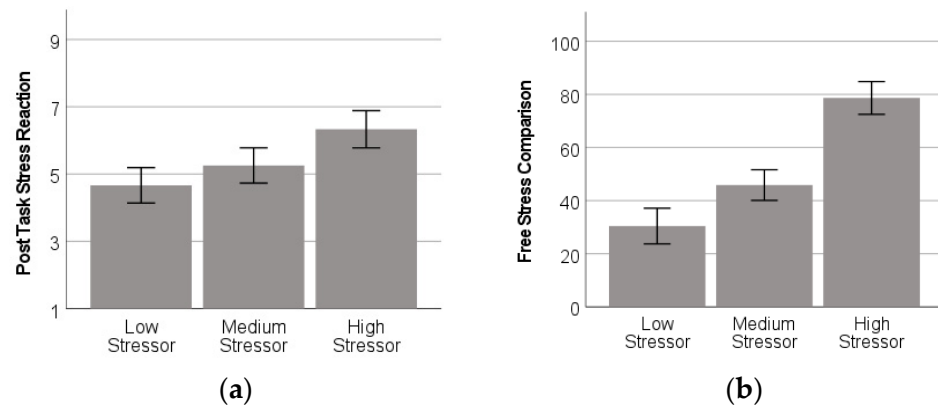
Figure 9. Subjective stress measures: (**a**) Post Task Stress Reaction (PTSR); (**b**) Free Stress Comparison. Error bars representing 95% confidence intervals.

The main effect of stressor level on subjective stress measured by Free Stress Comparison ($N = 39$) was significant, $F(2,76) = 80.6$, $p < 0.001$, $d = 2.91$ (Figure 9b). Pairwise comparison indicated the subjective stress was significantly higher ($p < 0.001$, $d = 2.46$) for participants in high stressor ($M = 78.7$, $SD = 18.8$) compared to low stressor ($M = 30.4$, $SD = 20.5$), significantly higher ($p < 0.001$, $d = 1.8$) for high stressor compared to medium stressor ($M = 45.9$, $SD = 17.6$); and significantly higher ($p < 0.001$, $d = 0.81$) for medium stressor compared to low stressor.

5.5. SSSQ: Stress State

The main effect of stressor level on the change in engagement ($N = 23$) was not significant, $F(2,44) = 0.47$, $p = 0.63$ (Figure 10a) for the low stressor ($M = -0.024$, $SD = 0.51$), medium stressor ($M = 0.089$, $SD = 0.61$), or high stressor ($M = 0.012$, $SD = 0.65$).

Figure 10. The change in stress state measured by the SSSQ: (**a**) Engagement; (**b**) Distress; (**c**) Worry. Error bars representing 95% confidence intervals.

The main effect of stressor level on the change in distress ($N = 23$) was significant, $F(2,44) = 7.05$, $p = 0.002$, with a large effect size, $d = 1.13$ (Figure 10b). Pairwise comparison indicated the change in distress was marginally significantly higher ($p = 0.072$, $d = 0.37$) for participants in high stressor ($M = 0.82$, $SD = 0.94$) compared to low stressor ($M = 0.50$,

$SD = 0.79$), significantly higher ($p = 0.01$, $d = 0.55$) for high stressor compared to medium stressor ($M = 0.39$, $SD = 0.59$), but not significantly different ($p = 0.73$) for medium stressor compared to low stressor.

The main effect of stressor level on the change in worry ($N = 23$) was marginally significant, $F(2,44) = 3.13$, $p = 0.054$, with a medium effect size, $d = 0.76$ (Figure 10c). Pairwise comparison indicated the change in worry was marginally significantly higher ($p = 0.096$, $d = 0.23$) for participants in medium stressor ($M = -0.38$, $SD = 0.59$) compared to high stressor ($M = -0.22$, $SD = 0.8$), but not significantly different ($p = 0.30$) for medium stressor compared to low stressor ($M = -0.28$, $SD = 0.73$), and not significantly different ($p = 0.99$) for the high stressor compared to low stressor.

5.6. NASA-TLX: Workload

The main effect of stressor level on workload ($N = 39$) was significant, $F(1.63,62.1) = 12.5$, $p < 0.001$, with a large effect size, $d = 1.15$ (Figure 11). Pairwise comparison indicated the workload was significantly higher ($p = 0.002$, $d = 0.7$) for participants in high stressor ($M = 57.5$, $SD = 17.6$) compared to low stressor ($M = 45.1$, $SD = 17.8$), significantly higher ($p < 0.001$, $d = 0.61$) for high stressor compared to medium stressor ($M = 46.9$, $SD = 17.4$), but not significantly ($p = 0.99$) different for medium stressor compared to low stressor.
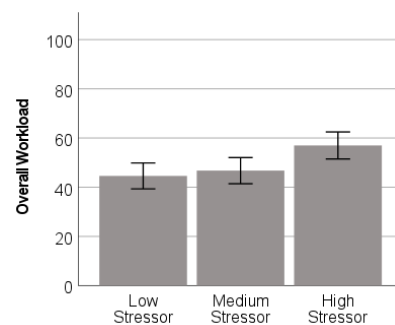
Figure 11. Overall workload for different levels of stressors obtained by NASA Task Load Index (TLX). Error bars representing 95% confidence intervals.

Within the NASA-TLX subscales, mental workload was significantly different, $F(1.7,66.2) = 4.28$, $p = 0.022$, $d = 0.67$, with the high stressor ($M = 55.6$, $SD = 21.9$) being marginally significantly higher than the low stressor ($M = 45.8$, $SD = 23.1$; $p = 0.064$, $d = 0.44$). Physical workload was significantly different, $F(2,90) = 7.78$, $p = 0.001$, with the high stressor ($M = 45.5$, $SD = 26.6$) being significantly higher than the low stressor ($M = 34.7$, $SD = 26.2$; $p = 0.01$, $d = 0.41$). Temporal workload was significantly different, $F(2,76) = 14.1$, $p < 0.001$, with the high stressor ($M = 68.2$, $SD = 26.6$) being significantly higher than the low stressor ($M = 47.8$, $SD = 23.8$; $p < 0.001$, $d = 0.81$), high stressor being significantly higher than the medium stressor ($M = 58.4$, $SD = 25.3$; $p = 0.011$, $d = 0.38$), and medium stressor being significantly higher than the low stressor ($p = 0.036$, $d = 0.43$). Performance was significantly different, $F(1.8,68.6) = 5.22$, $p = 0.01$, $d = 0.74$, with the high stressor ($M = 63.4$, $SD = 32.8$) being significantly higher than the medium stressor ($M = 43.7$, $SD = 32.9$; $p = 0.008$, $d = 0.6$). Effort was significantly different, $F(1.63,62.1) = 5.1$, $p = 0.013$, with the high stressor ($M = 67.8$, $SD = 22.8$) being significantly higher than the medium stressor ($M = 59.4$, $SD = 23.1$; $p = 0.001$, $d = 0.37$), and the high stressor being marginally significantly higher than the low stressor ($M = 59$, $SD = 21.4$; $p = 0.053$, $d = 0.40$). Frustration was significantly different, $F(1.77,67.3) = 6.53$, $p = 0.004$, with the high stressor ($M = 45.3$, $SD = 31.2$) being significantly higher than the medium stressor ($M = 32.4$, $SD = 28$; $p = 0.012$, $d = 0.43$) and low stressor ($M = 34.4$, $SD = 27.5$; $p = 0.043$, $d = 0.37$).

5.7. STAI: Anxiety

The main effect of stressor level on post-trial anxiety ($N = 23$) was significant, $F(2,44) = 5.25$, $p = 0.009$, with a large effect size, $d = 0.98$ (Figure 12). Pairwise comparison indicated the anxiety was significantly higher ($p = 0.005$, $d = 0.43$) for participants in high stressor ($M = 40.6$,

SD = 10.7) compared to medium stressor (*M* = 36.2, *SD* = 9.58), but not significantly different (*p* = 0.92) for the medium stressor compared to the low stressor (*M* = 37.5, *SD* = 9.76) or (*p* = 0.24) for the high stressor compared to the low stressor.
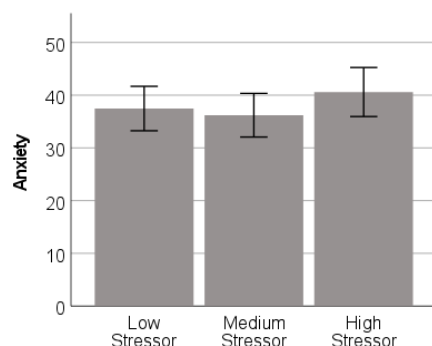
Figure 12. Anxiety obtained by the State Trait Anxiety Inventory (STAI). Error bars representing 95% confidence intervals.

6. Discussion

A spaceflight emergency procedure was conducted in VR-ISS with the aim at evaluating VR stressors. The hypothesis that manipulating different levels of VR environmental stressors during an operational task can induce different levels of stress was partially supported. The results demonstrated that levels of indicators of physiological stress (HR, RMSSD, EDA, SBP), subjective stress (PTSR, Free Stress Comparison), and ratings of distress, workload, and anxiety were significantly different for those training simulations. Some measures did not find significant differences, including pNN50 and DBP. Further, most measures showed differences between the low stressor and high stressors but could not discriminate the medium stressor level.

Several of the physiological stress indicators were found to be different for the stressor levels. The HR was different between the low and high stressor levels and similarly different for the RMSSD. The decrease in RMSSD for the high stress indicates parasympathetic inhibition and reflects an inability to relax. Further, there was an increase in EDA which indicates sympathetic activation, which is the likely cause for the increase in HR and SBP. Altogether, this suggests the high stressor simulation was effective at inducing a robust physiological stress response.

In contrast, the low and medium stressor levels were difficult to distinguish from each other and had varied physiological responses. The RMSSD remained elevated and the HR remained depressed for both levels. However, the medium level did have an increase in EDA, suggesting there was some sympathetic activation concurrent with parasympathetic attenuation. Research on cardiac autonomic balance versus cardiac regulatory capacity [65] suggests that the coactivity is still representative of a stress response, albeit a mild response due to the lack of support from other measures.

The subjective stress was found to be different for the stressor levels based on immediate ratings after each trial (PTSR), ratings after the experiment (Free Stress Comparison), and changes in distress during the trials (SSSQ). The results suggest that design of the VR simulations using environmental stressors was successful at manipulating trainee stress levels. However, the distress results show that low and medium were hard to distinguish between. The cause may be attributed to which stressors were selected for the simulation, the magnitude of each environmental stressors, or the combination of different stressors (e.g., noise, smoke). Previous research has shown that different stressors can elicit varying stress responses and can have a cumulative effect that may be greater than the individual effects of the stressors alone [66,67]. The present experiment used expert opinion to inform how the stressors training simulations but had little empirical support. By changing the stressors, magnitude, or combination, it may be possible to have the VR stressor scenarios result in physiological and distress that are more discriminable in future research. Never-

theless, the large effect of the stressor levels on the subjective stress measures demonstrates distinguishable training levels that can be used to induce stress in VR.

The results show that the workload was higher for the high stressor compared to the other stressor levels. Because the simulations were designed to only change environment stressors and not task load, it was expected that workload would not change between the three stressor levels because the procedure was the same for each. Since workload can be thought of as a stress derived from the stressor of task load, it was predicted that the TLX rating for frustration and performance may be different, but not enough to impact the overall workload. However, it is unexpected to find a difference for the high stressor level from the TLX mental workload, physical workload, temporal, and effort. A possible explanation can be discerned from the increase in the high stressor anxiety levels. The attentional control theory [10] suggests that an increased environmental stressor resulted in more stress, less resources for emotion regulation, and thereby, increases in the perceived demand of the workload. Similarly, previous research that found heightened stress reactivity and threat sensitivity when individuals are in conditions with high amount of stress versus a lower threat sensitivity and reactivity in conditions with no stress [68]. The implications are that that anxiety can have adverse effects on processing efficiency and central executive capacity, thus, a subjective workload may increase concurrently with increasing stress levels, even when the task load stays constant.

While VREs are a promising modality for training operations in stressful environments, questions still exist on how VRE designs effectively translate to the real world. The VR-ISS was created with focus on place illusion, plausibility, and virtual body ownership to enhance immersion and the potential to induce a stress response. However, a meta-analysis on stress biomarkers in VR by [69] found inconsistent responses across studies that induce a robust stress response with both high and low levels of immersion. Further, the individual difference may have a large impact on general biomarker responsivity; thus, new methods may be needed for induction to discriminate finer levels of stress.

This study had several limitations that should be considered with regards to interpretation of results and future work. First, the sample size may have been small and statistically underpowered. Several of the measures (SBP, change in worry) had large effects sizes but marginally significant results, such that there is a possibility the effect exists when tested with more subjects. Running more participants would increase the statistical power and confidence in the results. Second, the study recruited participants from the general population rather than astronauts. The general population is less familiar with the ISS layout and procedures, which could possibly lead to stress or confusion from being trained in a short period of time. Further, astronauts may have stronger associations between threat cues and spaceflight hazards, and simultaneously, possible development of coping skills to manage the threat appraisal. Third, while Latin squares was used to minimize the influence of training order, statistical analysis to verify that assumption for six training-order sequences would be underpowered. Therefore, analysis of the counterbalance was not conducted, and it is unknown to the extent that the effects of the training-order influenced the differences among effects in the stressor levels [60]. Future work will include evaluating the simulations with a participant sample similar to the age range, education level, and demographic of astronauts. Future investigation of other training effects would be beneficial and include memory consolidation/retention, task performance, and physiological habituation over multiple sessions.

7. Conclusions

This research found that stressors present in an emergency procedure could be manipulated in VR and approximate the physiological and subjective responses expected during a stress response. Results from this experiment were mixed. The high and low stressor levels had distinguishable results, but the medium level was harder to distinguish.

The findings complement past research that has investigated levels of stressors in VREs [16], which also found varying levels of stress when individuals are exposed to

gradually increasing stressor. However, this experiment expands on those studies by adding the context of an operational task procedure and relevant environmental stressors.

While the measurement of stress is empirically straightforward, questions still remain about stressor underpinnings, association with immersion, and impact of VRE design on stressor appraisal. VR stressor scenarios may have more discernable physiological and distress by changing the stressors, magnitude, or combination. Future research should investigate how the combinations of stressors may influence the resulting stress response and how similar the stress is to that felt in the real environment.

Astronauts may benefit from training with VREs that simulate procedures on the interior of the ISS. Simulating stressors with a VRE allows for an immersive experience that is not constrained by resources (e.g., money, staff) or physical infrastructure. A training environment that can induce specific stress similar to the operational environment may help individual manage cognitive resources and emotional states with less consequence than in a real situation. By validating stressor levels, astronauts may someday use VR for training emergency fires and other stressful spaceflight procedures.

References

1. Garcia, A.D.; Schlueter, J.; Paddock, E. Training astronauts using hardware-in-the loop simulations and virtual reality. In Proceedings of the AIAA Scitech 2020 Forum, Orlando, FL, USA, 6–10 January 2020; p. 167. [CrossRef]
2. Homan, D.; Gott, C. An integrated EVA/RMS virtual reality simulation, including force feedback for astronaut training. In Proceedings of the Flight Simulation Technologies Conference, San Diego, CA, USA, 29–31 July 1996; p. 3498. [CrossRef]
3. Cater, J.P.; Huffman, S.D. Use of the Remote Access Virtual Environment Network (RAVEN) for Coordinated IVA—EVA Astronaut Training and Evaluation. *Presence Teleoperators Virtual Environ.* **1995**, *4*, 103–109. [CrossRef] [PubMed]
4. Eichler, P.; Seine, R.; Khanina, E.; Schön, A. Astronaut Training for the European ISS Contributions Columbus Module and ATV. *Acta Astronaut.* **2006**, *59*, 1146–1152. [CrossRef]
5. Marciacq, J.B.; Bessone, L. Crew Training Safety: An Integrated Process. In *Safety Design for Space Systems*; Musgrave, G., Larsen, A., Sgobba, T., Eds.; Butterworth-Heinemann: London, UK, 2009; pp. 745–815.
6. Driskell, J.E.; Salas, E.; Johnston, J.H.; Wollert, T.N. Stress Exposure Training: An Event-Based Approach. In *Performance under Stress*; Hancock, P., Szalma, J.L., Eds.; Ashgate: London, UK, 2008; pp. 271–286.
7. Staal, M. *Stress, Cognition, and Human Performance: A Literature Review and Conceptual Framework*; NASA Technical Memorandum 212824; NASA Ames Research Center: Moffett Field, CA, USA, 2004.
8. Hockey, G.R.J. Compensatory Control in the Regulation of Human Performance under Stress and High Workload: A Cognitive-Energetical Framework. *Biol. Psychol.* **1997**, *45*, 73–93. [CrossRef]

9. Gaillard, A.W.K. Stress, workload, and fatigue as three biobehavioral states: A general overview. In *Stress, Workload, and Fatigue*; Hancock, P.A., Desmond, P.A., Eds.; Lawrence Erlbaum: Mahwah, NJ, USA, 2001; pp. 623–640.

10. Eysenck, M.W.; Derakshan, N.; Santos, R.; Calvo, M.G. Anxiety and Cognitive Performance: Attentional Control Theory. *Emotion* **2007**, *7*, 336. [CrossRef] [PubMed]

11. Kilby, C.J.; Sherman, K.A.; Wuthrich, V. Towards Understanding Interindividual Differences in Stressor Appraisals: A Systematic Review. *Personal. Individ. Differ.* **2018**, *135*, 92–100. [CrossRef]

12. Clifford, R.M.S.; Jung, S.; Hoermann, S.; Billinghurst, M.; Lindeman, R.W. Creating a Stressful Decision Making Environment for Aerial Firefighter Training in Virtual Reality. In Proceedings of the 2019 IEEE Conference Virtual Real 3d User Interfaces (VR), Osaka, Japan, 23–27 March 2019; pp. 181–189. [CrossRef]

13. Keren, N.; Bayouth, S.T.; Franke, W.D.; Godby, K.M. Examining the Effect of Level of Stress on Firefighters' Time-to-Decision in Virtual Reality. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, San Diego, CA, USA, 30 September–4 October 2013; Volume 57, pp. 299–303. [CrossRef]

14. Binsch, O.; Bottenheft, C.; Landman, A.; Roijendijk, L.; Vermetten, E.H. Testing the Applicability of a Virtual Reality Simulation Platform for Stress Training of First Responders. *Mil. Psychol.* **2021**, *33*, 182–196. [CrossRef]

15. Frederiksen, J.G.; Sørensen, S.M.D.; Konge, L.; Svendsen, M.B.S.; Nobel-Jørgensen, M.; Bjerrum, F.; Andersen, S.A.W. Cognitive Load and Performance in Immersive Virtual Reality versus Conventional Virtual Reality Simulation Training of Laparoscopic Surgery: A Randomized Trial. *Surg. Endosc.* **2020**, *34*, 1244–1252. [CrossRef]

16. Parsons, T.D.; Rizzo, A.A.; Courtney, C.G.; Dawson, M.E. Psychophysiology to Assess Impact of Varying Levels of Simulation Fidelity in a Threat Environment. In *Advances in Human-Computer Interaction*; Hindawi Limited: London, UK, 2012; pp. 1–9. [CrossRef]

17. Driskell, J.; Johnston, J.; Salas, E. Does Stress Training Generalize to Novel Settings? *Hum. Factors* **2001**, *43*, 99–110. [CrossRef]

18. Uhlig, T.; Roshani, F.-C.; Amodio, C.; Rovera, A.; Zekusic, N.; Helmholz, H.; Fairchild, M. ISS Emergency Scenarios and a Virtual Training Simulator for Flight Controllers. *Acta Astronaut.* **2016**, *128*, 513–520. [CrossRef]

19. Finseth, T.T.; Keren, N.; Dorneich, M.C.; Franke, W.D.; Anderson, C.C.; Shelley, M.C. Evaluating the Effectiveness of Graduated Stress Exposure in Virtual Spaceflight Hazard Training. *J Cogn. Eng. Decis. Mak.* **2018**, *12*, 248–268. [CrossRef]

20. Olbrich, M.; Graf, H.; Keil, J.; Gad, R.; Bamfaste, S.; Nicolini, F. Virtual reality based space operations–a study of ESA's potential for VR based training and simulation. In Proceedings of the International Conference on Virtual, Augmented and Mixed Reality (VAMR) 2018, Held as Part of HCI International 2018, Las Vegas, NV, USA, 15–20 July 2018; pp. 438–451. [CrossRef]

21. Lazarus, R.S.; Folkman, S. *Stress, Appraisal, and Coping*; Springer: New York, NY, USA, 1984.

22. Pallavicini, F.; Argenton, L.; Toniazzi, N. Virtual Reality Applications for Stress Management Training in the Military. *Aerosp. Med. Hum. Perform.* **2016**, *87*, 1021–1030. [CrossRef]

23. Serino, S.; Triberti, S.; Villani, D.; Cipresso, P.; Gaggioli, A.; Riva, G. Toward a Validation of Cyber-Interventions for Stress Disorders Based on Stress Inoculation Training: A Systematic Review. *Virtual Real.* **2014**, *18*, 73–87. [CrossRef]

24. Koolhaas, J.M.; Bartolomucci, A.; Buwalda, B.; de Boer, S.F.; Flügge, G.; Korte, S.M.; Meerlo, P.; Murison, R.; Olivier, B.; Palanza, P.; et al. Stress Revisited: A Critical Evaluation of the Stress Concept. *Neurosci. Biobehav. Rev.* **2011**, *35*, 1291–1301. [CrossRef] [PubMed]

25. Dickerson, S.S.; Kemeny, M.E. Acute Stressors and Cortisol Responses: A Theoretical Integration and Synthesis of Laboratory Research. *Psychol. Bull.* **2004**, *130*, 355–391. [CrossRef]

26. Blascovich, J.; Tomaka, J. The biopsychosocial model of arousal regulation. *Adv. Exp. Soc. Psychol.* **1996**, *28*, 1–51. [CrossRef]

27. Miller, S.M. Predictability and Human Stress: Toward a Clarification of Evidence and Theory. *Adv. Exp. Soc. Psychol.* **1981**, *14*, 203–256. [CrossRef]

28. Slater, M. Place Illusion and Plausibility Can Lead to Realistic Behaviour in Immersive Virtual Environments. *Philos. Trans. R. Soc. B Biol. Sci.* **2009**, *364*, 3549–3557. [CrossRef]

29. Held, R.M.; Durlach, N.I. Telepresence. *Presence Teleoperators Virtual Environ.* **1992**, *1*, 109–112. [CrossRef]

30. Sheridan, T.B. Musings on telepresence and virtual presence. *Presence Teleoperators Virtual Environ.* **1992**, *1*, 120–126. [CrossRef]

31. Barfield, W.; Weghorst, S. The sense of presence within virtual environments: A conceptual framework. In *Human-Computer Interaction: Software and Hardware Interfaces*; Salvendy, G., Smith, M., Eds.; Elsevier: Amsterdam, The Netherlands, 1993; Volume 19, pp. 699–704.

32. Slater, M.; Wilbur, S. A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence Teleoperators Virtual Environ.* **1997**, *6*, 603–616. [CrossRef]

33. Bergstrom, I.; Azevedo, S.; Papiotis, P.; Saldanha, N.; Slater, M. The plausibility of a string quartet performance in virtual reality. *IEEE Trans. Vis. Comput. Graph.* **2017**, *23*, 1332–1339. [CrossRef] [PubMed]

34. Botvinick, M.; Cohen, J. Rubber hands "feel" touch that eyes see. *Nature* **1998**, *391*, 756. [CrossRef] [PubMed]

35. Peperkorn, H.M.; Diemer, J.E.; Alpers, G.W.; Mühlberger, A. Representation of Patients' Hand Modulates Fear Reactions of Patients with Spider Phobia in Virtual Reality. *Front. Psychol.* **2016**, *7*, 268. [CrossRef] [PubMed]

36. Kirschbaum, C.; Pirke, K.; Hellhammer, D. The 'Trier Social Stress Test'—A Tool for Investigating Psychobiological Stress Responses in a Laboratory Setting. *Neuropsychobiology* **1993**, *28*, 76–81. [CrossRef]

37. Plarre, K.; Raij, A.; Hossain, S. Continuous Inference of Psychological Stress from Sensory Measurements Collected in the Natural Environment. In Proceedings of the 10th ACM/IEEE International Conference on Information Processing in Sensor Networks, Chicago, IL, USA, 12–14 April 2011; pp. 97–108.

38. Johnston, J.; Cannon-Bowers, J.A. Training for stress exposure. In *Stress and Human Performance*; Driskell, J., Salas, E., Eds.; Lawrence Erlbaum: Mahwah, NJ, USA, 1996; pp. 223–256.

39. Dorneich, M.C.; Mathan, S.; Ververs, P.M.; Whitlow, S.D. Cognitive State Estimation in Mobile Environments. In *Augmented Cognition: A Practitioner's Guide*; Schmorrow, D., Stanney, K., Reeves, L., Eds.; HFES Press: Santa Monica, CA, USA, 2008; pp. 75–111.

40. Healey, J.A.; Picard, R.W. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Trans. Intell. Transp. Syst.* **2005**, *6*, 156–166. [CrossRef]

41. Chabal, S.; Bohnenkamper, A.; Reinhart, P.; Quatroche, A. *Stressors Present in a Disabled Submarine Scenario: Part 1. Identification of Environmental, MENTAL, and Physical Stressors*; Naval Submarine Medical Research Laboratory: Groton, CT, USA, 2019.

42. United States National Aeronautics and Space Administration (NASA). *International Space Station, Emergency Procedures 1a: Depress, Fire, Equipment Retrieval (No. JSC-48566)*; NASA Johnson Space Center: Houston, TX, USA, 2013.

43. Bartone, P.T.; Roland, R.R.; Bartone, J.V.; Krueger, G.P.; Sciarretta, A.A.; Johnsen, B.H. Human Adaptability for Deep Space Missions: An Exploratory Study. *J. Hum. Perform. Extreme Environ.* **2019**, *15*, 5. [CrossRef]

44. IGOAL, NASA/JSC, "ISS (Internal)," NASA 3D Resources. 27 March 2017. Available online: https://nasa3d.arc.nasa.gov/detail/iss-internal (accessed on 27 March 2017).

45. Finseth, T.; Dorneich, M.C.; Keren, N.; Franke, W.D.; Vardeman, S. Designing Training Scenarios for Stressful Spaceflight Emergency Procedures. In Proceedings of the 2020 AIAA/IEEE 39th Digital Avionics Systems Conference (DASC), San Antonio, TX, USA, 11–15 October 2020; pp. 1–10. [CrossRef]

46. McCraty, R.; Shaffer, F. Heart Rate Variability: New Perspectives on Physiological Mechanisms, Assessment of Self-Regulatory Capacity, and Health Risk. *Glob. Adv. Health Med.* **2015**, *4*, 46–61. [CrossRef]

47. Sharma, N.; Gedeon, T. Modeling Observer Stress for Typical Real Environments. *Expert Syst. Appl.* **2014**, *41*, 2231–2238. [CrossRef]

48. Mestanik, M.; Visnovcova, Z.; Tonhajzerova, I. The Assessment of the Autonomic Response to Acute Stress Using Electrodermal Activity. *Acta Med. Martiniana* **2014**, *14*, 5–9. [CrossRef]

49. Braithwaite, J.; Watson, D.; Jones, R. A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments. *Psychophysiology* **2013**, *49*, 1017–1034.

50. Penley, J.A.; Tomaka, J. Associations among the Big Five, Emotional Responses, and Coping with Acute Stress. *Personal. Individ. Differ.* **2002**, *32*, 1215–1228. [CrossRef]

51. Singh, R.R.; Conjeti, S.; Banerjee, R. A comparative evaluation of neural network classifiers for stress level analysis of automotive drivers using physiological signals. *Biomed. Signal Process.* **2013**, *8*, 740–754. [CrossRef]

52. Carpenter, R. A Review of Instruments on Cognitive Appraisal of Stress. *Arch. Psychiatr. Nurs.* **2016**, *30*, 271–279. [CrossRef] [PubMed]

53. Helton, W. Validation of a Short Stress State Questionnaire. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, New Orleans, LA, USA, 20–24 September 2004; Volume 48. [CrossRef]

54. Matthews, G.; Joyner, L.; Gilliland, K.; Campbell, S.; Falconer, S.; Huggins, J. Validation of a Comprehensive Stress State Questionnaire: Towards a State "Big Three". In *Personality Psychology in Europe*; Mervielde, I., Deary, I.J., De Fruyt, F., Ostendorf, F., Eds.; Tilburg University Press: Tilburg, The Netherlands, 1999; Volume 7, pp. 335–350.

55. Helton, W.; Näswall, K. Short stress state questionnaire. *Eur. J. Psychol. Assess.* **2015**, *31*, 20–30. [CrossRef]

56. Hart, S.G.; Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Adv. Psychol.* **1988**, *52*, 139–183. [CrossRef]

57. Nygren, T.E. Psychometric Properties of Subjective Workload Measurement Techniques: Implications for Their Use in the Assessment of Perceived Mental Workload. *Hum. Factors* **1991**, *33*, 17–33. [CrossRef]

58. Spielberger, C.D.; Gorsuch, R.L.; Lushene, R.; Vagg, P.R.; Jacobs, G.A. *Manual for the State-Trait Anxiety Inventory*; Consulting Psychologists Press: Palo Alto, CA, USA, 1983.

59. Tartarisco, G.; Carbonaro, N.; Tonacci, A.; Bernava, G.; Arnao, A.; Crifaci, G.; Cipresso, P.; Riva, G.; Gaggioli, A.; Rossi, D.D.; et al. Neuro-Fuzzy Physiological Computing to Assess Stress Levels in Virtual Reality Therapy. *Interact. Comput.* **2015**, *27*, 521–533. [CrossRef]

60. Reese, H.W. Counterbalancing and Other Uses of Repeated-Measures Latin-Square Designs: Analyses and Interpretations. *J Exp. Child Psychol.* **1997**, *64*, 137–158. [CrossRef]

61. Kim, H.Y. Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis. *Restor. Dent. Endod.* **2013**, *38*, 52. [CrossRef]

62. Gelman, A. Commentary: *P* values and statistical practice. *Epidemiology* **2013**, *24*, 69–72. [CrossRef]

63. Cohen, J. *Statistical Power Analysis for the Behavioural Sciences*; Lawrence Earlbaum Associates: Hillside, NJ, USA, 1988.

64. Matthews, G.; Campbell, S.E.; Falconer, S.; Joyner, L.A.; Huggins, J.; Gilliland, K.; Grier, R.; Warm, J.S. Fundamental Dimensions of Subjective State in Performance Settings: Task Engagement, Distress, and Worry. *Emotion* **2002**, *2*, 315–340. [CrossRef] [PubMed]

65. Berntson, G.G.; Norman, G.J.; Hawkley, L.C.; Cacioppo, J.T. Cardiac Autonomic Balance versus Cardiac Regulatory Capacity. *Psychophysiology* **2008**, *45*, 643–652. [CrossRef] [PubMed]

66. Abdelall, E.S.; Eagle, Z.; Finseth, T.; Mumani, A.A.; Wang, Z.; Dorneich, M.C.; Stone, R.T. The Interaction Between Physical and Psychosocial Stressors. *Front. Behav. Neurosci.* **2020**, *14*, 63. [CrossRef] [PubMed]

67. Pedrotti, M.; Mirzaei, M.A.; Tedesco, A.; Chardonnet, J.-R.; Mérienne, F.; Benedetto, S.; Baccino, T. Automatic Stress Classification with Pupil Diameter Analysis. *Int. J. Hum.-Comput. Int.* **2014**, *30*, 220–236. [CrossRef]

68. Akinola, M.; Mendes, W.B. Stress-Induced Cortisol Facilitates Threat-Related Decision Making among Police Officers. *Behav. Neurosci.* **2012**, *126*, 167–174. [CrossRef]

69. van Dammen, L.; Finseth, T.; McCurdy, B.H.; Barnett, N.P.; Conrady, R.; Leach, A.G.; Deick, A.F.; Stennis, A.V.; Gardner, R.; Smith, B.; et al. Evoking stress reactivity in virtual reality: A systematic review and meta-analysis. *Neurosci. Biobehav. Rev.* 2022; *submitted*.