Original Paper

An Integrated Platform Combining Immersive Virtual Reality and Physiological Sensors for Systematic and Individualized Assessment of Stress Response (bWell): Design and Implementation Study

Budhachandra Khundrakpam[1], PhD; Melanie Segado[1], PhD; Jesse Pazdera[1], PhD; Vincent Gagnon Shaigetz[1], MSc; Joshua A Granek[2], PhD; Nusrat Choudhury[1], MEng

[1]National Research Council Canada, Boucherville, QC, Canada
[2]Defence Research and Development Canada, Toronto, ON, Canada

Corresponding Author:
Nusrat Choudhury, MEng
National Research Council Canada
75 de Mortagne
Boucherville, QC, J4B 6Y4
Canada
Phone: 1 450 978 6198
Email: nusrat.choudhury@nrc-cnrc.gc.ca

Abstract

Background: Stress is a pervasive issue in modern society, manifesting in various forms such as emotional, physical, and work-related stress, each with distinct impacts on individuals and society. Traditional stress studies often rely on psychological, performance, or social tests; however, recently, immersive virtual reality (VR), which provides a sense of presence and natural interaction, offers the opportunity to simulate real-world tasks and stressors in controlled environments. Despite its potential, the use of VR to investigate the multifaceted manifestations of stress has not been thoroughly explored.

Objective: This study aimed to explore the feasibility of using a VR-based platform, *bWell*, to elicit multifaceted stress responses and measure the resulting behavioral and physiological changes. Specifically, we aimed to design various VR stress exercises based on neurocardiac models to systematically test cardiac functioning within specific contexts of self-regulation (executive functioning, physical efforts, and emotional regulation).

Methods: The development process adhered to guidelines for VR clinical trials and complex health interventions, encompassing 3 phases: preparation, development, and verification. The preparation phase involved a comprehensive literature review to establish links between stress, the heart, and the brain, leading to the formulation of a conceptual model based on the Neurovisceral Integration Model (NVIM) and Vagal Tank Theory (VTT). The development phase involved designing VR exercises targeting specific stressors and integrating physiological sensors such as photoplethysmography (PPG) and electromyography (EMG) to capture heart rate variability (HRV) and facial expressions. The verification phase, conducted with a small number of trials, aimed to design a study and implement a workflow for testing the feasibility, acceptability, and tolerability of the VR exercises. In addition, the potential for capturing physiological measures along with subjective ratings of stress for specific dimensions was assessed.

Results: Verification trials demonstrated that the VR exercises were well tolerated, with negligible cybersickness and high user engagement. The different VR exercises successfully elicited the intended stress demands, along with the physiological responses.

Conclusions: The study presents a novel VR-based experimental setup that allows a systematic and individualized assessment of stress responses, paving the way for future research to identify features that confer stress resilience and help individuals manage stress effectively. While our conceptual model highlights the role of HRV in providing valuable insights into stress responses, future research will involve multivariate and machine learning analyses to predict individual stress responses based on comprehensive sensor data, including EMG and the VR-based behavioral data, ultimately guiding personalized stress management interventions.

Introduction

Background

A rising problem in our society is *stress*, with serious burdens imposed on individuals and society by different types of stressors, such as emotional stress, physical stress, and work stress [1,2]. The multifaceted nature of stress makes it even more problematic, with different stressors yielding different responses and consequences. For example, increased workload and time pressure has led to increased mental stress [3]. Similarly, in recent years, there has been an increase in emotional, behavioral, and psychological stress due to the COVID-19 pandemic [4,5].

Cumulative evidence from studies of stress neurobiology indicates individual variability in how people respond to stress. While for some individuals, stress can lead to negative outcomes (eg, deterioration of physical and mental health), other individuals may not be affected to that extent [6,7]. Identifying the risk factors that make a person vulnerable to negative consequences of stress would allow timely detection and intervention. By contrast, identifying and delineating the individual features that potentially confer stress resilience (defined as the ability of most people to maintain normal functioning even under extreme stress) is important in helping individuals reach their full potential. Factors that confer resilience include genetic predisposition, age, sex, prenatal environment, sleep, and exercise [8]. In terms of biological mechanisms, stress resilience has been attributed to a coordinated activity of the autonomic nervous system (ANS); the hypothalamic-pituitary-adrenal axis; the immune system; and the brain regions that control their activity, including the hypothalamus, amygdala, hippocampus, medial prefrontal cortex, and ventral tegmental area [9-11]. In particular, stress-induced activation of ANS has been linked to changes in heart rate and vasoconstriction [9]. In addition, stress responses via activation of sympathetic-adrenomedullary and hypothalamic-pituitary-adrenal axes have been observed in response to psychosocial and physical demands during intense exercise [12].

Given the complex coordinated involvement of several bodily systems (both central and peripheral), studies on stress resilience require measurement of attributes that can reflect these interactions. One such attribute is heart rate variability (HRV), which captures heart-brain interactions and ANS dynamics [13]. HRV is defined as the variation in the time intervals between consecutive heartbeats and is an emergent property of complex interactions among several interdependent physiological systems. Several studies have shown the usefulness of HRV in the analysis of stress that the body experiences during physical training and to increase insight into physiological recovery after training [14,15]. Measurement of HRV (using electrocardiogram [ECG] or photoplethysmography [PPG]) is an easy noninvasive technique, and therefore, it has been used to investigate stress-related changes in cardiovascular health in different conditions, including exercise [16] and sleep [17]. More importantly, studies on individuals with stress-related disorders (such as anxiety and depression) revealed altered HRV metrics, indicating impairment of neurocardiac function mediated by changes to the ANS [18]. In addition, integrating HRV with other modalities such as electroencephalography have shown the impact of stress on cardiac and brain health [19,20].

Traditionally, the impact of stress on biological measures such as HRV and brain activity is investigated after inducing a specific stressor (mainly in live scenarios) using psychological, performance, or social tests [21,22]. However, in recent years, the advancement of virtual reality (VR) and head-mounted displays has allowed the investigation of different stressors that are otherwise difficult to reproduce in real life. Using head-mounted displays, immersive VR offers a sense of presence and a more natural interaction, allowing multisensorial (visual, audio, and touch) simulation of real-world tasks with tight experimental control, enabling the study of specific stressors. In a study, VR was used to elicit stress in high-altitude simulations, and stress-related responses in the heart and brain were observed using ECG and electroencephalography, respectively [23]. In a VR study using ECG signals, a multilevel stress assessment (3 levels of stress) was performed using spectrograms and 1D ECG signals that performed better than traditional machine learning (ML) models [24]. In another example, playing 3D games (a combination of the first-person shooter and horror genres) in immersive VR was shown as an effective stressor for practicing stress management skills [25]. In alignment with these findings, a meta-analysis of 52 studies revealed that stress tasks in VR elicit a variety of physiological stress reactivity, highlighting the scope of VR in the study of stress [26].

Objectives

Despite the advantages offered by VR (multisensorial—visual, audio, and touch simulation of real-world tasks), previous studies have mostly used VR to induce particular stressor at a time, for example, cognitive stress [27] and emotional stress [28]. In this work, we explored the feasibility of implementing a VR-based experimental setup to (1) elicit multifaceted stress response and (2) measure the resulting behavioral and physiological response. Using our validated VR platform *bWell* [29], we aimed to leverage specific exercises among its battery for eliciting different stressor responses, specifically for physical, cognitive, and emotional stress. Particular emphasis was given on designing VR stress exercises to systematically test cardiac functioning within specific contexts of self-regulation (executive functioning and emotional regulation) based on neurocardiac models [30,31]. In addition, by integrating VR with PPG and electromyography (EMG), we captured various physiological measures, including HRV and facial expressions (collected via EMG), along with subjective ratings of stress (namely, the Perceived Stress Scale scores for

overall stress and NASA-Task Load Index [NASA-TLX] for specific stress dimensions) and mood.

The overarching goal is that the study design or paradigm tailor to different stress domains and measurable behavioral and physiological outputs. The long-term goal is to perform multivariate and ML-based analysis for predicting the subjective perceived stress of an individual. In addition, it should be adaptable to the specific needs of each individual, allowing a systematic and individualized assessment of stress response.

Methods

Development Process

Because the work involved immersive VR in a health context, the "Recommendations for Methodology of Virtual Reality Clinical Trials in Health Care by an International Working Group" were followed during the development process [32]. Following the recommendations, we focused on the first two phases: (1) VR1 phase with an emphasis on content development of different VR stress exercises and (2) VR2 phase with an emphasis on verification trials testing feasibility, acceptability, and tolerability. In addition, the study development process also followed guidelines from the Medical Research Council, namely *"Guidance on how to develop complex interventions to improve health and healthcare,"* [33] which recommends combining the opinions or input elicited during content development with existing or available theories (eg, the rationale for intervention, expected changes, and how changes may be achieved). On the basis of the guidelines, our development study included a precursor step to identify the prevailing theories and frameworks in the domain to formulate hypotheses for the study.

Consequently, our development process comprised three phases: (1) a *preparation phase,* during which an extensive literature review was performed to establish the links between heart, brain, and stress, leading to the formulation of a conceptual model and formation of hypotheses (precursor phase); (2) a *development phase,* which involved iterative design of VR exercises for targeted stressors and hardware integration (VR1 phase); and (3) a *verification phase* (VR2 phase), focusing on implementing a workflow able to test the outcome measures associated with VR2 trials—tolerability, acceptability, and feasibility (Figure 1).

Figure 1. Schema of the development process. The development process comprised three phases: preparation phase, virtual reality (VR)-1 development phase and VR-2 verification phase.
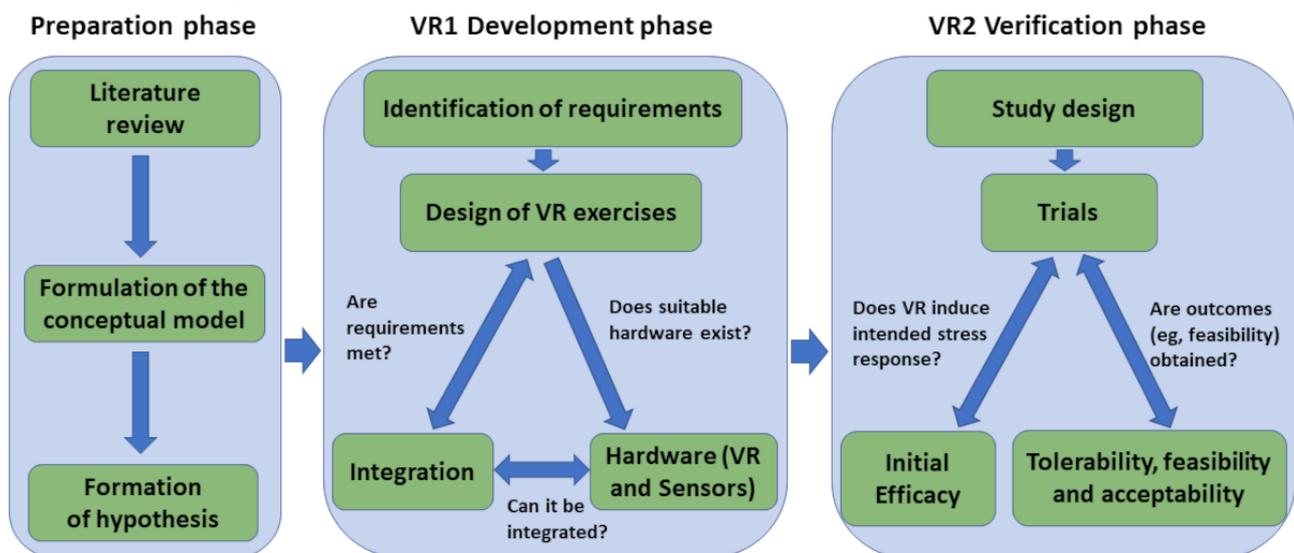
Preparation Phase

Literature Review and Survey

A literature survey was performed to establish links between stress, the heart, the brain, and cognition. Given the involvement of several bodily systems in stress, we observed diverse findings: on the one hand, associations with mental load [34] and emotional stress [35], on the other hand, associations with cognitive performance [36,37]. With the goal of linking these diverse findings, we next focused on publications that have reported integrated models of the heart, the brain, and cognition (especially self-regulation, a cognitive process implicated in stress). The first such model is the *Neurovisceral Integration Model(NVIM)* [38,39]. According to the NVIM, the heart and the brain are connected by direct and indirect pathways that influence the various processes (physiological, behavioral, emotional, and cognitive) related to self-regulation and adaptability. More importantly, the NVIM postulates vagal tone as the common index that is linked to physiological, affective, and cognitive regulation. The second model is based on the *Vagal Tank Theory (VTT)* [31]. The VTT suggests that a higher vagal tone is associated with better self-regulation and adaptive functioning.

It is worth noting that the theories of interest focus on vagal tone, which refers to the activity of the vagus nerve. The vagus nerve is the largest autonomic nerve, innervating nearly every organ in the body. More importantly, the vagus nerve regulates metabolic homeostasis by controlling the heart rate [40] as well as the interval between 2 heartbeats (HRV) [41]. While direct measurement of vagal activity (ie, physiological studies of the vagus nerve) has been limited due to difficulties in recording chronic signals in the vagus [42,43], HRV provides an indirect

measure of vagal tone (a clinical metric to specify overall levels of vagal activity) because it is largely regulated by cardiac vagal control [44,45]. HRV has been shown to be a sensitive metric for detecting abnormalities in vagal tone, as observed in clinical conditions such as diabetes, heart failure, and hypertension [46-48]. More broadly, HRV, as an indirect measure of vagal tone, has shown to be useful in many aspects of psychophysiology, such as self-regulation mechanisms linked to cognitive, affective, social, and health phenomena [13,30,49]. In summary, HRV is commonly used in research as a practical index for vagal tone, and therefore, we will be using HRV in our study to link vagal tone with stress.

Formulation of the Conceptual Model

To formulate our model, we first examined the hypotheses of the NVIM and VTT. Specifically, the NVIM suggests that greater parasympathetic nervous system activity (ie, vagal tone) increases regulatory control in the brain and body [38,39]. This control supports certain biological processes such as digestion, while facilitating brain functioning related to cognitive control and emotional regulation. In contrast, withdrawal of vagal tone supports physical action, such as a fight-or-flight response. The parasympathetic nervous system is thought to monitor and balance feedback from numerous subsystems to determine the optimal level of vagal tone to deploy, given one's current context, active goal, and unmet needs [50]. HRV reflects vagal tone and, as such, should correlate with the integrity of cognitive and emotional control processes [30]. By contrast, VTT suggests that there are *3* states of vagal control, the "3 Rs": *resting, reactivity,* and *recovery* [31]. Optimal vagal tone is high at rest, so a healthy system should settle to a high baseline level of parasympathetic activity. In response to mental demands, vagal control systems should react by maintaining a high vagal tone to promote executive control. In response to more active demands, these systems should react by withdrawing vagal tone to promote physical performance. In practice, vagal tone often drops below resting levels even during cognitive tasks, possibly due to the strain on and gradual exhaustion of vagal control systems. Once demands are met, vagal control should transition into a recovery state, where vagal tone returns to resting levels.

The NASA-TLX provides a formal framework for categorizing different aspects of workload and task demands [51]. Therefore, we identified this index as a useful tool for understanding the demands that vagal control systems must continuously monitor and react to, as per the *VTT* and the *NVIM*. Specifically, the NASA-TLX defines 3 types of task demands—*mental*, *physical*, and *temporal*—and 3 additional factors—*performance*, *effort*, and *frustration*—that contribute to perceived workload. On the basis of the *NVIM* model of vagal control, *mental*, *physical*, and *temporal* task demands should be associated with different optimal levels of vagal tone and, therefore, different states of reactivity. *Mental* demands are best solved by promoting executive control through high levels of parasympathetic nervous system activity; in contrast, active *physical* and *temporal* demands are best addressed by a withdrawal of vagal tone.

Formation of Hypothesis

On the basis of previous publications, we formed specific hypotheses for self-regulation for each level of cardiac vagal control (Table 1). At *resting (baseline)*, we hypothesized that the higher the vagal tone, the better the self-regulation [49,52-54]. For *reactivity*, different hypotheses were formed for emotional, physical, and cognitive demands. It may be noted that for consistency with terminology in neuroscience, we have used cognitive demand as representing mental demand and emotional demand as representing frustration in the NASA-TLX. For emotional demand, (1) a complete vagal withdrawal would elicit a fight-or-flight response [52,55], (2) a small increase in vagal tone would indicate successful emotion regulation [53,54], and (3) a small decrease would indicate that the task did not elicit a sufficiently large stress response [49,56]. For cognitive demand, (1) we hypothesized an initial increase in vagal tone under low cognitive load (self-regulatory effort) and a decrease in vagal tone under high cognitive load [31,57], and (2) in the context of vagal withdrawal (ie, high cognitive load), we hypothesized that a small decrease in vagal tone would predict better performance [31,58]. For physical demand, a small vagal withdrawal would indicate quick adaptation to meet the demands of the situation [59,60]. Finally, for *recovery*, we hypothesized that the faster the vagal tone returns to the baseline level, the better the self-regulation [61].

Table 1. Description of the hypotheses corresponding to our conceptual model based on the Neurovisceral Integration Model (NVIM) and Vagal Tank Theory (VTT). A distinct hypothesis was made for each phase—resting (baseline), reactivity, and recovery.

Cardiac vagal control level	Hypothesis of self-regulation, as indicated by cardiac vagal control
1. Resting (baseline)	• A higher vagal tone indicates better self-regulation
2. Reactivity	
2a. Emotional demand	• Complete vagal withdrawal indicates that a fight-or-flight response was elicited • A small increase indicates successful emotional regulation • A small decrease indicates that the task did not elicit a sufficiently large stress response
2b. Cognitive demand	• An initial increase in vagal tone under low cognitive load (self-regulatory effort) and a decrease under high cognitive load • A smaller decrease (under high cognitive load) is associated with better performance
2c. Physical demand	• Small vagal withdrawal indicates quick adaptation to meet demands of the situation
2d. Physical and cognitive demand	• Combination of hypotheses 2c (vagal withdrawal) and 2b (initial increase under low cognitive load with a more pronounced effect)
3. Recovery (from demand)	• The faster the vagal tone returns to the baseline level, the better the coping mechanisms

VR1: Development Phase

Identification of Requirements

The fundamental objective of this work is to harness the affordances of immersive VR to provide new opportunities for practicing stress management skills, particularly the synthetic recreation of scenes that are otherwise too difficult or resource intensive to replicate in real life. The latter is of particular interest in the application toward training stress resilience, a complex and dynamic process. Combined with physiological monitoring, which can help to capture additional information on the user state, the use of VR-based technologies can enable research in support of better understanding the impacts of well-being and individual characteristics, such as mental and psychological resilience.

To study and validate the different parts of our conceptual model, we identified four main requirements:

1. *Requirement 1:* recreation of each level of cardiac vagal control (3 "Rs": resting, reactivity, and recovery)
2. *Requirement 2:* ability to elicit different types of mental demands
3. *Requirement 3:* ability to elicit physical demand
4. *Requirement 4:* integration of physiological sensing with immersive VR, essential for the integration of HRV measures as an index of vagal tone

Design of VR Exercises

We have previously developed *bWell*, an interactive and immersive VR platform designed as a broadly applicable toolkit delivering multisensory tasks, targeting general aspects of cognition and everyday functioning [29,62-64]. For this study, we made use of *3* bWell exercises, namely *Tent*, *Mole,* and *Stroll,* which were adapted to systematically test adaptive aspects

of HRV within specific contexts of self-regulation (executive functioning, emotional regulation, and physical demand) and for different levels of cardiac vagal control, as identified in our conceptual model and requirements 1 and 2. The bWell platform and exercise battery are described in detail in our previous publication [29]. The following is a brief summary of the three exercises:

1. *Tent*—an exercise designed for self-guided stress management in which the user is immersed in scenic views and asked to look around while focusing on their breathing
2. *Mole*—an exercise targeting response inhibition and cognitive control in which the user has a hammer in each hand and has to hit cylinders that pop up in front of them (only when the cylinder and hammer are of the same color)
3. *Stroll*—an exercise targeting sustained attention with a continuous performance test (CPT) in which the user has to press a button for each new shape that appears, except when it is a green diamond. The user is represented by a self-avatar and is passively displaced for a stroll in a natural scene.
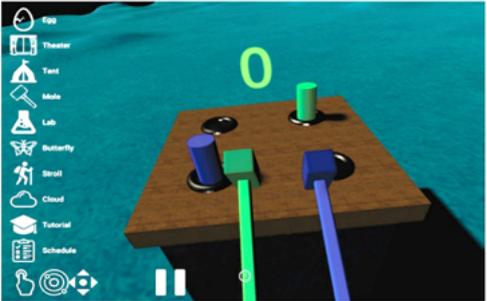
To test our specific hypotheses for different stress demands or stressors (Table 1), we adapted the aforementioned bWell exercises (except for the *Mole* exercise for cognitive demand) as follows:
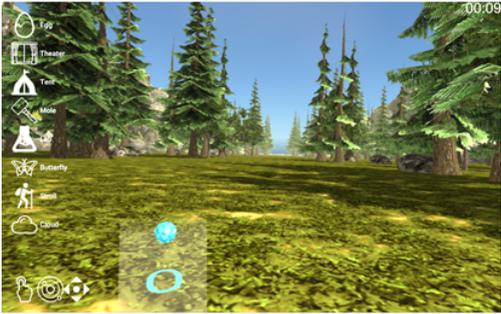
1. *Tent*
   - Nature scene to collect data on user
     - *Resting:* As per recommended guidelines for experimental settings, baseline HRV should be recorded in a "nontask situation that best controls for the presence of task comparison" [65]. Consequently, the administration of the exercise was modified such that the user was asked to simply immerse themselves in the scene, limit

physical movement, and breathe spontaneously for 5 minutes.

- *Recovery:* this scene was also selected to be administered after a given task to allow a period of recovery after stress demand. As a measure of recovery, a Likert scale implemented in VR was added to collect a report on mood right after the exposure to a stressor and after the recovery period in a relaxing nature environment.

- *City*—in addition to the nature scenes, to elicit emotional demand, we adapted the *Tent* exercise to include an environment such that the user is immersed in a busy city with noise pollution—roads with ongoing traffic, including audio honking cars and sirens.

2. *Stroll*—to exert physical demand, we adapted the Stroll exercise such that the user must physically displace themselves to advance in the scene. The user was asked to jog, keeping their heart rate at a moderate level of exertion (the predefined threshold was displayed as a gauge based on individual heart rate). The exercise could include a CPT, as described earlier, to exert both physical and cognitive demand (*Stroll+CPT*) or could be administered without the test as a purely physical task (*Stroll*).

The formulated hypotheses shown in Table 1 can now be systematically matched with our designed VR exercises (Table 2).

Table 2. Description of the hypotheses corresponding to our conceptual model with the designed virtual reality (VR) exercises. A distinct hypothesis was made for each phase—resting (baseline), reactivity, and recovery—and these hypotheses were tested with different VR exercises.

Cardiac vagal control level	Requirement	bWell exercise
1. Resting (baseline): Tent—nature scene, no specific task given	User to immerse themselves, limit physical movement, and breathe spontaneously as a baseline state	
2. Reactivity		
2a. Emotional demand: city—(passive) city scene	User to immerse in a busy city with excessive noise pollution to elicit emotional demand	
2b. Cognitive demand: mole—(active) targeting response inhibition	User to engage in a dynamic and bimanual go or no-go task to elicit cognitive demand	
2c. Physical demand: stroll—(active) physical running task	User to jog to create a physical demand, with heart rate biofeedback guiding them to maintain a consistent moderate level of activity	

Cardiac vagal control level	Requirement	bWell exercise
2d. Physical and cognitive demand: stroll+CPT[a]–dual task with running task as 2c along with a cognitive task (response inhibition)	User to jog and engage in the task to elicit both physical and cognitive demands	
3. Recovery from demand: Tent—nature scene	User to immerse themselves and limit physical movement to recover	

[a]CPT: continuous performance test.

Hardware Identification (VR+Sensors)

Based on the identified requirements, we aimed to identify and integrate hardware for use with the bWell exercises. We focused on evaluating the suitability of the use of immersive VR technologies. The choice of hardware considered specialized equipment aiming to maximize the feeling of "*presence*" in virtual environments as well as the potential to democratize the use of VR for training more broadly. Consequently, we evaluated both a high-end system—*HTC Vive Pro Eye*—or use within dedicated facilities as well as a more accessible option—*Pico Neo 3 Eye*. The latter targets the opportunity for individuals to practice at home or in settings outside the laboratory (eg, on the field, classroom, and clinical).

We aimed to reproduce the required factors in the simulated environment with the integration of specialized hardware. A VR treadmill (KatVR) was identified to recreate the physical demands. We also aimed to capture stress via the inclusion of physiological sensing, with a focus on integrating heart-based measures. The *EmteqPRO* facemask was identified for use in this project because it is a commercially available "plug-and-play" system for use with immersive VR, provides researchers with access to raw data, and is equipped with automatic classification algorithms for processing raw heart rate data into HRV as well as classification of emotion [66]. The system can accommodate multimodal measures of EMG and HRV via the use of a facemask that can be directly integrated onto the VR headset. The Emteq device does not add any additional steps to the workflow for carrying out trials with immersive VR; instead, the user simply must put on the headset as they normally would.

Hardware Integration

A first evaluation was conducted to determine the suitability of the identified hardware.

Immersive VR Headset

The *HTC Vive Pro Eye* and the *Pico Neo 3 Eye* headsets were identified because they both offer eye tracking. In addition, the Emteq solution is available for both. The HTC Vive Pro Eye is a tethered device, meaning it requires a high-performance, VR-ready PC to operate. A tethered VR headset operates only when connected to a PC through a cable connection and offers superior graphics and fidelity in comparison to stand-alone headsets. The *Pico Neo 3 Eye* is a stand-alone headset, which is wireless, and the VR is run entirely within the headset. The benefits of a stand-alone headset include increased comfort and mobility, a simple setup, and a lower cost compared to tethered devices. However additional efforts to optimize the VR experience in terms of graphical rendering and design are required to minimize lags related to hardware performance.

It was determined that both headsets are suitable for use in the proposed study with certain adaptations. A wireless adapter was integrated into the HTC Vive Pro Eye, enabling untethered use with the treadmill. For the Pico Neo 3 Eye, an external dongle was required to enable screen casting to an external monitor, allowing the test administrator to view the participant's perspective within the VR during trials.

Emteq Integration With bWell in Unity

The Emteq face mask was integrated within *bWell* through the use of the "EmteqLabs Unity" plugin. Using the plugin, the EMG amplitude as well as the heart rate (beats per minute) of each participant was extracted. Using these extracted values,

several widgets (Figure 2) were implemented in the 2 existing layers of the *bWell* user interface (UI): test administrator UI and participant UI. The first widget was part of the EmteqLabs plugin package and was used to display EMG amplitude (Figure 2A). The same widget also provided the administrator with an indication of sensor contact. Good contacts were represented by green dots on the widget, while bad contacts were represented by gray dots. The second widget was composed of 2 simple text boxes that were used to inform the test facilitator of the current connection status of the Emteq (Figure 2B). The third widget was a speedometer that was used to inform the participant whether they were in a heart rate zone of moderate effort (Figure 2C). The widget was implemented as a gauge to prompt participants to raise or lower their effort level, with the zone in the widget turning green when participants are at the correct range. Specifically, (zone turns green at correct range). A heart icon that grew and shrank based on a speed that matched the current beats per minute was also included in the speedometer.

Figure 2. Integration of physiological sensor with virtual reality (VR). Implementation of widgets in the user interface for integration of EmteqPRO (physiological sensor for heart rate variability) into VR platform (bWell): (A) widget for displaying electromyography (EMG) amplitude in VR, (B) widget for connection status of EmteqPRO, and (C) widget for displaying heart rate.

Treadmill Integration

The usability of donning a wireless immersive VR headset while running on the treadmill was tested. Running speed (vigorous vs moderate) was evaluated, and it was found that both vigorous and moderate running was acceptable in terms of usability for both headsets. However, it was seen that running on a treadmill was not intuitive and that a warm-up time was required. A test run was added to the protocol, such that the user could have some time to familiarize themselves with running on the treadmill before putting on the VR headset.

It was also seen that vigorous running may prove to be difficult for participants running on the treadmill for the first time. In addition, running on the treadmill while performing a cognitive exercise (*Mole* exercise) occasionally resulted in inadvertent errors from the user. As a result, the protocol was adjusted to have users run at a moderate pace (light jog) during the study. For the dual task (Stroll+CPT), a cognitive exercise with simpler user interaction was use. The exercise was changed from the initial choice to use *Mole* (requiring users to swing with their arms) to the *Stroll* (requiring users to pull a trigger on the controller).

Finally, the physical demand exercise (*Stroll*) was originally designed for passive displacement. In this task, it is the self-avatar taking a nature stroll while the actual user is not moving. To promote a higher sense of *presence* for the user, we added advancement in the VR scene based on the user's own steps while advancing on the treadmill. The running pace was added as an output measure in the VR logs generated automatically after a session with bWell.

The conclusion from the hardware integration tests was that both the *HTC Vive Pro Eye* and *Pico Neo 3 Eye* were suitable.

However, because the primary use at this stage was for study purposes, it was decided that it would be better to use the hardware that is designed for external administration of VR versus self-administration with a headset. In the workflow for administration, the bWell platform provides a UI that allows the test facilitator to dynamically change the exercise settings while a user is immersed in VR [29]. Consequently, it would be better suited for the iterative approach in finding the optimal exercise settings for the study. Thus, the decision was to use the *HTC Vive Pro Eye* for the study and the ensuing verification trials.

VR2: Verification Phase

In this phase, we focused on conducting early trials for evaluation and verification of feasibility, acceptability, tolerability, and initial efficacy (VR2 phase).

Study Design

We designed a study to systematically test our hypotheses (Figure 3). The study protocol consisted of a 5-minute initial baseline consistent with what is used in the HRV literature (Table S1 in Multimedia Appendix 1). During baseline, participants were in a standing position with their eyes open to optimally match the body position used during task conditions. Participants experienced a neutral nature scene. The baseline HRV was collected to establish a reference point. Participants then experienced alternating blocks of stressor tasks interspersed with recovery (Tent nature scenes). Stressor task presentation order was pseudorandomized across participants. Each participant performed all 4 exercises (stress conditions). Each stressor condition was followed by a 2-minute recovery period in the *Tent* (Table S1 in Multimedia Appendix 1).

Figure 3. Description of the study design. A within-participants repeated measures study design was implemented. It comprised an initial baseline (5 minutes), followed by 4 blocks of stressor (2 minutes each) and recovery (2 minutes each) periods. The stressors were 4 virtual reality (VR) exercises targeting emotional demand, cognitive demand, physical demand, and cognitive+physical demand. The sequence of stressors was pseudorandomized. CPT: continuous performance test.

The measures of VR2 trials were captured as follows:

- *Acceptability:* user experience, as assessed by the game user experience satisfaction scale (GUESS)-18 gamer user experience satisfaction scale [67] and the System Usability Scale [68]
- *Tolerability:* user comfort and any side effects from the use of technology, as assessed by the Simulator Sickness Questionnaire (SSQ; captured both before and after VR exposure)
- *Feasibility:* ability to elicit the intended stress response, as measured via physiological response as well as self-report

For the self-report measure, the NASA-TLX [51] and single-item rating for stress questionnaires were used.

Mood Likert scales were also automatically administered in immersive VR before and after the recovery exercise. Participant and VR views were also recorded. After completing the bWell session, participants reviewed a screen recording of the completed virtual task, provided qualitative feedback, and completed the set of questionnaires.

It may be noted that while the NVIM and VTT predict the dependence of vagal tone on task conditions, the NASA-TLX measures 6 aspects of workload. The study design involved 5 conditions, 1 without a stressor and 4 others in which each condition features a different exercise representing specific stressors based on the NASA-TLX (summarized in Table 3). To effectively assess the effects of each condition on stress responses, we proposed a within-*participants* design, where each participant completed all conditions of the experiment. The order of the stressor conditions was pseudorandomized across participants. The approach taken to experimentally isolate the contributions of each type of stress required the removal of performance criteria (eg, score) from the tasks, as they may create confounds in the stressor type not related to the experimental manipulation of the condition. However, the cumulative effects of stressors were also investigated using a physical+cognitive stressor dual task, as the inclusion of multiple simultaneous stressors is more representative of everyday stressful situations.

Table 3. Description of the different types of stressor demands for different virtual reality (VR) exercises. bWell VR exercises elicited different types of demands, as defined by the NASA-Task Load Index workloads. The main stress or demand captured for each VR exercise is shown in italics.

	Mental demand	Physical demand	Temporal demand	Performance	Effort	Frustration
Relaxing nature environment (*tent*: nature scene)	No	No	No	No	No	No
Emotional stressor (*city*: busy city scene)	No	No	No	No	No	*Yes* [a]
Cognitive stressor (*mole*: dynamic bimanual go or no-go task)	*Yes* [a]	No	Yes[a]	Yes[a]	Yes[a]	No
Physical stressor (*stroll*: running task)	No	*Yes* [a]	No	Yes[a]	Yes[a]	No
Physical+cognitive stressor (*stroll+CPT:* running with continuous performance test)	*Yes* [a]	*Yes* [a]	No	Yes[a]	Yes[a]	No

[a]Demands elicited by the bWell virtual reality exercises.

We set thresholds to the NASA-TLX as a self-reported measure to ensure that the designed VR exercises targeted the intended type of demand, that is, *cognitive, emotional, physical,* and *dual (ie, physical+cognitive)*. We defined the following threshold in terms of NASA-TLX score (score range=0-10): a score>5 indicates that the demand is being elicited (eg, *City* for emotional demand, *Mole* for cognitive demand, *Stroll* for physical demand, and *Stroll+CPT* for physical and cognitive demand; Table 4).

An additional aim was to not elicit demand related to poor performance because this may lead to a potential confound. In the NASA-TLX, performance is reverse coded (ie, 0=perfect performance). As such the threshold was set to<4, so that that participant should feel as if they performed well in all exercises. Finally, since all exercises invoked a type of demand except the one involving emotional demand, they were expected to involve an element of effort.

Table 4. Description of the thresholds for NASA-Task Load Index (NASA-TLX) scores for eliciting different stress demands. We set the NASA-TLX subscale scores to above the threshold, for the different types of stress demands that were elicited by each virtual reality exercise in Table 3.

	Frustration	Physical demand	Mental demand	Performance
City	>5[a]	—[b]	—	<4
Mole	<4	<4	>5[a]	<4
Stroll	—	>5[a]	—	<4
Stroll+CPT[c]	—	>5[a]	>5[a]	—

[a]NASA-TLX subscale scores that were above the threshold.

[b]Not available.

[c]CPT: continuous performance test.

More specifically, we targeted the following scores:

- *City*—a NASA-TLX score of >5 for *frustration,* which would indicate an emotional demand. We also intended for all other NASA-TLX scores to remain <4 to solely evoke an emotional response. Given the passive nature of this exercise, other demands should not be induced.
- *Mole*—a NASA-TLX score of >5 for *mental,* which would indicate a cognitive demand. A score of >5 would indicate that these demands were elicited. We also targeted *frustration* and *physical* NASA-TLX scores of <4 to solely target demands related to cognitive performance.
- *Stroll*—a NASA-TLX score of >5 for *physical,* which would indicate a physical demand. A demand of >5 would indicate that this was the case. We also targeted *mental* and *frustration* NASA-TLX scores of <4 to solely target demands related to physical performance.
- *Stroll+CPT*–a NASA-TLX score of >5 for *mental* and *physical,* which would indicate a dual cognitive and physical demand. We also intended to maintain a *frustration* score

of <4 because we wanted to target solely cognitive and physical demands.

In addition, we sought to verify that *mental* demands are comparable between the *Mole* and *Stroll+CPT* tasks.

Verification Trials

Verification trials were conducted using HTC Vive Pro Eye on 3 healthy participants according to the protocol defined during the study design.

Ethical Considerations

Ethics approval for the study was obtained from the Research Ethics Board (REB) of the National Research Council (NRC 2020-2023 and NRC 2023-46) and from Defence Research and Development Canada Human Research Ethics Committee (HREC 2022-032). Informed consent was obtained before the start of the study. Data were deidentified before any analysis. There was no compensation paid to the participants.

Results

The results of the verification trials are reported below in terms of tolerability, acceptability, and feasibility.

Tolerability: Cybersickness During the VR Sessions

We first checked whether the VR exercises designed in our study induced cybersickness using SSQ [69]. Table S2 in Multimedia Appendix 1 shows the scores for SSQ for pre- and post-VR sessions for 3 participants. Total SSQ scores were obtained from the scores of the 16 items of the questionnaire [70]. Results showed negligible score (total SSQ score: 0-5), which are within the acceptable range in the literature [71]. Thus, our results show that the participants did not experience adverse effects, especially cybersickness during the VR exercises.

Acceptability: User Experience

Using the GUESS-18 scale [67], we checked whether the VR exercises designed in our study were immersive. GUESS-18 has 9 subscales with 2 statements associated with each subscale. Of the 9 subscales, we did not include "social connectivity," as there was no interaction between participants for the VR exercises. Table S3 in Multimedia Appendix 1 shows the scores of the 8 subscales of GUESS-18 for the 3 participants. The subscale scores of the GUESS were computed by averaging the scores within each subscale. Then, the overall GUESS score was computed by summing the subscale scores [67]. We observed an overall GUESS score of 43.17 (Table S3 in Multimedia Appendix 1), which aligns well with previous findings [72] and indicates the usability and immersiveness of the VR exercises designed in our study. Regarding the individual subscales, we observed greater average scores for *usability*, *play engrossment,* and *personal gratification* and lower average scores for *creative freedom* and *enjoyment* (Figure 4).

Figure 4. Findings of user acceptability. User ratings on the game user experience satisfaction scale (GUESS)-18 questionnaires showing greater average scores for the subscales usability, play engrossment, and personal gratification and lower average scores for creative freedom and enjoyment. An overall GUESS score of 43.17 (Table S3 was observed, which aligns well with previous findings, indicating the usability and immersiveness of the virtual reality exercises designed in our study.
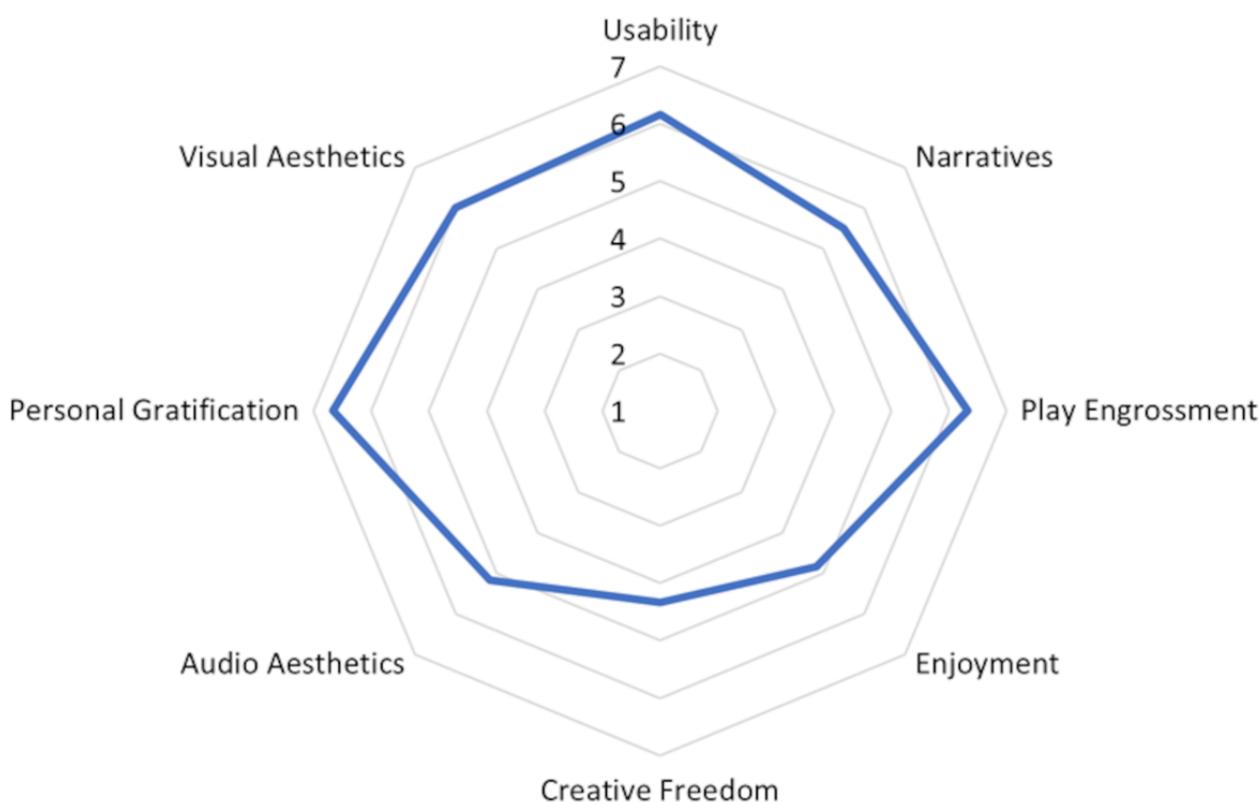
Feasibility: Do the VR Exercises Induce Genuine Stress Responses (Perceived)?

We evaluated the NASA-TLX scores to investigate the types of stressors and their associated demands. Figure 5 summarizes the data for NASA-TLX from verification trials of 3 participants.

Figure 5. Comparative analysis of different types of demands for different stressors. User ratings on the NASA-Task Load Index (NASA-TLX) revealed that participants perceived different types of demands (eg, mental, physical, and temporal) for the different VR exercises. Here, P1, P2, and P3 denote participant 1, participant 2, and participant 3, respectively, and the dashed line represents the threshold set for different demands associated with the different stressors (exercises). CPT: continuous performance test.
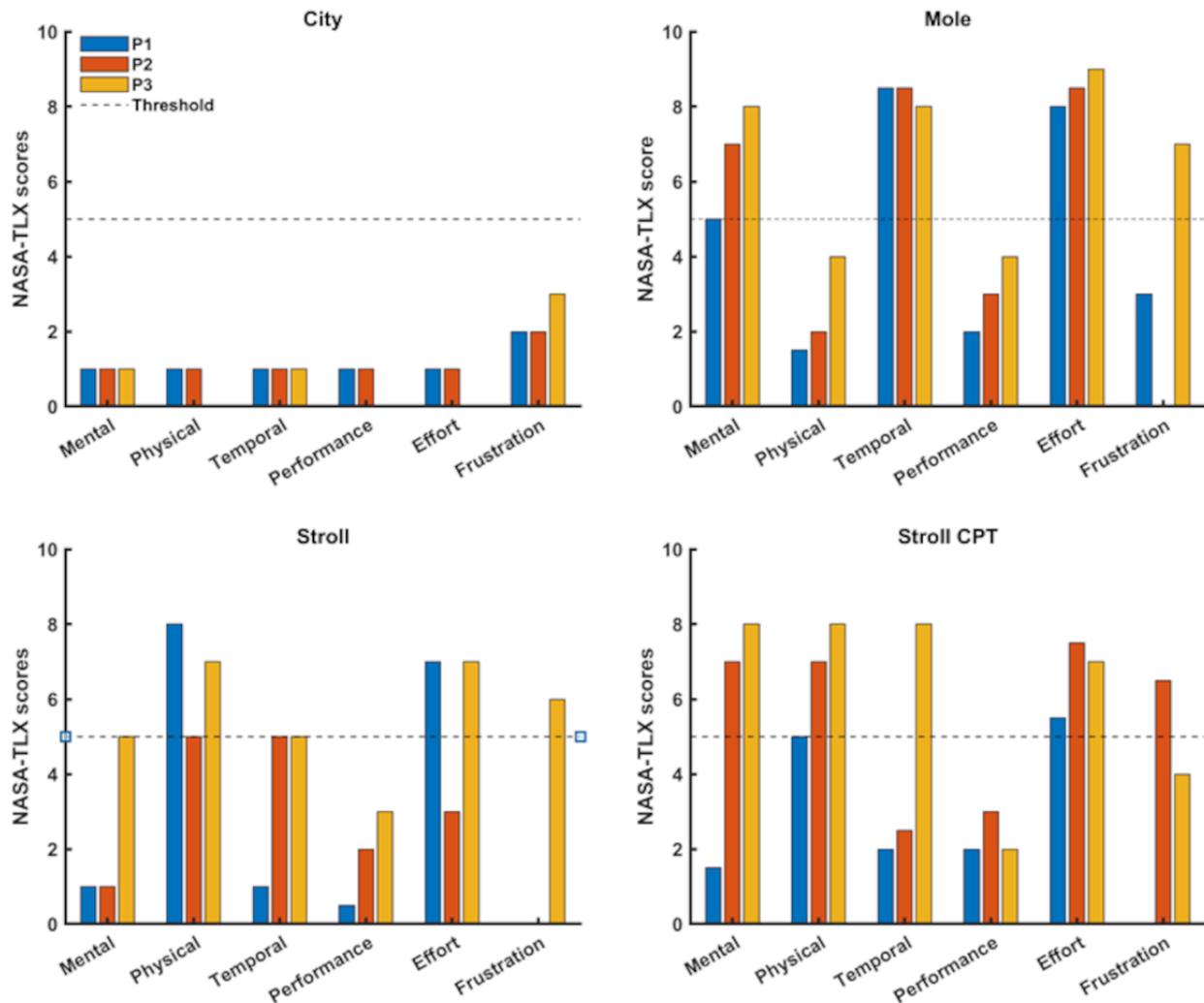
On the basis of the designed thresholds for the different stressors and their associated demands (refer to the Methods section), we observed the following for each of the bWell exercises (Figure 5), which led to further iterative development:

- *City*: a NASA-TLX rating of >5 for *frustration* was not achieved, indicating a need for adapting the exercise settings and creating new functionalities. All other demands were also not elicited (<4). In particular, users felt that they performed well (<4), and as such, it was verified that there would not be any confounds in *frustration* being elicited due to emotional response related to feelings of poor performance rather than VR stimuli.
- *Mole*: a cognitive demand was elicited with a *mental* demand of >5. We also observed a high level of *temporal* demand and associated *effort*.
- *Stroll*: a *physical* demand and associated *effort* was elicited (>5).
- *Stroll+CPT*: a dual cognitive (*mental*) and *physical* demand was elicited (>5), along with associated *effort*.

Because cognitive demand was elicited in both tasks, similar to the Mole task, the 2 can be compared in the analysis of

cognitive demand, even though they involve different cognitive tasks. It appears that *temporal* demand may also be elicited but was not conclusive at this stage because of high variability in the ratings. This may be clarified with results from additional participant trials from the full study. From Table 4, we observed that this demand was not consistently present. This may be attributed to the fact that this was a dual task, where the *temporal* demand may be felt for one aspect (eg, running) but not for the other (eg, cognitive task), and when prompted by the NASA-TLX, the participant may respond with only 1 task in mind.

Overall, the results showed that we achieved eliciting different stress demands as intended in our study design: for *Stroll* to exert a *physical* demand, *Mole* to exert a *mental* demand, and a combination of exercises (*Stroll+CPT*) to exert both *mental* and *physical* demands. At the same time, we observed that participants felt they performed well in all exercises, minimizing confounds due to feelings of poor performance. These verification trials allowed us to finalize the exercise configurations to be used for the full study trials.

Feasibility: Do the VR Exercises Induce Genuine Stress Responses (Physiological)?

To evaluate the physiological stress responses, we examined the HRV measures obtained, specifically the root mean square of successive differences (RMSSD) between normal heartbeats. We adopted RMSSD as the HRV measure because previous studies have shown that RMSSD reflects vagal tone [38,73] and is relatively free of respiratory influences. It was seen that the data were quite noisy, artifacts were present, and the RMSSD values from the raw PPG data were too high. For example, values in the range of 150 to 300 were obtained, which is much higher than normal. Discussion with the Emteq team revealed that this was likely because the measures were sampled in a dynamic VR environment, particularly where in some instances, users were physically running during PPG data acquisition. Emteq advised using their signal quality metric and using only the data where good quality was indicated.

To have better control of the data processing, the team decided instead to implement an in-house data processing pipeline. For this, the raw PPG signals were processed using the Python library *HeartPy* [74]. On the basis of the *HeartPy* documentation, which states that RMSSD is extremely sensitive to incorrectly detected heartbeat locations, any RMSSD value artifacts (RMSSD does not typically exceed 130) were removed and peak interpolation was used to minimize the number of incorrectly placed peaks due to low sampling rate [74]. After these steps, an improved signal-to-noise ratio was observed and the resultant RMSSD values fell well within the normal range expected (between 30 and 70).

We then compared HRV across each of the different conditions for 2 individuals. In the example shown in Figure 6, it can be seen that a given task appears to elicit a change in response. It is also evident that baseline, reactivity, and recovery HRV patterns were distinct for the 2 sample individuals. As can be seen in the lower panel of Figure 6, some instances of residual noise remain in the Mole exercise after processing. Given that the data are being collected during interactive tasks with a good degree of movement, a high noise-to-signal ratio was expected. The next steps will involve exploring other techniques, particularly for motion artifact removal, with the goal of further reducing noise in the signal.

Figure 6. Characterization of distinct patterns of heart rate variability (HRV) for different virtual reality (VR) exercises. HRV data (root mean square of successive differences [RMSSD]) for the different VR (bWell) exercises are shown for 2 participants (row 1 for participant 1 and row 2 for participant 2) along y-axis and time along x-axis. The plots illustrate distinct HRV responses for the different exercises as well as individual variability in the HRV responses. CPT: continuous performance test.
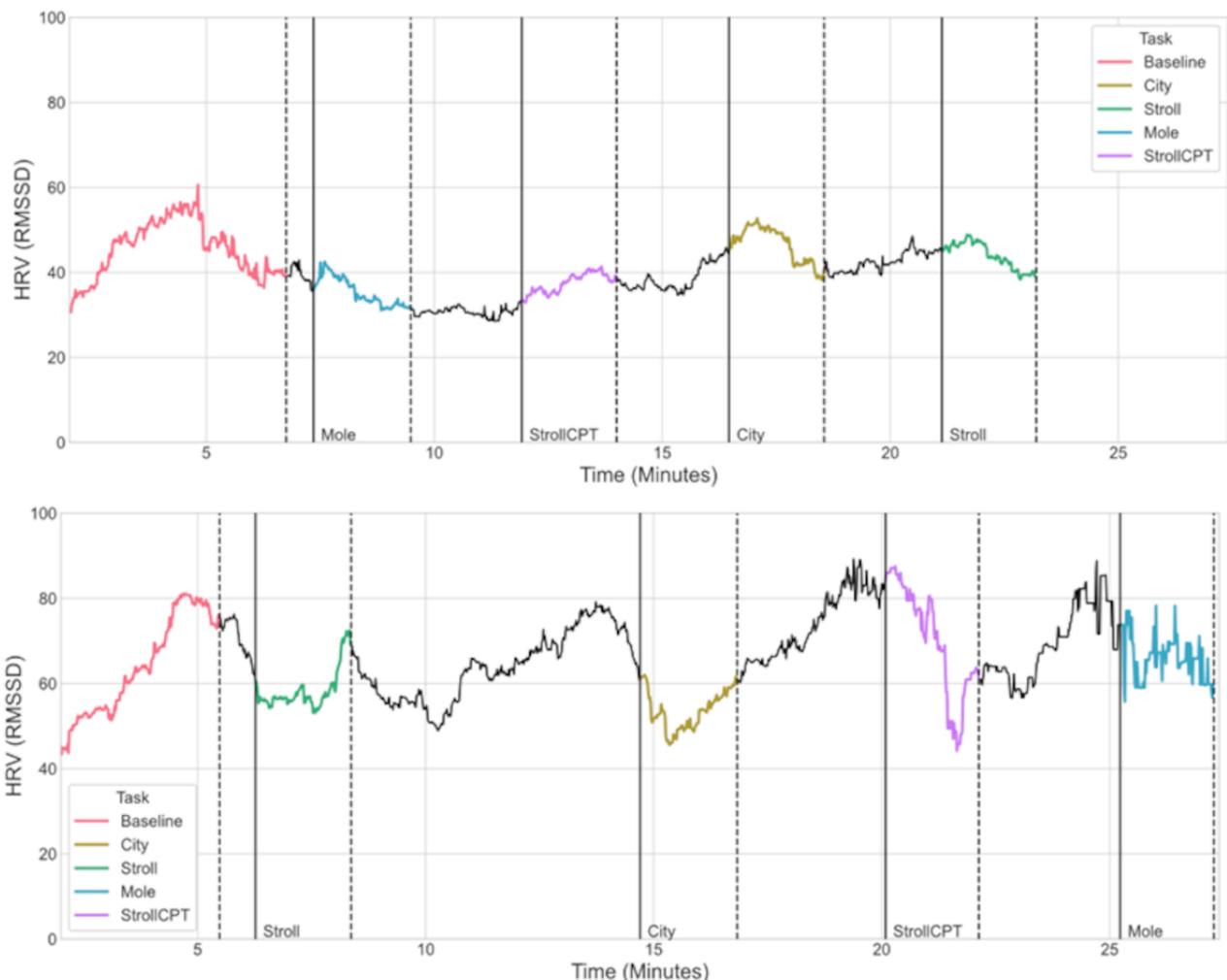
Discussion

Principal Findings

In this paper, we described the development process of a study design for testing the ability of immersive VR exercises to elicit genuine stress responses. The study development followed recommendations for use of VR in clinical trials [32] as well as guidance in developing complex interventions for improving health care [33]. On the basis of these guidelines, our study comprised a combination of *prevailing theories and frameworks* (eg, NVIM and the VTT), *experiences of users,* and *iterative actions during development,* which resulted in concrete outcomes [33]. More specifically, we built on our validated VR platform *bWell* [29] to integrate physiological sensors (PPG) and assessed how different types of VR exercises target different types of stressors (characterized by the NASA-TLX). The results of verification trials showed that the designed VR exercises induced negligible cybersickness, were well tolerated, and demonstrated high user engagement. In addition, in alignment with our hypotheses, the different VR exercises successfully elicited the intended stress demands (as measured with NASA-TLX subscales, eg, mental, physical, and temporal demands), along with distinct physiological responses.

In line with the recommended guidelines for VR trial design in the health domain [44], our study design considered both tolerability and acceptability. Keeping in mind user comfort during VR, the exercises were designed in such a way that the duration for the entire session lasted only 25 minutes. In addition, during the VR session, every exercise was followed by a 2-minute recovery state. By allowing the participants to experience alternating blocks of stressor tasks interspersed with recovery, this study design was used as a guide to evaluate the technical feasibility and scope for the phase of verification trials. Indeed, the results of the verification trials showed user comfort and tolerability, especially VR as an acceptable framework with *negligible* experience of cybersickness (assessed with SSQ scores; Table S2 in Multimedia Appendix 1). As a measure of acceptability, results of verification trials using the GUESS-18 scores showed a high degree of "*presence*" during our VR sessions (Table S3 in Multimedia Appendix 1 and Figure 4), consistent with earlier findings [72]. Because the sense of "presence" (or the subjective feeling of "being there") characterizes the experience of users, results showing a high degree of presence indicate that the VR exercises designed for the study created the sensation of physical presence and resulted in users behaving and reacting as in the real world, highlighting the importance of a good user experience and engagement. Best practices for HRV data collection were also considered for the study design. In experimental settings, it is recommended that the baseline HRV should be recorded in a "nontask situation that best controls for the presence of task comparisons" [65]. Consequently, in this study, baseline was recorded in VR and in a similar position to how the participants would complete the task. Specifically, baseline was recorded with the user on the treadmill and passively reexperiencing the VR environments. The session to collect baseline HRV was designed as follows: eyes open, standing, spontaneous breathing, and free to look around and immerse themselves in the VR environment. In addition

and according to guidance on HRV reporting, RMSSD was selected as the HRV measure because it is appropriate for short sampling times (ie, <5 minutes) [41].

While the results of verification trials determined the feasibility and validity of our study protocol, some results indicated that modifications were also required. For example, the NASA-TLX scores showed high levels of *Frustration,* which captured participants' physical demands or frustration related to performance and not necessarily due to an emotional stressor. Because our original goal was to investigate the effects of emotion, we made certain modifications in our bWell exercises to elicit emotional responses in the participants, such as amplifying the noise pollution in the virtual environment. The results of the verification trials also indicated that we need to modify the administration of the NASA-TLX, which currently probes *Frustration* related to both performance and emotion. Instead of asking *"How insecure, discouraged, irritated, stressed and annoyed were you,"* the question will be modified to include only the elements related to emotion *"How irritated and/or annoyed were you?"*

Implications

It may be mentioned that the overarching aim for the integration of behavioral and physiological outputs was to eventually enable systematic and individualized assessment of stress response. Responses to stress are often multifaceted with a profile specific to the individual. A given stressor may cause a response in one person but not necessarily in another, which may be attributed to the influence of personal experiences. The result is that everyone does not experience stress in the same way. In an effort to standardize the demands placed on individuals, we relied on several mechanisms to adapt the VR exercises in real time and customized them to the participant. For cognitive demand, we aimed to elicit a uniform cognitive demand among varied individuals. To do so, we leveraged the adaptive algorithms in *bWell* to control the difficulty level based on performance (for *Mole* exercise). The adaptive mechanism is designed to maintain an 80% success rate, ensuring the exercise is neither too easy (boring) nor too difficult (demotivating). For physical demand, we made use of the heart rate biofeedback to ensure consistent physical demand because running at a specified speed may be more or less exerting for one individual versus another. To achieve this, a heart rate gauge was displayed to the participant to encourage them to maintain a level of physical activity that corresponds to moderate intensity. This measure is additionally tailored to the individual because it considers the age of the participant, where moderate physical activity is calculated at 50% to 70% of maximum heart rate (maximum heart rate=220 – age).

Study Limitations

There were certain limitations in our study. First, participants reported issues with disengagement and a lack of frustration during the VR exercise (city). The reason for this could be the low resolution of the visual assets in the scene, a concern that will be addressed in future work by improving the resolution. Second, there could be a learning effect in terms of interactions in VR for participants who had no prior VR experience. In the full study (planned next steps), participants will be given

tutorials for immersion in VR. Finally, the verification trials were done with few participants (N=3), limiting our ability to generalize findings. However, we would like to emphasize that our study was designed to identify, create, and implement an experimental setup for testing the ability of immersive VR exercises to elicit genuine responses to different types of stressors. In addition, it aimed to assess the feasibility, tolerability, and acceptability of these exercises through verification trials. This is consistent with previous studies that focused on study development [75]. Because the purpose of the verification trials was to show the feasibility, tolerability, and user acceptability, we collected data from only 3 participants, with the intention to run full trials in the future. It is worth noting that, despite the limited number of participants, we were able to demonstrate the feasibility, tolerability, and user acceptability of the study.

Conclusions

In conclusion, we identified, designed, and implemented an experimental setup for testing the ability of immersive VR exercises to elicit genuine responses corresponding to different types of stressors and tested its feasibility, tolerability, and acceptability using verification trials. While only a small number of trials were carried out with participants, the verification trials provided the support required to carry out a full study (planned next steps).

A strong point of our study design was the collection of several physiological measures that permitted the collection of quantitative measures complementing that of self-report with the NASA-TLX. With the capture of measures such as HRV and facial expressions, it will be possible to compare how the bWell exercises may physiologically elicit response versus what users perceive to be elicited. Such evaluations may serve to obtain a more complete picture of the in-situ user state and may permit gaining insights on the use of traditional self-report, which has been critiqued as an unreliable index of functional outcomes [76]. In this regard, future work will explore multivariate analysis such as partial least squares [77] and ML techniques such as random forest [78] for predicting individual user responses to different stressors based on multiple input data, including sensor data (eg, HRV and facial expressions) and VR-based behavioral data (eg, response time). Such work will help in identifying and delineating the individual features that confer stress resilience and will guide us in helping individuals reach their full potential.

Data Availability

The datasets generated and analyzed during this study are not publicly available due to legislation regarding privacy and security. None of the data for the experiments reported here are available, and none of the experiments were preregistered.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Description of the timeline for different task and recovery, Simulator Sickness Questionnaire scores for pre– and post–virtual reality session, and subjective ratings of game user experience satisfaction scale-18 questionnaires.
[DOCX File , 390 KB-Multimedia Appendix 1]

References

1. Yaribeygi H, Panahi Y, Sahraei H, Johnston TP, Sahebkar A. The impact of stress on body function: a review. EXCLI J. 2017;16:1057-1072. [FREE Full text] [doi: 10.17179/excli2017-480] [Medline: 28900385]

2. Chen B, Wang L, Li B, Liu W. Work stress, mental health, and employee performance. Front Psychol. Nov 8, 2022;13:1006580. [FREE Full text] [doi: 10.3389/fpsyg.2022.1006580] [Medline: 36425815]

3. Alberdi A, Aztiria A, Basarab A. Towards an automatic early stress recognition system for office environments based on multimodal measurements: a review. J Biomed Inform. Mar 2016;59:49-75. [FREE Full text] [doi: 10.1016/j.jbi.2015.11.007] [Medline: 26621099]

4. Pedrosa AL, Bitencourt L, Fróes AC, Cazumbá ML, Campos RG, de Brito SB, et al. Emotional, behavioral, and psychological impact of the COVID-19 pandemic. Front Psychol. Oct 2, 2020;11:566212. [FREE Full text] [doi: 10.3389/fpsyg.2020.566212] [Medline: 33117234]

5. Saladino V, Algeri D, Auriemma V. The psychological and social impact of COVID-19: new perspectives of well-being. Front Psychol. Oct 2, 2020;11:577684. [FREE Full text] [doi: 10.3389/fpsyg.2020.577684] [Medline: 33132986]

6. Ebner K, Singewald N. Individual differences in stress susceptibility and stress inhibitory mechanisms. Curr Opin Behav Sci. Apr 2017;14:54-64. [doi: 10.1016/j.cobeha.2016.11.016]

7. Sapolsky RM. Stress and the brain: individual variability and the inverted-U. Nat Neurosci. Oct 25, 2015;18(10):1344-1346. [doi: 10.1038/nn.4109] [Medline: 26404708]

8. Bangasser DA, Valentino RJ. Sex differences in stress-related psychiatric disorders: neurobiological perspectives. Front Neuroendocrinol. Aug 2014;35(3):303-319. [FREE Full text] [doi: 10.1016/j.yfrne.2014.03.008] [Medline: 24726661]

9. Russo SJ, Murrough JW, Han MH, Charney DS, Nestler EJ. Neurobiology of resilience. Nat Neurosci. Nov 14, 2012;15(11):1475-1484. [FREE Full text] [doi: 10.1038/nn.3234] [Medline: 23064380]

10. McEwen BS. Sex, stress and the hippocampus: allostasis, allostatic load and the aging process. Neurobiol Aging. 2002;23(5):921-939. [doi: 10.1016/s0197-4580(02)00027-1] [Medline: 12392796]

11. Ulrich-Lai YM, Herman JP. Neural regulation of endocrine and autonomic stress responses. Nat Rev Neurosci. Jun 13, 2009;10(6):397-409. [FREE Full text] [doi: 10.1038/nrn2647] [Medline: 19469025]

12. Clark A, Mach N. Exercise-induced stress behavior, gut-microbiota-brain axis and diet: a systematic review for athletes. J Int Soc Sports Nutr. Apr 01, 2016;13(1):43. [FREE Full text] [doi: 10.1186/s12970-016-0155-6] [Medline: 27924137]

13. McCraty R, Shaffer F. Heart rate variability: new perspectives on physiological mechanisms, assessment of self-regulatory capacity, and health risk. Glob Adv Health Med. Jan 2015;4(1):46-61. [FREE Full text] [doi: 10.7453/gahmj.2014.073] [Medline: 25694852]

14. Amano M, Kanda T, Ue H, Moritani T. Exercise training and autonomic nervous system activity in obese individuals. Med Sci Sports Exerc. Aug 2001;33(8):1287-1291. [FREE Full text] [doi: 10.1097/00005768-200108000-00007] [Medline: 11474328]

15. Dong JG. The role of heart rate variability in sports physiology. Exp Ther Med. May 2016;11(5):1531-1536. [FREE Full text] [doi: 10.3892/etm.2016.3104] [Medline: 27168768]

16. Mongin D, Chabert C, Extremera MG, Hue O, Courvoisier DS, Carpena P, et al. Decrease of heart rate variability during exercise: an index of cardiorespiratory fitness. PLoS One. Sep 2, 2022;17(9):e0273981. [FREE Full text] [doi: 10.1371/journal.pone.0273981] [Medline: 36054204]

17. Tobaldini E, Nobili L, Strada S, Casali KR, Braghiroli A, Montano N. Heart rate variability in normal and pathological sleep. Front Physiol. Oct 16, 2013;4:294. [FREE Full text] [doi: 10.3389/fphys.2013.00294] [Medline: 24137133]

18. Siepmann M, Weidner K, Petrowski K, Siepmann T. Heart rate variability: a measure of cardiovascular health and possible therapeutic target in dysautonomic mental and neurological disorders. Appl Psychophysiol Biofeedback. Dec 22, 2022;47(4):273-287. [FREE Full text] [doi: 10.1007/s10484-022-09572-0] [Medline: 36417141]

19. Ahn JW, Ku Y, Kim HC. A novel wearable EEG and ECG recording system for stress assessment. Sensors (Basel). Apr 28, 2019;19(9):23. [FREE Full text] [doi: 10.3390/s19091991] [Medline: 31035399]

20. Attar ET, Balasubramanian V, Subasi E, Kaya M. Stress analysis based on simultaneous heart rate variability and EEG monitoring. IEEE J Transl Eng Health Med. 2021;9:1-7. [doi: 10.1109/jtehm.2021.3106803]

21. Shilton AL, Laycock R, Crewther SG. The Maastricht acute stress test (MAST): physiological and subjective responses in anticipation, and post-stress. Front Psychol. Apr 19, 2017;8:567. [FREE Full text] [doi: 10.3389/fpsyg.2017.00567] [Medline: 28469586]

22. Frisch JU, Häusser JA, Mojzisch A. The trier social stress test as a paradigm to study how people respond to threat in social interactions. Front Psychol. Feb 02, 2015;6:14. [FREE Full text] [doi: 10.3389/fpsyg.2015.00014] [Medline: 25698987]

23. Aspiotis V, Miltiadous A, Kalafatakis K, Tzimourta KD, Giannakeas N, Tsipouras MG, et al. Assessing electroencephalography as a stress indicator: a VR high-altitude scenario monitored through EEG and ECG. Sensors (Basel). Aug 03, 2022;22(15):5792. [FREE Full text] [doi: 10.3390/s22155792] [Medline: 35957348]

24. Ahmad Z, Rabbani S, Zafar MR, Ishaque S, Krishnan S, Khan N. Multilevel stress assessment from ECG in a virtual reality environment using multimodal fusion. IEEE Sensors J. Dec 1, 2023;23(23):29559-29570. [doi: 10.1109/jsen.2023.3323290]

25. Bouchard S, Bernier F, Boivin É, Guitard T, Laforest M, Dumoulin S, et al. Modes of immersion and stress induced by commercial (off-the-shelf) 3D games. J Def Model Simul. May 04, 2012;11(4):339-352. [doi: 10.1177/1548512912446359]

26. van Dammen L, Finseth TT, McCurdy BH, Barnett NP, Conrady RA, Leach AG, et al. Evoking stress reactivity in virtual reality: a systematic review and meta-analysis. Neurosci Biobehav Rev. Jul 2022;138:104709. [doi: 10.1016/j.neubiorev.2022.104709] [Medline: 35644278]

27. Hemmeter U, Stormer R, Mager R, Kuntze M, Mueller-Spahn F, Hennig J, et al. Modification of virtual reality combined with a mental task stimulates cortisol in volunteers. Neuropsychobiology. May 2, 2005;51(3):165-172. [doi: 10.1159/000085208] [Medline: 15838187]

28. Breuninger C, Sláma DM, Krämer M, Schmitz J, Tuschen-Caffier B. Psychophysiological reactivity, interoception and emotion regulation in patients with agoraphobia during virtual reality anxiety induction. Cogn Ther Res. Oct 15, 2016;41(2):193-205. [doi: 10.1007/S10608-016-9814-9]

29. Shaigetz VG, Proulx C, Cabral A, Choudhury N, Hewko M, Kohlenberg E, et al. An immersive and interactive platform for cognitive assessment and rehabilitation (bWell): design and iterative development process. JMIR Rehabil Assist Technol. Nov 03, 2021;8(4):e26629. [FREE Full text] [doi: 10.2196/26629] [Medline: 34730536]

30. Thayer JF, Hansen AL, Saus-Rose E, Johnsen BH. Heart rate variability, prefrontal neural function, and cognitive performance: the neurovisceral integration perspective on self-regulation, adaptation, and health. Ann Behav Med. Apr 8, 2009;37(2):141-153. [doi: 10.1007/s12160-009-9101-z] [Medline: 19424767]

31. Laborde S, Mosley E, Mertgen A. Vagal tank theory: the three Rs of cardiac vagal control functioning - resting, reactivity, and recovery. Front Neurosci. Jul 10, 2018;12:458. [FREE Full text] [doi: 10.3389/fnins.2018.00458] [Medline: 30042653]

32. Birckhead B, Khalil C, Liu X, Conovitz S, Rizzo A, Danovitch I, et al. Recommendations for methodology of virtual reality clinical trials in health care by an international working group: iterative study. JMIR Ment Health. Jan 31, 2019;6(1):e11973. [FREE Full text] [doi: 10.2196/11973] [Medline: 30702436]

33. O'Cathain A, Croot L, Duncan E, Rousseau N, Sworn K, Turner KM, et al. Guidance on how to develop complex interventions to improve health and healthcare. BMJ Open. Aug 15, 2019;9(8):e029954. [FREE Full text] [doi: 10.1136/bmjopen-2019-029954] [Medline: 31420394]

34. Delliaux S, Delaforge A, Deharo JC, Chaumet G. Mental workload alters heart rate variability, lowering non-linear dynamics. Front Physiol. May 14, 2019;10:565. [FREE Full text] [doi: 10.3389/fphys.2019.00565] [Medline: 31156454]

35. Low A, McCraty R. Heart rate variability: new perspectives on assessment of stress and health risk at the workplace. Heart Mind. 2018;2(1):16-27. [doi: 10.4103/hm.hm_11_18]

36. Eggenberger P, Annaheim S, Kündig KA, Rossi RM, Münzer T, de Bruin ED. Heart rate variability mainly relates to cognitive executive functions and improves through exergame training in older adults: a secondary analysis of a 6-month randomized controlled trial. Front Aging Neurosci. Jul 15, 2020;12:197. [FREE Full text] [doi: 10.3389/fnagi.2020.00197] [Medline: 32760267]

37. Grässler B, Hökelmann A, Cabral RH. Resting heart rate variability as a possible marker of cognitive decline. Kinesiol. 2020;52(1):72-84. [doi: 10.26582/k.52.1.9]

38. Thayer JF, Lane RD. A model of neurovisceral integration in emotion regulation and dysregulation. J Affect Disord. Dec 2000;61(3):201-216. [doi: 10.1016/s0165-0327(00)00338-4] [Medline: 11163422]

39. Thayer JF, Lane RD. Claude Bernard and the heart-brain connection: further elaboration of a model of neurovisceral integration. Neurosci Biobehav Rev. Mar 2009;33(2):81-88. [doi: 10.1016/j.neubiorev.2008.08.004] [Medline: 18771686]

40. Pavlov VA, Tracey KJ. The vagus nerve and the inflammatory reflex--linking immunity and metabolism. Nat Rev Endocrinol. Dec 21, 2012;8(12):743-754. [FREE Full text] [doi: 10.1038/nrendo.2012.189] [Medline: 23169440]

41. Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Circulation. Mar 01, 1996;93(5):1043-1065. [Medline: 8598068]

42. Liu CY, Mueller MH, Grundy D, Kreis ME. Vagal modulation of intestinal afferent sensitivity to systemic LPS in the rat. Am J Physiol Gastrointest Liver Physiol. May 2007;292(5):G1213-G1220. [FREE Full text] [doi: 10.1152/ajpgi.00267.2006] [Medline: 17204546]

43. Ozaki N, Sengupta JN, Gebhart GF. Mechanosensitive properties of gastric vagal afferent fibers in the rat. J Neurophysiol. Nov 01, 1999;82(5):2210-2220. [FREE Full text] [doi: 10.1152/jn.1999.82.5.2210] [Medline: 10561400]

44. Laborde S, Mosley E, Thayer JF. Heart rate variability and cardiac vagal tone in psychophysiological research - recommendations for experiment planning, data analysis, and data reporting. Front Psychol. 2017;8:213. [FREE Full text] [doi: 10.3389/fpsyg.2017.00213] [Medline: 28265249]

45. Shaffer F, Ginsberg JP. An overview of heart rate variability metrics and norms. Front Public Health. Sep 28, 2017;5:258. [FREE Full text] [doi: 10.3389/fpubh.2017.00258] [Medline: 29034226]

46. Dekker JM, Crow RS, Folsom AR, Hannan PJ, Liao D, Swenne CA, et al. Low heart rate variability in a 2-minute rhythm strip predicts risk of coronary heart disease and mortality from several causes: the ARIC Study. Atherosclerosis risk in communities. Circulation. Sep 12, 2000;102(11):1239-1244. [doi: 10.1161/01.cir.102.11.1239] [Medline: 10982537]

47. Grassi G. Assessment of sympathetic cardiovascular drive in human hypertension: achievements and perspectives. Hypertension. Oct 2009;54(4):690-697. [doi: 10.1161/hypertensionaha.108.119883]

48. Ronkainen E, Korpelainen JT, Heikkinen E, Myllylä VV, Huikuri HV, Isojärvi JI. Cardiac autonomic control in patients with refractory epilepsy before and during vagus nerve stimulation treatment: a one-year follow-up study. Epilepsia. Mar 06, 2006;47(3):556-562. [FREE Full text] [doi: 10.1111/j.1528-1167.2006.00467.x] [Medline: 16529621]

49. Porges SW. The polyvagal perspective. Biol Psychol. Mar 2007;74(2):116-143. [FREE Full text] [doi: 10.1016/j.biopsycho.2006.06.009] [Medline: 17049418]

50. Smith R, Thayer JF, Khalsa SS, Lane RD. The hierarchical basis of neurovisceral integration. Neurosci Biobehav Rev. May 2017;75:274-296. [doi: 10.1016/j.neubiorev.2017.02.003] [Medline: 28188890]

51. Hart SG, Staveland LE. Development of NASA-TLX (task load index): results of empirical and theoretical research. Adv Psychol. 1988;52:139-183. [doi: 10.1016/S0166-4115(08)62386-9]

52. Porges SW. Cardiac vagal tone: a physiological index of stress. Neurosci Biobehav Rev. 1995;19(2):225-233. [doi: 10.1016/0149-7634(94)00066-a] [Medline: 7630578]

53. Thayer JF, Åhs F, Fredrikson M, Sollers JJ, Wager TD. A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. Neurosci Biobehav Rev. Mar 2012;36(2):747-756. [doi: 10.1016/j.neubiorev.2011.11.009] [Medline: 22178086]

54. Park G, Vasey MW, Van Bavel JJ, Thayer JF. When tonic cardiac vagal tone predicts changes in phasic vagal tone: the role of fear and perceptual load. Psychophysiology. May 27, 2014;51(5):419-426. [doi: 10.1111/psyp.12186] [Medline: 24571084]

55. Porges SW, Doussard-Roosevelt JA, Portales AL, Greenspan SI. Infant regulation of the vagal "brake" predicts child behavior problems: a psychobiological model of social behavior. Dev Psychobiol. Dec 1996;29(8):697-712. [doi: 10.1002/(SICI)1098-2302(199612)29:8<697::AID-DEV5>3.0.CO;2-O] [Medline: 8958482]

56. Porges SW. A phylogenetic journey through the vague and ambiguous Xth cranial nerve: a commentary on contemporary heart rate variability research. Biol Psychol. Mar 2007;74(2):301-307. [FREE Full text] [doi: 10.1016/j.biopsycho.2006.08.007] [Medline: 17055142]

57. Crameri L, Hettiarachchi IT, Hanoun S. Effects of dynamic resilience on the reactivity of vagally mediated heart rate variability. Front Psychol. Jan 20, 2020;11:579210. [FREE Full text] [doi: 10.3389/fpsyg.2020.579210] [Medline: 33551903]

58. Beauchaine TP, Gatzke-Kopp L, Mead HK. Polyvagal Theory and developmental psychopathology: emotion dysregulation and conduct problems from preschool to adolescence. Biol Psychol. Mar 2007;74(2):174-184. [FREE Full text] [doi: 10.1016/j.biopsycho.2005.08.008] [Medline: 17045726]

59. Fisher JP. Autonomic control of the heart during exercise in humans: role of skeletal muscle afferents. Exp Physiol. Mar 02, 2014;99(2):300-305. [FREE Full text] [doi: 10.1113/expphysiol.2013.074377] [Medline: 23995102]

60. Fu Q, Levine BD. Exercise and the autonomic nervous system. Handb Clin Neurol. 2013;117:147-160. [doi: 10.1016/B978-0-444-53491-0.00013-4] [Medline: 24095123]

61. Kimhy D, Crowley O, McKinley PS, Burg M, Lachman M, Tun P, et al. The association of cardiac vagal control and executive functioning--findings from the MIDUS study. J Psychiatr Res. May 2013;47(5):628-635. [FREE Full text] [doi: 10.1016/j.jpsychires.2013.01.018] [Medline: 23434176]

62. Proulx C, Cabral A, Choudhury N, Debergue P. Acceptability study of a novel immersive cognitive care platform for remediation of cognitive deficits. Annu Rev Cyber Therapy Telemed. 2018;16:43-49.

63. Hernandez Hernandez ME, Michalak E, Choudhury N, Hewko M, Torres I, Menon M, et al. Co-design of a virtual reality cognitive remediation program for depression (bWell-D) with patient end users and clinicians: qualitative interview study among patients and clinicians. JMIR Serious Games. May 07, 2023;11:e43904. [FREE Full text] [doi: 10.2196/43904] [Medline: 37027183]

64. Cabral A, Choudhury N, Proulx C, Harmouche R, Kohlenberg E, Deberguevaluation OC. Evaluation of cybersickness in a passive walking virtual reality cognitive exercise. Annu Rev CyberTherapy Telemed. 2019;17:59-64. [FREE Full text]

65. Quintana DS, Heathers JA. Considerations in the assessment of heart rate variability in biobehavioral research. Front Psychol. Jul 22, 2014;5:805. [FREE Full text] [doi: 10.3389/fpsyg.2014.00805] [Medline: 25101047]

66. Gnacek M, Broulidakis J, Mavridou I, Fatoorechi M, Seiss E, Kostoulas T, et al. emteqPRO—fully integrated biometric sensing array for non-invasive biomedical research in virtual reality. Front Virtual Real. Mar 11, 2022;3:35. [doi: 10.3389/frvir.2022.781218]

67. Keebler Assoc JR, Shelstad WJ, Google DC, Chaparro BS, Phan Google MH. Validation of the GUESS-18. J Usability Stud. 2020;16(1):49-62. [doi: 10.5555/3532747.3532752]

68. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. London, UK. CRC Press; 1996.

69. Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG. Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. Int J Aerosp Psychol. Jul 1993;3(3):203-220. [doi: 10.1207/s15327108ijap0303_3]

70. Bouchard S, Robillard G, Renaud P. Revising the factor structure of the simulator sickness questionnaire. Annu Rev CyberTherapy Telemed. 2007;5:128-137.

71. Bouchard S, St-Jacques J, Renaud P, Wiederhold B. Side effects of immersions in virtual reality for people suffering from anxiety disorders. J Cyber Ther Rehabil. 2009;2(2):127-137. [FREE Full text]

72. Carman WA. Validation of the GUESS-18 for the usability of a virtual reality racing game. LSU Master's Theses. URL: https://digitalcommons.lsu.edu/gradschool_theses/5606 [accessed 2023-06-14]

73. Kleiger RE, Stein PK, Bigger JT. Heart rate variability: measurement and clinical utility. Ann Noninvasive Electrocardiol. Jan 2005;10(1):88-101. [FREE Full text] [doi: 10.1111/j.1542-474X.2005.10101.x] [Medline: 15649244]

74. van Gent P, Farah H, van Nes N, van Arem B. HeartPy: a novel heart rate algorithm for the analysis of noisy signals. Transp Res Part F Traffic Psychol Behav. Oct 2019;66:368-378. [doi: 10.1016/j.trf.2019.09.015]

75. Watland S, Nes LS, Hanson E, Ekstedt M, Stenberg U, Børøsund E. The caregiver pathway, a model for the systematic and individualized follow-up of family caregivers at intensive care units: development study. JMIR Form Res. May 25, 2023;7:e46299. [FREE Full text] [doi: 10.2196/46299] [Medline: 37097744]

76. Huckvale K, Venkatesh S, Christensen H. Toward clinical digital phenotyping: a timely opportunity to consider purpose, quality, and safety. NPJ Digit Med. Sep 6, 2019;2(1):88. [FREE Full text] [doi: 10.1038/s41746-019-0166-1] [Medline: 31508498]

77. McIntosh AR, Lobaugh NJ. Partial least squares analysis of neuroimaging data: applications and advances. Neuroimage. 2004;23 Suppl 1:S250-S263. [doi: 10.1016/j.neuroimage.2004.07.020] [Medline: 15501095]

78. Khundrakpam B, Booij L, Jeon S, Karama S, Tohka J, Evans AC. Individualized prediction of future cognition based on developmental changes in cortical anatomy. NeuroImage Rep. Dec 2022;2(4):100127. [doi: 10.1016/j.ynirp.2022.100127]

Abbreviations

ANS: autonomic nervous system
CPT: continuous performance test
ECG: electrocardiogram
EMG: electromyography
GUESS: game user experience satisfaction scale
HRV: heart rate variability
ML: machine learning
NASA-TLX: NASA-Task Load Index
NVIM: Neurovisceral Integration Model
PPG: photoplethysmography
RMSSD: root mean square of successive differences
SSQ: Simulator Sickness Questionnaire
UI: user interface
VR: virtual reality
VTT: Vagal Tank Theory

XSL•FO
RenderX