Training for Stressful Operations using Adaptive Systems: Conceptual Approaches and Applications

Tor Finseth, Michael Dorneich, Nir Keren, Warren Franke, and Stephen Vardeman
Iowa State University
Ames, Iowa
tfinseth@iastate.edu, dorneich@iastate.edu, nir@iastate.edu, wfranke@iastate.edu, vardeman@iastate.edu

ABSTRACT

Training systems are a potential stress countermeasure by simulating high stress conditions in a safe and controlled environment. Training often involves increasing the complexity of scenarios over time until trainees can reliably execute the task. However, several limitations reduce the long-term retention and robustness of training during intense acute stress felt in-situ. These limitations include generalized training practices that are not tailored to the individual, the unreliability of self-reported subjective stress, over-trained skills that are inflexible and not robust to novel stressors, training pedagogies that focus too much on task proficiency rather than how the individual manages stress during task execution, and ambiguity for when trainers should increase/decrease training difficulty.

Adaptive systems are proposed as a supplement while training individuals to maintain task performance. An adaptive system is a joint human-computer system that is able to automate functions/tasks to varying degrees to help the user, often without explicit instruction. In the context of training for stressful tasks, an adaptive system could detect and monitor stress using physiological sensors and machine learning and use this information to modify scenarios to provide individualized training. This would allow coping skills to be practiced without overwhelming or under-stimulating the trainee's stress tolerance, adapt training according to proficiency in both task execution and physiological stress, and offer clear benchmarks for when to increase/decrease training difficulty. When coupled with a simulated training environment, an adaptive system could adapt training by altering the task procedure and implicitly changing the task environment to help the user build resilience to novel stressors.

This paper presents conceptual approaches and applications for training stressful operations using adaptive systems. A generic adaptive stress training system framework is described along with recommendations based on an experimental example in the spaceflight domain for training emergency fire procedures in a virtual reality International Space Station.

ABOUT THE AUTHORS

Dr. Tor Finseth earned his PhD at Iowa State University in Aerospace Engineering and Human-Computer Interaction. His research interests include the study of human stress response in context of occupational training and advanced technologies such as virtual reality, time-series machine learning, stress detection, adaptive systems.

Dr. Michael Dorneich is an Associate Professor in the Department of Industrial and Manufacturing Systems Engineering at Iowa State University. Specific research and application areas include human factors, cognitive engineering, adaptive automation and adaptive interfaces, distributed systems, interactive learning environments, and decision-support systems.

Dr. Nir Keren is an Associate Professor of Occupational Safety and a graduate faculty at the Virtual Reality Application Center at Iowa State University. He develops virtual reality environment to study mission critical occupation employees' performance under stress.

Dr. Warren Franke is a Professor in the Department of Kinesiology at Iowa State University. His expertise is in cardiovascular physiology and his research interests are focused on health risks in emergency responders.

Dr. Stephen Vardeman is a Professor in the Department of Industrial and Manufacturing Systems Engineering and Department of Statistics at Iowa State University. His expertise are in Statistical Machine Learning and Engineering Analytics.

Training for Stressful Operations using Adaptive Systems: Conceptual Approaches and Applications

Tor Finseth, Michael Dorneich, Nir Keren, Warren Franke, and Stephen Vardeman
Iowa State University
Ames, Iowa
tfinseth@iastate.edu, dorneich@iastate.edu, nir@iastate.edu, wfranke@iastate.edu, vardeman@iastate.edu

INTRODUCTION

Training aims to develop necessary competences to ensure personnel are prepared to meet mission requirements. These competencies encompass not only the technical and procedural knowledge required to do tasks, but also the cross-functional skills required to succeed in a complex and frequently stressful environment such as multi-tasking, situational awareness, concentration, and adaptability (Robson & Manacapilli, 2014). However, exposure to high-levels of stress can degrade task performance and potentially place the operator in a dangerous situation. Even individuals with high proficiency in skill training can succumb to the stress in the operational environment (Orasanu & Backer, 1996). However, inadequate training to handle stress 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.

Several training limitations reduce the robustness of skills during intense acute stress felt *in-situ*. First, using generalized training practices to train individuals to handle stress is difficult because of the wide variation on what is perceived as stressful. Additional time and resources are often necessary to tailor training to the individual to help them build cognitive competencies (Robson & Manacapilli, 2014). Second, training using self-reports can have limited reliability as an assessment of the individual's subjective stress. Self-reported stress may not reflect the physiological stress response, may have low sensitivity to rapidly changing emotional states, and may be fabricated to avoid social judgement from colleagues or superiors (Campbell & Ehlert, 2012). Third, training pedagogies focus too much on task proficiency, which may be rigid and inflexible when faced with novel stressors, rather than how the individual manages stress during task execution (Delahaij et al., 2006). More effort is needed to train competencies to respond to unpredictable acute stress (Robson & Manacapilli, 2014). Replacing task-based rehearsal with competency training may be better suited to help individuals perform in a variety of tasks and the ability to deal with unexpected, untrained-for emerging stressors (Barshi & Dempsey, 2016). Lastly, there is ambiguity about when trainers should increase/decrease training difficulty. Controlling the training environment is important when preparing individuals to perform in high-stress environments. However, using task proficiency as a gauge for changing the difficulty level during repetitive training rehearsal may decrease training benefits by not accounting for the individual's cognitive state (e.g., confidence, anxiety level; Keinan & Friedland, 1996).

Automated training systems can be designed to be adaptive with the goal of helping reduce acute stress levels and train competencies in preparation for high-stress operations. Sensing the human state in real-time can enable the system to adapt a training environment based on the user's proficiency and momentary context. This can offer training advantages for individuals with different abilities to cope with stressful events. Systems would require minimal resources and trainer support, while increasing precision on when to change training difficulty. An adaptive system could ensure stress levels are maintained within a suitable physiological range so the individual is not overwhelmed. Adaptive stress training could promote competency and control, further enhancing the stress response and maintaining performance during life-threatening events.

This research describes several conceptual approaches for a system to train individuals for stressful operations, and presents recommendations based on the development and testing of an adaptive stress training system with virtual reality (VR) in the spaceflight domain. The approach described here proposes a training system in conjunction with an adaptive process that can alter the stressors impacting the individual, where the changes in stressor levels are based on the individual's stress response. Physiological sensors and machine learning can detect and classify the level of user stress. Sensing the human state in real-time can enable the system to adapt a VR training environment based on the user's stress tolerance and the momentary context of the task procedure. An evaluation of the adaptive system components and overall effectiveness using a simulated fire emergency aboard a VR International Space Station (VR-ISS) is presented as an example for training stressful operations.

Professionals may someday use adaptive automation in systems to train for emergencies in high-stress occupations such as astronauts, military, police officers, firefighters, surgeons, and aircraft pilots. This work helps identify how adaptive systems can be designed to supplement current task training.

BACKGROUND

Stress

Stress occurs when an environment imposes significant demands on an individual (Lazarus & Folkman, 1984). When an individual perceives a stressor, they initiate the stress appraisal process which balances the perceived magnitude of the stressor in comparison to the individual's ability to meet the demand. A stress appraisal results in a stress response that causes various physiological, emotional, cognitive, social, and/or behavioral effects (Salas et al., 1996). Physiological responses often include inhibiting the parasympathetic nervous system, which is responsible for calming the body, while simultaneously activating the sympathetic nervous system, which increases heart rate, respiratory rate, blood pressure, muscle tension, diversion of blood flow from the internal organs to the brain and muscles, perspiration, and pupil dilation (Giannakakis et al., 2019). Due to the significant physiological response, stress can play a large role in an individual’s ability to accomplish a task.

How Stress Decreases Performance

The presence of a stressor, defined as any perceived threatening stimuli, such as workload, noise, time pressure, or task ambiguity, often result in emotional states such as anxiety, fear, or frustration (Driskell & Johnston, 1998). The induced emotional states limit the attentional capacity by preferentially focusing the individual mind on the stressor. This in-turn limits the amount of cognitive bandwidth that can be used to solve novel tasks under strenuous conditions, and ultimately reduces performance (Eysenck et al., 2007). The key to regaining an individual’s unstressed performance in the presence of a stressor is to adjust the individual's stress appraisal process. Changing the appraisal of a stressor from a threat to a challenge (or eventually a benign stimuli) is the beginning of forming resilience to a stressor (Kalisch et al., 2015).

Methods to Detect and Monitor Stress

Stress detection continues to evolve and is motivated by the potential utility of continuously monitoring stress levels in real-time (Smets et al., 2019; Reimer et al, 2017). Stress detection often involves a comparison of an individual’s emotional/physiological/vocal indicators to those same stressed and unstressed indicators from an individual or group. This may require controlled laboratory experiments to develop a comparative benchmark for stress or no-stress conditions, commonly known as ground truth. To compare data of interest to ground truth, the most common means of comparison is through machine learning classification or regression. This type of classification is generally known as supervised machine learning, which establishes a model with a training dataset composed of ground truth data before it can be used to classify stress on a testing dataset. Stress detection systems have been developed for drivers in semi-urban scenarios (Singh, et al., 2014), patients undergoing virtual reality therapy (Tartarisco et al., 2015), individuals in working environments (Betti et al., 2018), and people that need help managing daily stress (Hovsepian et al., 2015; Reimer et al., 2017; Plarre et al., 2011).

Detection systems collect information about stress responses from either objective physiological sensors or subjective psychological metrics, preprocess the signals, and extract and select specific/relevant information, which they then use to classify the stress levels or states. Data collection can be accomplished with physiological sensors that include electrodermal activity (EDA), electrocardiogram (ECG), electroencephalogram (EEG), respiration (RSP), skin temperature (ST), and blood volume pulse (BVP; Giannakakis et al., 2019). Instead of using direct biosignals, many stress detection systems derive features from the times-series signals that may characterize stress patterns or behavior, including features that measure distribution of data points, variation, correlation properties, stationarity, entropy, and nonlinear properties (Finseth et al., 2021b; Fulcher, 2018). For example, stress indices have been primarily inferred from changes in the time intervals between ECG heartbeats, which measure Heart Rate Variability (HRV) using a time-domain, frequency-domain, or non-linear analysis (Shaffer & Ginsberg, 2017). These features are then used as predictor variables within a machine learning model to establish the conditional probability that the data of interest belongs to a class (e.g., stress or no-stress). To discriminate and classify the features, different classification algorithms can be implemented depending on the data structure, such as support vector machine, Bayesian network, decisions tree and random forest, neural net, and K-nearest neighbor. Stress classification has generally been designed to discriminate two-levels of stress (stress vs. no-stress), although research with more levels is becoming more common (Smets et al., 2019). Despite the advancements in detection methods and technology, the translation of post-hoc, offline stress detection to real-time detection remains challenging with more research needed to improve reliability and robustness under changing physiological and environmental conditions.

Adaptive Systems

Adaptive automation is the process of dynamically changing authority or agency of system functions between a human and/or computer with the purpose of optimizing system (human-machine) performance (Feigh et al., 2012; Kaber et al., 2005). Automation refers to the execution of a function previously carried out by a human, whereas adaptive automation refers to dynamic change between different levels of autonomy (Byrne & Parasuraman, 1996). Therefore, an adaptive system is the technology component of a human-machine system that uses adaptive automation. Adaptive systems have primarily been characterized as adaptations of function allocation where task load is distributed between the human and system dynamically during task execution to maximize task performance, often without explicit input from the user (Parasuraman et al., 2000). The human-automation considers the agents’ capability to assume authority for a function and individual responsibility to perform the collective set of functions (Pritchett et al., 2014). Beyond allocating functions to distribute workload, adaptive systems can also modify what, how, and when information is presented to the user (Feigh et al., 2012). Such adaptations can result in changes to interface display features, training scenario difficulty, simulated environments, and control of objects/vehicles (Zahabi & Razak, 2020).

An adaptive system has three main functions: (1) assess the context or performance of multiple inputs, (2) use logic in the form of rules and triggers to decide how and when to adapt the system, and (3) execute adaptive changes to the automation or human-machine interface (Feigh et al., 2012). Context and performance are assessments of the operator or operator state, system state or modes of system operations, environment or events external to operator, completion of task goals, or time and location. Adaptations are adjustable features that are changed to influence the behavior of the system and human. Many types of adaptations can be used to adapt the system’s behavior or interface, including temporally off-loading tasks for the user, changing the information shown to the user, changing the timing or priority of tasks, or changing how the users exchange information with the system (Feigh et al., 2012). Adaptation rules and triggers tell the system when to change by engaging or disengaging adaptations, and how long adaptations should persist (Feigh et al., 2012). Triggers typically rely on information from the context assessment. The overall system performance depends on the individual performance of each of the three functions.

Adaptations can be used for purposes other than distributing task load through function allocation, including the ability to help humans learn and control affective, physiological, or cognitive skills. Some adaptive systems can monitor the human state from physiological signals, including metrics such as workload, fatigue, and stress. Specifically, much research has been devoted to the monitoring of stress detection in real-time (Smets et al., 2019). By classifying stress into different states, the system can adapt functionality given specific conditions; moreover, the system can decide when to engage with users and vary the amount of interaction and authority for a task. Therefore, the integration of stress detection with an adaptive system may be an effective way to build resilience and prevent stress.

REVIEW OF PREVIOUS WORK FOR REDUCING STRESS WITH ADAPTIVE SYSTEMS

A number of researchers (Serino et al., 2014; Pallavicini et al., 2016) have recognized the benefits of using adaptive systems to modify VR environments to conduct stress interventions. Further, some researchers have proposed system designs or tested components for a VR adaptive system that integrates stress prevention or stress management strategies (Zahabi & Razak, 2020).

The VR Adaptive Simulation (VRAS) was developed to prevent serious mental health problems prior to military deployment (Ćosić et al., 2010a; Ćosić et al., 2010b; Ćosić et al., 2011). The VRAS system gradually exposes trainees to stressful stimuli with simultaneous practice of stress coping skills. Stimuli are generated through static pictures, sounds, and real-life video clips. The subjective emotional response of the trainees informs the adaptations and is measured through voice, speech, facial expressions, heart rate, skin conductance, and respiration. This research was proof of concept and only tested the component that estimates the emotional states of the user.

Another adaptive system for military application is the VR for Cognitive Performance and Adaptive Treatment (VRCPAT 2.0; Parsons & Reinebold, 2012). The purpose of the system is to test a soldier's readiness to return to duty following a traumatic experience or head injury. The user then engages in a head-mounted display VR simulation of driving a military Humvee. Psychophysiological signals and task performance use several adaptations to change the VR environment, including the lead Humvee's speed, explosive blasts, insurgent artificial intelligent (AI) characters, in-game weather conditions, and a scent machine that produces gunpowder odor. However, this research is proof of concept and classification accuracy and experimental results have yet to be published.

A system proposed by Jones and Dechmerowski (2016) leverages mobile technology to objectively measure stress levels during augmented reality (AR) and VR training. The system proposes to use unobtrusive physiological stress monitoring, including cardiovascular and respiratory measures and electrodermal activity. The system uses a linear stochastic gradient descent binomial classifier that differentiates stress during a public speaking task at 95% accuracy (Winslow et al., 2015). They propose that a closed-loop system can create an ever-changing environment, thus pushing users to create coping mechanisms for use in a real-life environment. However, similar to the other training systems mentioned previously, this research is proof of concept and experimental results have yet to be published.

CONCEPTUAL SYSTEM ARCHITECTURE FOR ADAPTIVE STRESS TRAINING

The conceptual aim of a training is to prevent stress and related psychological outcomes during task execution, which include workload, fatigue, and anxiety. Therefore, two principles help guide training system development: (1) allow individuals to become gradually more familiar with stressors and stressor intensities that simulate the real-world task environment, and (2) help the user develop strategies to increase regulatory control over their own stress response.

To address the first principle of gradually increasing familiarity to stressors, stress training frameworks such as Stress Inoculation Training (SIT; Meichenbaum, 2007) and Stress Exposure Training (SET; Driskell et al., 2008) provide insight into how an adaptive system can introduce stressors. The high demand of the stress environment can cause individuals to develop positive or negative expectations regarding their capacity to perform. Negative expectations about the ability to meet the stressor demands can degrade performance and lead to negative emotions such as despair or learned helplessness. Research indicates the stress training is more effective when a trainee experiences success or a sense of task mastery when training (Driskell & Johnston, 1998). Therefore, administering training requires careful attention to the individuals stress levels and judgment about the appropriate increases to stressor levels.

An adaptive system can appropriately change stress levels through the detection of physiological signals as an indication of the user's stress state, as well as the other system and performance assessments, to autonomously adapt the training environment. Adapting the training environment for the individual centers around the concept of adapting stressors to keep users within a "zone" where they are neither overwhelmed nor under stimulated (Driskell et al., 2008). To adapt the training environment, the system can use implicit feedback in which the user observes feedback in the form of subtle changes to some details(s)/features(s) of the training environment (e.g., task information, visual or auditory stressors; Gaume et al., 2016). The changes are a direct correlate of the user's physiological stress state, but the biosignals are not explicitly shown to the user. For example, physiological arousal levels can be used to manipulate road visibility during a driving game or changing the field of view for police engaging hostiles in a training simulator (Parnandi & Guiterrez-Osuna, 2015; Brammer et al., 2021). As an individual progresses in training, the training environment can continue to increase the implicit changes until the individual achieves mastery in conditions that replicate that of the real-world operational environment.

To help the user develop strategies to increase regulatory control over their own stress response, regulatory control can be developed either through autonomic regulation of subcognitive systems or through volitional control by initiating coping strategies while the system is in operation (Gaume et al., 2016). Implicit feedback is theorized to rely more on autonomic regulation rather than volitional control (i.e., executive function and conscious cognitive strategies), consequently using less cognitive resources and attentional processes (Gaume et al., 2016). This might be favorable for training high-demand tasks that require sustained attention. However, the development of volitional control may be more effective than autonomic regulation, as healthier and more reliable coping strategies can be intentionally implemented during training while preventing maladaptive coping strategies from developing (e.g., avoidance, rumination). Prior to system operation, coping skills can be taught to the trainee for the purpose of building volitional control and practicing them later with the system (see Driskell et al., 2008). Since stress can interfere with skill acquisition and memory, the trainee should be proficient in the task and skill before using an adaptive system to train with stressors (Keinan & Friedland, 1996).

Component Descriptions

The general structure of an adaptive system for training is illustrated in Figure 1. The system relies on four components: context assessment, adaptation manager, training module, and the human-machine interface.

Context Assessment – The context assessment component is responsible for collecting data from various sources including the system, task, and human state. For example, collecting heart rate to quantify the user's cardiovascular activity. Further, assessments can be combined and classified into quantitative performance metrics. For example, the continuous streaming of physiological data enables real-time classification of stress levels (Smets et al., 2018). The reliability and accuracy of signal processing and classification are primary determinants of the overall effectiveness of adaptive systems triggered by a user's physiologically state (Erdogmus et al., 2005).

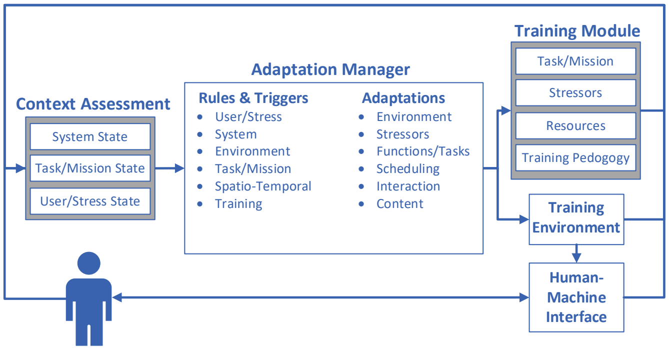


Figure 1. Framework for a generic adaptive stress training system

Adaptation Manager. The adaptation manager adapts the training environment to balance the joint human-automation abilities and is responsible for the rules and triggers as well as the level and type of adaptations. Based upon state information provided by the context assessment, rules and triggering mechanism determine how to change the systems behavior. The triggers can be related to an assessment of the user or their stress level, the modes of operation of the system, the training environment, the progress through a task/mission/procedure, time and location, or the predefined training pedagogy. For example, a system designed to optimize learning flow (Nakamura & Csikszentmihalyi, 2014) could use operator triggers to engage/disengage adaptations during training; when user anxiety is high, adaptations could engage to help reduce task difficulty. Whereas when user anxiety is low, adaptations could disengage (or others engage) to increase challenges and avoid boredom.

In the context of training for stressful operations, adaptations should be centered on helping the user build a resilience to environment stressors while maintaining performance. The adaptations can be related to modifications of the training environment, stressors, function/task allocation, the timing and duration of tasks, how the user interacts with the system (e.g., interface layout, how to access information), or what information is given to the user and amount of detail. For example, to modify the environmental stimuli, the system can adapt the intensity of stressors (unpredictable or uncontrollable threats, such as explosions, gunfire) or the amount of resources (e.g., malfunctioning equipment). Whereas, to modify task load, adaptations could include dynamically changing whether the user or the system (or AI teammate) performs tasks/subtasks within a procedure. If the user's stress level suggests compromised performance or the need for further stress inoculation, the adaptation manager can change the environmental stressor accordingly.

Training Module. The training module controls the normal operation of the training system and the training protocol both within a single training session and over multiple sessions. The module can contain different mission/procedure scenarios as well as the different levels or agency and authority in the different scenarios. These parameters are set before the training sessions begin, but can be modified by adaptations. For example, the training module may have the user experiences sessions that gradually increase in task function, details, or environmental stressors. However, adaptations during a session may intervene if user workload/stress is too high, which may command the training module to branch into a new scenario where the AI team members are granted more agency to complete mission objectives. Further, performance results from the session may trigger adaptations that change the scenario in the following user training session.

Training Environment and Human-Machine Interface (HMI). Virtual reality, augmented reality, mixed reality, or tactile simulations can be used as a training environment to help individuals acquire knowledge, skill, and abilities in highly realistic situations. The user interacts with the training environment through the HMI (e.g., head-mounted display, simulator console), although some applications may only have an HMI and not a training environment. Both the training environment and HMI can be adaptive with their set of triggers and rules that can be independently autonomously executed.

SYSTEM IMPLMENTATION

A VR stress training system was designed and evaluated in a series of experiments, each with a focus on testing the individual components as well as the overall system's effectiveness (Finseth et al., 2018; Finseth et al., 2020, Finseth et al., 2021a, Finseth et al., 2021b). The lessons learned from each experiment have been listed, along with recommendations for researchers wanting to implement similar training systems and an example from a spaceflight emergency fire procedure training.

An example of the adaptive system architecture for training with the stress of a spaceflight emergency is illustrated in Figure 2 (Finseth et al., 2021a). A VR-ISS simulation was developed based on NASA spaceflight procedures (e.g., fire, depressurization, toxic contaminants; Finseth et al., 2020). The user interacted with the environment with a head-mounted display and positioned-tracked controllers, with simulated locomotion that mocks microgravity. Head and controller tracking allowed the user to be fully immersed in the simulation while maintaining precision during task execution. Further, task-relevant equipment was modeled within the simulation for the user to practice locating oxygen masks, measuring contaminant levels, and operating a fire extinguisher.

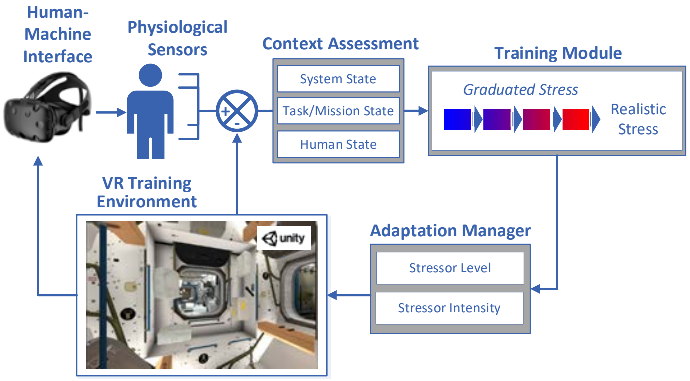


Figure 2. Example of an adaptive system architecture to adapt stressor level while training individuals for a spaceflight emergency procedure (Finseth et al., 2021a)

In this system, the goal of the context assessment was to produce a moment-to-moment stress classification. Data from physiological sensors were continuously streamed and used to derive hundreds of features to characterize stress across time windows (e.g., mean, variance, linearity, stationarity, entropy, frequency features). A small subset of features that were most discriminant of the individual's stress profile was selected. Features that showed class discrimination were extracted from the sensor signals and then used to train a machine learning classifier to detect multiple levels of stress (Finseth et al., 2021b).

The training module determined the target level of stress by following the SIT/SET protocol to introduce stressors over multiple user sessions (Finseth et al., 2018). Three levels of stressors were induced through a combination of environment changes (visibility, noise, and lights) to create low, medium, and high stressor levels (Figure 3).

The adaptation manager relied on a series of rules and triggering mechanisms. Changes in the user's stress triggered changes to environment content and stressor intensity through implicit feedback (Gaume et. al., 2016). Adaptations were designed to trigger only after the first minute and at 30-second increments thereafter. This was intended to allow adequate time for physiological responses to stabilize. Further, adaptations were not triggered until the stress had a conditional probability greater than 70% in two consecutive windows (60 sec). The stressor level was not allowed to exceed the SET level set for each of the five training sessions (low, medium, medium, high, high) to allow individuals to gradually be introduced to the stressors (Finseth et al., 2021a).

DEVELOPMENT RECOMMENDATIONS

1. Scenario Design: Task Relevant Manipulations

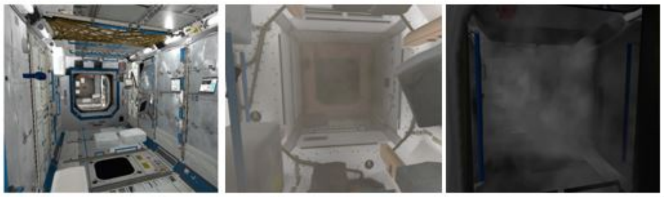
It is often unclear which stressors should be selected for the training simulations and which have the largest effects on users doing tasks in real-life situations. Some stressors are specific to the task (e.g., smoke during emergency fire training) and are therefore harder to identify as a potential influence on the individual's response. Relevant stressors identified by simulation designers may be different from those identified by operators who have experienced the real-life situation. To increase training transfer to real-life, stressors should be identified through subject matter experts, operators, interviews, and documented case studies. Although, not all stressors have the same effects on the procedure and some stressors may change the task procedure steps, which may be a useful consideration in training that aims to hold the skill criterion constant across individuals.

For example, a workshop was held with a panel of subject matter experts to identify stressors that contribute to the stress in a spaceflight environment to inform how VR spaceflight emergency fire training simulations could be designed. 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. Stressors were categorized by their ability to change the environment without substantially changing how the individual would perform the procedure. Selecting environmental stressors may help isolate effects from the training intervention, which may become confounded with differing task loads (Finseth et al., 2020). Three environmental stressors were selected for training an emergency fire procedure, where their intensity was varied among the training scenarios: alarms, flickering lights, and visibility from smoke (Figure 3).

2. Training Stressor Levels: Manipulation Testing

Multiple stressors may be involved in the operational environment. When designing training simulations, researchers need to decide the magnitude of each environmental stressors and the combination of different stressors (e.g., noise, smoke). To complicate the matter, physical and psychosocial stressors can have cumulative effects that are greater than the individual effects of the stressors alone (Abdelall et al., 2020). Therefore, stressor levels that will be used to manipulate the training simulations need to be measured to ensure they induce different stress levels (psychological and physiological stress) in the trainee. This process also determines the ground truth in which the data can be labeled as different stress levels and used as a reference for evaluating the accuracy of classification (Giannakakis, et al., 2019).

For example, a combination of different stressor intensities were used to create three distinct stressor levels (Figure 3). In order to verify that the scenarios are different before testing system effectiveness at detecting stress, the stressor levels for a simulated fire procedure were evaluated to ensure discrimination (Figure 4).



Stressor level	Smoke (visibility)	Alarm (noise)	Power Outage (lights)
LOW	None	None	None
MEDIUM	Light	Caution alarm	Node 2
HIGH	Heavy	Caution, warning, and fire alarm	All Modules

Figure 3. Stressor levels of a VR emergency procedure (Finseth et al., 2020)

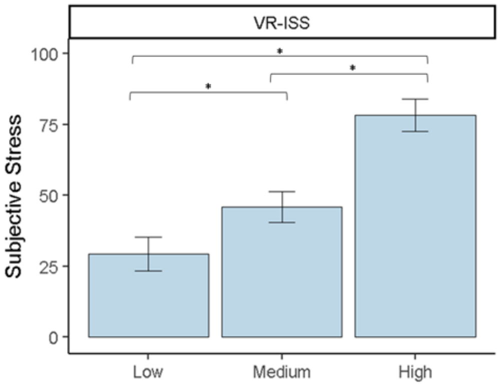


Figure 4. Subjective stress for stressor levels (Finseth et al., 2020)

Designing the Stress Detection

3. Personalized Versus Generalized Stress Detection Models

For stress detection, generalized classification refers to a model that is trained using predefined features (e.g., heart rate) from many individuals, whereas personalized classification refers to a model that is trained using only one individual's data (usually with features that are discriminative of that individual's stress response). There is a consistent pattern that individuals differ greatly in their physiological predictors and the physiological correlation to psychological states of stress. This may explain why personalized (i.e., individual-specific) models tend to outperform generalized (i.e., between-person) stress classification models (Osotsi et al., 2020). Although, personalized classifier models may be a more robust stress detection than generalized systems, generalized systems can start detecting stress immediately for a new user. In contrast, personalized systems require time and effort to train the model before it can start detecting stress. To avoid these limitations, some systems attempt personalization by calibrating the model intermittently using subjective feedback from the user (Plarre et al., 2011; Hovsepian et al., 2015; Reimer et al., 2017). Reinforcement learning, as an alternative to supervised machine learning, may also offer the potential to update the classification model (Zahabi & Razak, 2020). Researcher should consider using personalized and calibrated models as these methods are likely to continue to grow in popularity to account for changing personal physiological responses (Smets et al., 2019).

4. Stressors Used to Train the Classification Model

Supervised machine learning classification will need to establish a model with a training dataset before it can be used to classify stress on a testing dataset. The difficulty of stress detection is that physiological systems within the body may have differing responses for different types of stressors. For example, using social stress from a mock interview will elicit different physiological reactions than noise stimuli (Pedrotti et al., 2014). To reduce classification error, the training dataset should be collected from stressors that closely mirror the stressors encountered when the individual begins training scenarios. This will improve the similarity of the physiological features in both datasets, leading to more accurate stress detection.

For example, consider an experiment where participants conducted either a mental workload N-back task or a VR-ISS spaceflight emergency task during trials with three stressor levels: low, medium, and high (Finseth et al., 2021b). The physiological data collected during the trials were then used to extract and select features that were the most discriminative of the individual's stress response. The selected feature data was then used to train the supervised classification model to detect stress. The features selected for N-back were different than the VR-ISS. The N-back commonly selected on heart rate skew, heart rate partial autocorrelation (one lag). Whereas, the VR-ISS commonly selected systolic blood pressure, time-domain components of HRV, and power spectral density coefficient features. These features have different characterizations on the autonomic nervous system. This suggest that trying to use another task (e.g., N-back) for creating a stress detection model may result in the selection of features that do not represent the intended training environment (VR-ISS), and thus lower classification performance.

Adaptation Testing

5. Frequency of Adaptations

Frequent adaptations can be volatile and not allow enough time for the user to adjust (Fuchs et al., 2008). However, how much time should be placed between adaptations is dependent on the type of task. For stress training, adjusting training quickly may be better for participant needs, but the physiological system needs time to respond to the changing stress levels. Longer durations between adaptations can allow users to transition between stress levels. Researchers should balance adaptation frequency with training objectives. Further, having consistency in the timing of adaptations creates more transparency and allows users to interpret how the automation is behaving and a prediction of how it will behave in the future.

6. Safeguards for Adaptation Rules during Unforeseeable System Errors

Adaptation rules and triggers may perform as intended, however other system components may have unforeseeable issues. If the adaptation rules have too much flexibility, system wide errors caused by signal noise and incomplete information could result in adaptations that do not match the user's current state, thus leading to negative training effects. Systems should be developed with safeguard constraints that protect against propagating errors through adaptations. This can happen in situations where a physiological sensor becomes disconnected and leads to an erroneous stress prediction that would correctly trigger adaptations, but still result in the training overwhelming the user with stress.

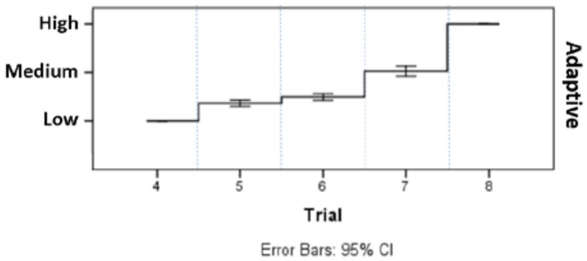


Figure 5. Stressor level adaptations made during trials by the adaptive system (Finseth et al., 2021a)

For example, the adaptation rules for the spaceflight emergency adaptive system acted as a safeguard and resulted in stress inoculation (Finseth et al., 2021a). The rules were constrained to only allow a maximum adaptation to be a predetermined stressor level that increased over multiple training trials (i.e., low, medium, medium, high, high). These rules allowed the trainees to move at their own pace, to increase the stressor level when they gained a specified proficiency, and increased the reliability that the system behaves as intended by minimize the risk from adapting in the wrong conditions. The experiment allowed adaptations in trials 5-7, with results showing that most participants adapted at a rate of once per trial, with some participants having two adaptations or no adaptations due to different rates of inoculation to stress (Figure 5).

7. Evaluation of Overall System Effectiveness

The utility of an adaptive training system to improve an individual's response to future stressful operations relies on the comparison of its effectiveness against other training methodologies, specifically those that do not use adaptations (Zahabi & Razak, 2020). Therefore, researchers should account for training effects that may be confounding the resulting stress response if rules and triggering do not work as intended within the adaptive system, and help to understand what types of adaptations result in the most improvement.

For example, repetitive training to learn new skills, such as training with constant low-levels of stress (or no stress) may reinforce skill development but lead to performance problems in a high stress environment. Whereas gradually increasing stress may prepare for those stressors and is promoted within stress training programs (SIT and SET), but may not adequately adjust for individual's needs (Friedland & Kenian, 1996). An experiment was conducted to compare the stress response from before and after receiving training between a constant low-levels group, a graduated group, and an adaptive system group. The adaptive group followed the same stressor level pedagogy as the graduated group (e.g., low, medium, medium, high, high), but was allowed to adapt the level based on user stress. Results showed that all training conditions lowered stress, but the preponderance of effects for the adaptive group suggest it is more successful in decreasing stress over multiple trials, with improved heart rate, HRV, and task performance (Finseth et al., 2021a). By using rules based on the graduated group stressor level, it is then possible to infer that the adaptations had a positive and sizable effect on the stress response of the individuals.

FUTURE DIRECTIONS

While adaptive systems continue to show promise for training applications, more research is needed. On the topic of stress detection, there is great difficulty in validating the classification accuracy of deployed models due to temporal effects (e.g., fatigue, stress reappraisal, physiological habituation). Testing datasets collected by deployed stress detection systems often show large differences over time compared to training data used to build models in the lab. Therefore, it is hard to verify detections against a ground truth because it violates machine learning assumptions of independent and identical distributions. Further, more robust stress detection is needed to handle unforeseen events such as signal artifacts or sensor disconnects, which could cause false stress classifications and alter task training such that individuals are negatively impacted. On the topic of adaptive systems to train for stressful operations, more evaluations are needed on the training effect of developing coping skills. Within the SIT and SET framework, coping skills are taught to users alongside task procedures before adaptive system training. More research is needed for the coping skill types, retention, duration of training (session length, or spacing), and how the combined coping skill training with adaptive system affects resilience development. A mobile training system utilizing AR or VR may minimize setup and implementation difficulty and require minimal trainer/psychologist support for developing resilience.

CONCLUSION

Training is an essential countermeasure tool to mitigate safety risks to vehicles and operators, as well as increase the probability of mission success. This paper presented research on the development and testing of an adaptive system for stressful operation training and lessons from the task training, system development, and experimental methods. The recommendations may be useful for future researchers looking to develop a training system or generate research questions.

Adaptive systems have a number of practical applications for occupations that encounter stressful or hazardous tasks. Real-time physiological stress detection shows the capacity to account for individual differences in stress responses, which may help align training to match the user's current state and increase the effectiveness of training in the operational environment. Adaptive training can consistently adapt to keep user in an optimal zone that challenges the user without under-whelming or over-whelming them. Future development of this system may someday supplement current task training by helping individuals develop the critical proficiencies to novel stressful conditions.

ACKNOWLEDGEMENTS

This work was funded by the National Aeronautics and Space Administration [grant number 80NSSC18K1572].

REFERENCES

Abdelall, E. S., Eagle, Z., Finseth, T., Mumani, A. A., Wang, Z., Dorneich, M. C., & Stone, R. T. (2020). The interaction between physical and psychosocial stressors. *Frontiers in behavioral neuroscience, 14*.

Barshi, I., & Dempsey, D. L. National Aeronautics and Space Administration (NASA). (2016). *Risk of Performance Errors Due to Training Deficiencies: Evidence Report.* (No. JSC-CN-35755). Houston, TX: NASA Johnson Space Center.

Betti, S., Lova, R.M., Rovini, E., Acerbi, G., Santarelli, L., Cabiati, M., Del Ry, S. & Cavallo, F. (2017). Evaluation of an integrated system of wearable physiological sensors for stress monitoring in working environments by using biological markers. *IEEE Transactions on Biomedical Engineering, 65*(8), 1748-1758.

Brammer, J. C., Van Peer, J. M., Michela, A., van Rooij, M. M., Oostenveld, R., Klumpers, F., ... & Roelofs, K. (2021). Breathing biofeedback for police officers in a stressful virtual environment: challenges and opportunities. *Frontiers in psychology, 12,* 401.

Byrne, E. A., & Parasuraman, R. (1996). Psychophysiology and adaptive automation. *Biological psychology, 42*(3), 249-268.

Campbell, J., & Ehlert, U. (2012). Acute psychosocial stress: does the emotional stress response correspond with physiological responses?. *Psychoneuroendocrinology, 37*(8), 1111-1134.

Ćosić, K., Popović, S., Horvat, M., Kukolja, D., Dropuljić, B., Kostović, I., Judaš, M., Radoš, M., Radoš, M., Vasung, L., & Spajić, B.B. (2011). Virtual reality adaptive stimulation in stress resistance training. Proceedings RTO-MP-HFM-205 on" Mental Health and Well-Being across the Military Spectrum.

Ćosić, K., Popović, S., Kostovic, I., & Judas, M. (2010a). Virtual reality adaptive stimulation of limbic networks in the mental readiness training. *Studies in health technology and informatics*, 154, 14-19.

Ćosić, K., Popović, S., Kukolja, D., Horvat, M., & Dropuljić, B. (2010b). Physiology-driven adaptive virtual reality stimulation for prevention and treatment of stress related disorders. *CyberPsychology, Behavior, and Social Networking, 13*(1), 73-78.

Delahaij, R., Gaillard, A.W.K., & Soeters, J.M.L.M. (2006). Stress training and the new military environment. In *Human Dimensions in Military Operations – Military Leaders' Strategies for Addressing Stress and Psychological Support* (pp. 17A-1 – 17A-10). Meeting Proceedings RTO-MP-HFM-134, Paper 17A. Neuilly-sur-Seine, France: RTO.

Driskell, J. E., & Johnston, J. H. (1998). Stress exposure training. *Making decisions under stress: Implications for individual and team training,* 191-217.

Driskell, J. E., Salas, E., Johnston, J. H., & Wollert, T. N. (2008). Stress exposure training: An event-based approach. *Performance under stress,* 271–286. London: Ashgate

Erdogmus, D., Adami, A., Pavel, M., Lan, T., Mathan, S., Whitlow, S., & Dorneich, M. (2005, March). Cognitive state estimation based on EEG for augmented cognition. In *Conference Proceedings. 2nd International IEEE EMBS Conference on Neural Engineering, 2005*. (pp. 566-569). IEEE.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. *Emotion, 7*(2), 336.

Feigh, K. M., Dorneich, M. C., & Hayes, C. C. (2012). Toward a characterization of adaptive systems: a framework for researchers and system designers. *Human Factors, 54*(6), 1008-1024.

Finseth, T., Dorneich, M. C., Keren, N., Franke, W. D., & Vardeman, S. (2020). Designing Training Scenarios for Stressful Spaceflight Emergency Procedures. In *2020 AIAA/IEEE 39th Digital Avionics Systems Conference (DASC)* (pp. 1-10). IEEE.

Finseth T., Dorneich, M.C., Keren, N., Franke, W., Vardeman, S., Segal, J., Deick, A., Cavanah, E., & Thompson, K. (2021a). *The Effectiveness of Adaptive Training for Stress Inoculation in a Simulated Astronaut Task.* In *Proceedings of 2021 Human Factors and Ergonomics Society (HFES) Annual Meeting.* Baltimore, MD.Finseth, T., Dorneich, M. C., Vardeman, S., Keren, N., & Franke, W. D. (2021b). *Physiologically Based Stress Detection For Hazardous Operations Using Approximate Bayes Algorithm.* [Manuscript submitted for publication]. Department of Aerospace Engineering, Iowa State University.

Finseth, T. T., Keren, N., Dorneich, M. C., Franke, W. D., Anderson, C. C., & Shelley, M. C. (2018). Evaluating the Effectiveness of Graduated Stress Exposure in Virtual Spaceflight Hazard Training. *Journal of Cognitive Engineering and Decision Making*, *12*(4), 248-268.

Fuchs, S., Hale, K. S., Berka, C., & Juhnke, J. (2008). A mitigation framework for enhancing situation awareness. *Augmented Cognition: A Practitioner's Guide*, 112-143.

Fulcher, B. D. (2018). Feature-based time-series analysis. In *Feature engineering for machine learning and data analytics* (pp. 87-116). CRC Press.

Gaume, A., Vialatte, A., Mora-Sánchez, A., Ramdani, C., & Vialatte, F. (2016). A psychoengineering paradigm for the neurocognitive mechanisms of biofeedback and neurofeedback. *Neuroscience & Biobehavioral Reviews, 68*, 891-910.

Giannakakis, G., Grigoriadis, D., Giannakaki, K., Simantiraki, O., Roniotis, A., & Tsiknakis, M. (2019). Review on psychological stress detection using biosignals. *IEEE Transactions on Affective Computing*.

Hovsepian, K., Al'Absi, M., Ertin, E., Kamarck, T., Nakajima, M., & Kumar, S. (2015, September). cStress: towards a gold standard for continuous stress assessment in the mobile environment. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing* (pp. 493-504).

Jones, D., & Dechmerowski, S. (2016, July). Measuring Stress in an Augmented Training Environment: Approaches and Applications. In *International Conference on Augmented Cognition* (pp. 23-33). Springer, Cham.

Kaber, D. B., Wright, M. C., Prinzel III, L. J., & Clamann, M. P. (2005). Adaptive automation of human-machine system information-processing functions. *Human factors, 47*(4), 730-741.

Kalisch, R., Müller, M. B., & Tüscher, O. (2015). A conceptual framework for the neurobiological study of resilience. *The Behavioral and brain sciences, 38*.

Keinan, G., & Friedland, N. (1996). Training effective performance under stress: Queries, dilemmas, and possible solutions. *Stress and human performance*, 257–277. Mahwah, NJ: Lawrence Erlbaum.

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

Meichenbaum, D. (2007). Stress Inoculation training: a preventative and treatment approach. In R. W. P.M. Lehrer, *Principles and Practice of Stress Management*, 3rd Edition. Guilford Press.

Nakamura, J., & Csikszentmihalyi, M. (2014). The concept of flow. In *Flow and the foundations of positive psychology* (pp. 239-263). Springer, Dordrecht.

Orasanu, J.M., Kraft, N., Tada, Y. (2006). *Augmented Cognition: Past, Present, and Future,* 2nd edn.

Osotsi, A., Oravecz, Z., Li, Q., Smyth, J., & Brick, T. R. (2020). Individualized modeling to distinguish between high and low arousal states using physiological data. *Journal of Healthcare Informatics Research, 4*(1), 91-109.

Pallavicini, F., Argenton, L., Toniazzi, N., Aceti, L., & Mantovani, F. (2016). Virtual reality applications for stress management training in the military. *Aerospace medicine and human performance, 87*(12), 1021-1030.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans, 30*(3), 286-297.

Parnandi, A., & Gutierrez-Osuna, R. (2015). A comparative study of game mechanics and control laws for an adaptive physiological game. *Journal on Multimodal User Interfaces, 9*(1), 31–42. Springer.

Parsons, T. D., & Reinebold, J. L. (2012). Adaptive virtual environments for neuropsychological assessment in serious games. *IEEE Transactions on Consumer Electronics, 58*(2).

Pedrotti, M., Mirzaei, M. A., Tedesco, A., Chardonnet, J. R., Mérienne, F., Benedetto, S., & Baccino, T. (2014). Automatic stress classification with pupil diameter analysis. *International Journal of Human-Computer Interaction*, *30*(3), 220-236.

Plarre, K., Raij, A., Hossain, S.M., Ali, A.A., Nakajima, M., Al'Absi, M., Ertin, E., Kamarck, T., Kumar, S., Scott, M. and Siewiorek, D. (2011, April). 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* (pp. 97-108). IEEE.

Pritchett, A. R., Kim, S. Y., & Feigh, K. M. (2014). Measuring human-automation function allocation. *Journal of Cognitive Engineering and Decision Making*, *8*(1), 52-77.

Reimer, U., Laurenzi, E., Maier, E., & Ulmer, T. (2017, January). Mobile stress recognition and relaxation support with smartcoping: User-adaptive interpretation of physiological stress parameters. In *Proceedings of the 50th Hawaii International Conference on System Sciences*.

Robson, S., & Manacapilli, T. (2014). *Enhancing Performance Under Stress: Stress Inoculation Training for Battlefield Airmen.* Santa Monica, CA: The Rand Corporation.

Salas, E., Driskell, J. E., & Hughes, S. (1996). The study of stress and human performance. *Stress and human performance(A 97-27090 06-53), Mahwah, NJ, Lawrence Erlbaum Associates, Publishers*, *1996*, 1-45.

Serino, S., Triberti, S., Villani, D., Cipresso, P., Gaggioli, A., & Riva, G. (2014). Toward a validation of cyber-interventions for stress disorders based on stress inoculation training: a systematic review. *Virtual Reality, 18*(1), 73–87.

Shaffer, F., & Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Frontiers in public health*, *5*, 258.

Singh, R. R., Conjeti, S., & Banerjee, R. (2014). Assessment of driver stress from physiological signals collected under real-time semi-urban driving scenarios. *International Journal of Computational Intelligence Systems*, *7*(5), 909-923.

Smets, E., De Raedt, W., & Van Hoof, C. (2019). Into the wild: the challenges of physiological stress detection in laboratory and ambulatory settings*. IEEE journal of biomedical and health informatics*, *23*(2), 463-473.

Tartarisco, G., Carbonaro, N., Tonacci, A., Bernava, G.M., Arnao, A., Crifaci, G., Cipresso, P., Riva, G., Gaggioli, A., De Rossi, D., & Tognetti, A. (2015). *Neuro-fuzzy physiological computing to assess stress levels in virtual reality therapy. Interacting with Computers, 27*(5), 521-533.

Winslow, B., Carroll, M. B., Martin, J. W., Surpris, G., & Chadderdon, G. L. (2015). Identification of resilient individuals and those at risk for performance deficits under stress. *Frontiers in neuroscience, 9*, 328.

Zahabi, M., & Razak, A. M. A. (2020). Adaptive virtual reality-based training: a systematic literature review and framework. *Virtual Reality*, 1-28.