Adaptive Virtual Environments for Neuropsychological Assessment in Serious Games

Thomas D. Parsons and James L. Reinebold

Abstract — *While advances in military relevant simulations provide potential for increasing assessment of Soldier readiness to Return-to-Duty (e.g., following a blast injury), little has been done to develop these simulations into adaptive virtual environments (AVE) for improved neurocognitive and psychophysiological assessment. Adaptive assessments offer the potential for dynamically adapting the difficulty level specific to the patient's knowledge or ability. We present an iteration of the Virtual Reality for Cognitive Performance and Adaptive Treatment (VRCPAT) that proffers a framework for adapting scenarios in a game engine based upon the user's neurocognitive (task performance) and psychophysiological (e.g., heart rate, skin response, heart rate, and pupil diameter) states[1].*

Index Terms — Neuroscience, Neuropsychological Assessment, Psychophysiology, Adaptive Virtual Environment

I. INTRODUCTION

Military clinicians are increasingly being asked to make statements regarding a military service member's ability to return to active duty. Given the increasing prevalence of blast injuries to the head, and the fact that many brain injuries may have no external marker of injury, there is need for the serious games community to research innovative assessment methods. Current "Return-to-Duty" assessments are based upon the "Return-to-Play" guidelines found in Sports Medicine. Both have incorporated the Automated Neuropsychological Assessment Metrics (ANAM) to aid in decisions related to resuming activities following a concussion. Unfortunately, the ANAM was not developed with the intention of 1) adapting the difficulty level specific to the patient's ability; or 2) tapping into everyday behaviors like navigating a Middle Eastern city or gathering supplies for a mission.

Some promise has been found in serious gaming environments that aim to 1) incorporate adaptive assessments that dynamically adapt the difficulty level specific to the user's performance; and 2) increase the ecological validity of neurocognitive batteries through the use of simulation technologies for assessment and treatment

planning [1]. The success of virtual reality-based neuropsychological assessment and psychophysiology research paradigms may lead to a psychophysiological computing approach, in which such data gleaned from persons interacting within a military relevant simulation may be used to develop adaptive virtual environments (AVE) for training and rehabilitation [1].

At the heart of psychophysiological computing are the tools available to measure activity within the human user. Psychophysiological measures offer measures of human psychophysiological responses such as cardiorespiratory activity (e.g., heart rate (HR) and heart rate variability (HVR)), as well as measures of electrodermal activity (e.g., skin conductance and galvanic skin response (GSR)), respiration, and pupilometry. For a psychophysiological sensor to be useful it must be sensitive to an aspect of patient state that has cognitive or performance implications (i.e., arousal; cognitive workload). A key aspect of cognitive workload is the relation between the task and the user's abilities. Previous studies [2], [3] have shown that psychophysiological indices can be useful in adapting difficulty settings in games to provide a more entertaining play experience. Our framework leverages these advances to provide a more sophisticated performance assessment platform that allows the user's performance to be evaluated over a specific range of stress levels.

Research on stress and performance has revealed that arousal is closely related to a subject's performance in mental tasks. According to the well-known Yerkes-Dodson Law [4], performance is a nonmonotonic function of arousal. Performance increases with arousal when the arousal level is low, then reaches its peak at the optimal arousal level, and then decreases as arousal continues to increase. In an adaptive virtual environment that aims to improve the user's performance in mental tasks, such as learning and cognitive workload, it is very important to be able to identify the user's optimal arousal level and to recognize whether or not the user's actual arousal level is close to that optimal level.

In earlier work [5], we modified a game called Full Spectrum Warrior to develop a Virtual Reality Cognitive Performance Assessment test (VRCPAT 1.0). The primary goal of VRCPAT 1.0 has been to psychometrically validate it as a system that builds upon (not replaces) current neuropsychological assessments. The use of a military relevant VE offers enhanced flexibility to gather a greater degree of information relative to a user's neurocognitive processing: 1) information load via the intensity and

T.D. Parsons is with the University of Southern California, 12015 Waterfront Drive, Playa Vista, CA 90094-2536. (e-mail: tparsons@usc.edu).

J.L. Reinebold is with the Institute for Creative Technologies, University of Southern California, 12015 Waterfront Drive, Playa Vista, CA 90094-2536 (e-mail: reinebold@ict.usc.edu).

complexity of target stimuli; 2) temporal constraints during varied sustained assessment conditions; 3) distracting activities during the neurocognitive assessments; 4) sensory modality of the information presentation that needs to be attended to; 5) presentation of aversive stimuli for stressed performance evaluations; and 6) the development of a Normative Database whereby neuropsychological performance could be assessed under known conditions supported by normative standards. The current iteration of VRCPAT (Virtual Reality for Cognitive Performance and Adaptive Treatment; VRCPAT 2.0) uses a visual scenario editor and scripting language [6] to develop the scenarios: virtual Iraqi/Afghani scenarios (City, Checkpoint, Humvee). The platform was used due to its robust fidelity simulation, ease of modification, context relevant in game assets, and adoption by many military forces.

II. THE NORMATIVE DATABASE (VRCPAT 1.0)

We first developed VRCPAT 1.0, which includes a "normative database" drawn from over 400 subjects, including West Point cadets, college age civilians, and military service members recruited from large Army installations in the continental U.S [5], [7], [8]. Although there was some variation across studies, in general participants received part or all of the following: 1) VRCPAT battery; 2) 1.5-hour in-person neuropsychological assessment; and 3) ANAM.

A. Traditional Neuropsychological Measures

The neuropsychological battery typically included the following assessments: Attention: Digit Span (Forward and Backward) from the Wechsler Adult Intelligence Scale–Third edition (WAIS III); and the Paced Auditory Serial Addition Task (PASAT). Processing speed and visual-motor functioning: Digit Symbol Coding from the WAIS III; and Trail Making Test Part A. Executive function: Trail Making Test Part B; PASAT, and Color-Word Interference Test (i.e., Stroop) from the Delis-Kaplan Executive Function System (D-KEFS). General intellectual function: Wechsler Test of Adult Reading. Verbal learning and memory: Hopkins Verbal Learning Test – Revised (HVLT). Nonverbal learning and memory: Brief Visuospatial Memory Test – Revised (BVMT). Verbal fluency: we used letter and category fluency from the D-KEFS. The following ANAM subtests were given: TBI Questionnaire; Sleepiness Scale; Mood Scale; Simple Reaction Time; Code Substitution; Procedural Reaction Time; Math Processing; Matching to Sample; Go/No-Go; and Stroop.

B. Psychophysiological Measures

We also developed indices of valence and arousal using psychophysiological measures: Electrodermal activity (EDA) was assessed because greater levels of fear responding have been found to result in greater skin conductance response potentiation and increased skin conductance levels [9]. Skin conductance has also been found to increase as workload increases [10]. We looked at electrocardiographic activity (ECG) because during highly arousing and fearful situations ECG is generally associated with defensive responding, which results in increased

heart rate (HR), as opposed to orienting responses which reduce HR [11]. Increased arousal caused by fearful situations tends to result in HR increases. Increased HR is associated with increased cognitive workload [12], [13]. We also monitored respiration rate (RSP) because RSP has consistently been shown to increase in response to heightened levels of arousal associated with fear [14]. Further increase in respiratory rate has been associated with increased cognitive demand [12].

Convergent and discriminant validation of this module with a 1.5-hour neuropsychological assessment were conducted in four published studies (total sample number, N=137) and 3 unpublished studies (N=200) with a total sample of 337 healthy adults, between the ages of 21 and 36, that included men and women from ethnically diverse populations [5], [15], [16], [17]. Results across studies have consistently supported construct (both convergent and discriminant) validity. For example, in Parsons et. al. [5] convergent validity results indicated that the module correlated significantly with traditional neuropsychological 1) Learning Composite (HVLT 1-3; and BVMT 1-3; r=0.72, p<0.001), with 52% variance shared between the two indices; 2) Memory Composite (HVLT Total Recall; and BVMT Total Recall; r=0.68, p<0.001), with 46% variance shared between the two indices. Discriminant validity was evident in that no significant correlations were found between the module and tests of executive functions, attention, or processing speed.

C. Virtual Reality Stroop Task

We have also addressed construct validation using a Virtual Reality Stroop Task (VRST) [18]. Like the paper-and-pencil Stroop, the VRST assesses simple attention, gross reading speed, divided attentional abilities, and executive functioning. Like the ANAM, the VRST automates the paper-and-pencil Stroop task and allows for assessment of reaction time to single-item presentations of Stroop stimuli. The VRST goes beyond the ANAM and paper-and-pencil versions of the Stroop by assessing for the impact of training, stimulus intensity (arousal from threatening stimuli), and stimulus complexity (endogenous and exogenous presentations) in a virtual environment that replicates real world activities (see Figure 1).

Fig. 1. A Virtual Reality Stroop Task in which the subject is immersed in a virtual High-Mobility Multipurpose Wheeled Vehicle (HMMWV) as Stroop stimuli appear on the windshield.

The VRST involves the subject being immersed in a virtual High-Mobility Multipurpose Wheeled Vehicle (HMMWV) as Stroop stimuli appear on the windshield. The VRST is a measure of supervisory attentional processing (executive functioning) and was designed to emulate aspects of both the paper-and-pencil and ANAM versions of the Stroop test. Like the ANAM version, the VRST uses single-item presentation

of Stroop stimuli and requires an individual to press one of three computer keys to identify each of three colors, (i.e., red, green, or blue), however, the VRST adds a simulation environment with military relevant stimuli in high and low threat settings. As is the case with the D-KEFS, the VRST includes both simple and complex presentation of the interference stimuli.

III. The apparatus used for the virtual HMMWV included a desktop computer with a 3 GHz Processor; 6 GB of RAM; and a high performance graphic processing unit. Two monitors were used: 1) one for displaying the Launcher application which is used by the examiner administering the test; and 2) another for displaying the participant's view of the virtual environment in the Head Mounted Display (HMD). Participants wore a HMD with a multi-purpose sensor integrating nine discreet miniature inertial sensing elements with advanced Kalman filtering algorithms. A steering wheel was clamped on to the edge of a table in front of the monitors. Accelerator and brake pedals were positioned under the table. To increase the potential for sensory immersion, a tactile transducer was built using a three foot square platform with six 50W bass speakers attached. The tactile transducer was powered by an amplifier with 100 Watts per Channel x 2 in Stereo Mode.

Animation software was utilized for development of the virtual Iraqi/Afghani environment. The environments were rendered in real time using a graphics engine with a fully customizable rendering pipeline, including vertex and pixel shaders, shadows, bump maps, and screenspace geometric primitives [18]. The NeuroSim Interface (NSI) developed in the Neuroscience and Simulation Laboratory (NeuroSim) at the University of Southern California was employed for data acquisition, to guide stimulus presentation, and for psychophysiological monitoring. The NSI also allowed for key events in the environment to be logged and time stamped with millisecond temporal accuracy.

Psychophysiological responses were recorded throughout the participants' period of immersion within the VRST. Electriocardiogram (ECG), electrodermal activity (EDA), and respiration were the psychophysiological measures recorded as participants rode in the simulated HMMWV through alternating zones of low threat and high threat. Amongst low threat zones little activity in the virtual environment was presented aside from driving down a desert road. During the high threat zones, gunfire, explosions, and shouting amongst other stressors were presented throughout. The participants experienced four low threat and four high threat zones designed to manipulate levels of arousal. The order of threat level presentation was counterbalanced across participants. The VRST was completed during exposure to the high and low threat zones. The VRST consisted of four conditions: 1) word-reading, 2) color-naming, 3) simple interference (stimuli presented in the middle of windshield), and 4) complex interference (stimuli presented in variable locations on windshield). Each Stroop condition was experienced once in a high threat zone and once in a low threat zone. As with the ANAM Stroop test, the presentation speed of individual stimuli was user-defined, meaning that the subsequent stimulus did not appear until the appropriate key was pressed for the stimulus currently being viewed.

Our goal was to compare the VRST with paper-and-pencil (D-KEFS) and computer automated (ANAM) versions of the Stroop on behavioral measures. We also utilized psychophysiological measures to assess the effects of threat on arousal. Finally, we aimed to assess different aspects of supervisory attentional control within the VRST. The primary results were: (1) the typical Stroop effect pattern found in the D-KEFS and ANAM was replicated in the VRST; (2) the exogenous and endogenous attentional processing effects were similar in the complexity condition of the VRST and the D-KEFS Interference/Switching; (3) the two single-item presentation formats (i.e., VRST and ANAM) of the Stroop produced similar effects; 4) there was greater psychophysiological arousal in the VRST high threat zones than the low threat zones; and 5) while there was a learning effect present in the VRST across conditions, it was limited to the high threat zones. No negative side effects were associated with use of the VRST [18].

A. General Cognitive and Arousal Indices

Given the fact that blast injuries may cause a user's values to deviate from the norms found in the normative database, a user specific "General Cognitive Index" (GCI) and a "General Arousal Index" (GAI) were developed through comparison with baseline assessment (see "Profile Module" below). The GCI is the standardized score that represents how far the user deviates from the norm. Cognitive workload resource capacity (as measured by the GCI) essentially represents how much information can be maintained and manipulated in working memory. Individual differences in cognitive workload capacity are consistently found and these differences are strongly correlated with performance on a number of different cognitive tasks. In a similar manner, the GAI is the standardized score representing how far the user deviates from the norm for arousal in various VRCPAT scenarios. The GAI represents the user's psychophysiological profile relative to VE stimuli. Individual differences in arousal have been consistently found that strongly correlated with performance on a number of different cognitive tasks.

IV. THE ADAPTIVE FRAMEWORK (VRCPAT 2.0)

We have also developed an AVE: Virtual Reality for Cognitive Performance and Adaptive Treatment [1]. The VRCPAT 2.0 aims to progress the user from suboptimal cognitive and affective states toward an optimal state that enhances the user's performance. To do so requires the simulation to first infer user's current cognitive state based on psychophysiological signals and task performance and then create changes in the virtual environment to steer the user towards the desired state [19]. Like VRCPAT 1.0, this application also used the NSI for data acquisition, stimulus presentation, psychophysiological monitoring, and communication between the psychophysiological recording hardware and the virtual environment. Configuration parameters are saved to files using the NSI and automatically

loaded through its control module, allowing the experimenter to rapidly switch configurations in order to perform specific experimental sequences. The NSI also enables the sending of event markers from the stimulus presentation computer to the data recording device. Finally, the NSI compiles scripts to filter the incoming psychophysiological data in real-time. The VRCPAT 2.0 is composed of three modules (see Figure 2).
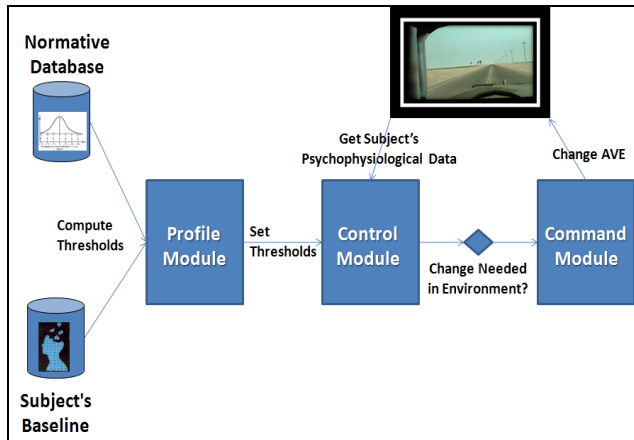

Fig. 2. The VRCPAT 2.0 Adaptive Framework. Users first establish set points during the profile module and then enter the adaptive simulation where the control and command modules guide the simulation based on the user's performance.

A. Profiling Module

The Profiling Module (PM) uses a baseline assessment that places the user in a simplified VE consisting of both high and low stress zones. Determination of whether the user's arousal level is above or below pre-established (normative database) thresholds is done using a comparison of features extracted from the following psychophysiological indices:

1) Cardiorespiratory Activity - Cardiovascular Activity is assessed in the Profiling Module using interbeat intervals (mean IBIs) that are calculated as the time difference in milliseconds between successive R waves in the ECG signal. Higher mean results indicate higher levels of arousal. To calculate the number of beats, we run a discrete Fourier transform and bandpass filter between 0.61 and 1.31 Hz. The dominant (highest magnitude) frequency detected within the band is the frequency of the onset of the R-waves. Next, the IBI is calculated based on the size of the window sampled. This allows for the profiling of the user's heart rate response during highly arousing and threatening situations that are found in the high threat zones of the HMMWV Stroop task. Such responses reflect defensive responding that results in increased heart rate. Contrariwise, orienting responses reduce heart rate [11]. In general, increased arousal caused by the threatening situations found in the high threat zones result in heart rate increases. Additionally, heart rate is used by the Profiling Module to provide a useful index of cognitive workload, which can be applied to the response characteristics related to performing the HMMWV Stroop task [12], [13].

2) Electrodermal Activity – Electrodermal Activity is assessed in the Profiling Module using a low pass filter applied to the electrodermal signal in order to minimize the effects of random noise in the signal. Median skin conductance response levels (SCLs) are calculated for each zone (safe and ambush) and analyzed, as are spontaneous fluctuations (SFs) in the skin conductance trace. SFs were defined as an increase of greater than 0.01 μS, with peak amplitude coming within a window of 1 to 3 seconds following the initial increase in the signal. The number of SFs occurring in each sample window is calculated (see Figure 3).
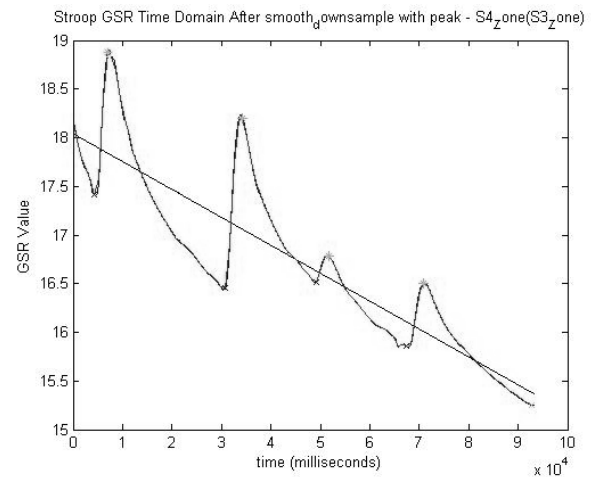

Fig. 3. Matlab scoring program output for one zone in the HMMWV Stroop task (VRST) scenario. Asterisks denote the presence of a spontaneous fluctuation. The downward trending straight line is the average slope of the signal across the zone. This data represents a user experiencing a low threat zone and performing the interference portion of the Stroop task.

The eccrine sweat glands, which are the basis of skin conductance measurement, provide an atypical and useful index of autonomic functioning in that they are innervated exclusively by the sympathetic nervous system, as opposed to cardiovascular measures which are influenced by parasympathetic vagal tone as well. Thus, skin conductance responding can be utilized as a direct measure of sympathetic activation. Of particular interest for military relevant VE applications is the way skin conductance responses are affected by fear inducing or threatening stimuli [9], in which greater levels of fear responding result in greater skin conductance response potentiation and increased skin conductance levels. Additionally, skin conductance will generally increase as workload increases. During cognitive workload tasks, as the task difficulty increases, skin conductance responses become larger [10], [13].

3) Respiration — Respiration is recorded in the Profiling Module using interbreath interval (IBrI), which is scored for each sample within a window by calculating the median interval in milliseconds between peak amplitudes in the raw data signal, signifying one full breath cycle. As with IBI, the dominant frequency of the signal is obtained by bandpass filtering (between 0.134 and 1.01 Hz) and returning the

highest magnitude frequency bin. The frequency is then used to calculate IBrI within the sample window which is compared to the mean IBrI from the previous window to determine the necessity of a change in the adaptive environment. Respiration rate has consistently been shown to increase in response to heightened levels of arousal associated with arousing stimuli [14], [20]. Mental imagery of fearful events has also been shown to increase rates of respiration [21]. Further, an increase in respiratory rate has been consistently associated with increased cognitive demand [12], [22]. During an air traffic control simulation with three levels of task difficulty, air traffic controllers exhibited increased rates of respiration as task difficulty increased [23].

4) Pupilometry — (electrooculographic; EOG) the median pupil radius (measured in mm) for all samples within a window. The EOG signals are evaluated by laboratory-developed software that detects blinks and provides interblink intervals. Higher EOG responses indicate higher levels of arousal.

Each of the features discussed is a statistically meaningful aggregation of data across spans of time in the AVE. Care must be taken in determining the duration of the sampling window. Smaller windows allow the virtual environment to be adapted more rapidly than larger windows, however selecting too small of a window size may not provide enough samples to extract meaningful features. Further, for AVEs, each step of feature extraction must be done in real-time. Thus, the algorithms used for feature extraction are limited by time constraints. For example, peak detection of each R-wave in an electrocardiogram recording for calculation of IBIs is not feasible, as movement artifacts and other producers of noise in the signal would typically require manual scoring which is inappropriate for real-time scoring. For this reason, we are continually incorporating increasingly robust feature extraction algorithms, such as frequency domain analyses making use of discrete Fourier transforms (DFT). For the features that require performing a DFT on the data (i.e., IBI and IBrI), we used the open source software library "Fastest Fourier Transform in the West" that ranks among the fastest performing DFT algorithm suites [24]. For SCL median values are calculated in order to reduce the effects of random noise in the signal.

To summarize, neurocognitive and psychophysiological metrics provide a means of obtaining objective and ongoing measures of user-state through noninvasive and nonconscious methods to improve experimental control. Arousal and cognitive workload are two aspects of participant-state that provide vital information for the successful implementation of adaptive systems that can be applied to improve real-world performance.

B. Controller Module

After the profiling run, the user is placed in the AVE. During the simulation, all the aforementioned psychophysiological features are computed across a fifteen second sliding window. To combine all psychophysiological features, we use an approach similar to [25], in which the user is said to be in a hyper- or hypo-aroused state if and only if all

psychophysiological features indicate that level of arousal. The median values for each of the psychophysiological features are computed across all windows occurring in each zone. These values form the thresholds to determine how to change the adaptive environment.

In addition to psychophysiological metrics, the user's cognitive performance on a task is also measured across the sliding window using a Humvee following simulation. Game players were instructed to maintain their initial separation from the lead vehicle (LV) while traveling through a virtual city despite changes in speed of the LV. The driver's vehicle is positioned 23 meters behind the LV and s/he is instructed to maintain the distance (with a tolerance of 5 meters) through acceleration and deceleration (see Figure 4).

Fig. 4. Humvee following simulation. Game players are instructed to maintain their initial separation from the lead vehicle (LV) despite changes in speed of the LV. The driver's vehicle is positioned 23 meters behind the LV and s/he is instructed to maintain the distance through acceleration and deceleration.

An important aspect of computing the cognitive and psychophysiological features was classification of the user from various signals and commands. A difficulty that we found inherent in this process was that the classifier used for establishing a user's level of arousal and cognitive performance was trained from data drawn from other subjects (Normative Database). It was challenging to develop generic cognitive and arousal recognition algorithms with model parameters fitting all subjects. In our initial attempts to do this, we used a support vector machine (SVM) to classify arousal levels that were drawn from the profiling module while users were immersed in VRCPAT [19]. Results revealed that when each user is considered separately, an average classification rate of 96.5% can be obtained; however, the average classification rate was much lower (36.9%, close to random guess) when a user's performance was predicted using data of other users' arousal levels.

Given the individual differences in arousal on VRCPAT tests, we needed to develop optimal classification strategies from a minimum number of user-specific training examples. We attempted to use the Normative Database to maximize the user-specific training data and then re-tuning the parameters of the profiling and control modules. Unfortunately,

continually returning to the Normative Database for user-specific data is time-consuming. As a result, we set out to minimize this effort and identify the optimal parameters of profile and control algorithms using a minimum number of user-specific training examples.

First, we tailored the profile and control algorithms by collecting a batch of user-specific training examples at a time, and then we estimated the recognition performance by leave-one out cross-validation, until the cross-validation performance was satisfactory, or the maximum number of iterations/training examples was reached. To do this we first used a transfer learning approach in which we made use of the information contained in other subjects' training examples [26]. Although training examples from other subjects may not be completely consistent with a new subject's profile, they still contain some useful information, as users exhibit similar (though usually not exactly the same) behaviors at the same arousal state. This is especially significant when the number of user-specific training examples is very small. We have found that transfer learning offers us a promising technique for handling individual differences in our control module, and that it could help tailor our system for individual users with very few user-specific training examples [27].

In addition to using data from other subjects through transfer learning, we found it necessary to optimally generate user specific training examples that are independent and complementary. Of primary importance was the need to have a way to process the information when the class labels needed to be determined before the training examples and features can be obtained. For example, in the arousal classification, the arousal (class label) needs to be specified before the corresponding responses can be generated and recorded. To overcome this challenge, we decided to explore the use of active class selection (ACS) because ACS was developed to optimally select the classes to obtain training examples so that a good classifier could be constructed from a small number of training examples [28]. Our experimental findings using a kNN classifier revealed consistent results in higher classification accuracy than a uniform sampling approach [29].

Determination of what subset of features must be above or below set point in order to create change in the AVE is also necessary. The first option was to make adaptations in the VE solely if all extracted features were above set point. This resulted in a VE that was comparatively resistant to change, but was not ideal for situations in which we desired a flexible training scenario that used adaptive assessment of the user's individual response patterns. Some individuals were very reactive according to variations in heart rate, but also showed little to no change in SCL over time. Next, we tried to allow for adaptation in the VE when any one feature surpassed the set point. This lead at times to an environment that was too prone to change and at times created an adaptive assessment scenario that became confusing and frustrating. The optimal solution that we found was to create adaptive changes to the environment based on the level of the majority of features relative to the set point. If four out of five selected features are above set point, the user is determined to be in a hyper-aroused state and the environment can be adjusted to decrease the intensity of the threatening stimuli, or the difficulty of the task.

Using a message broker which fully implements clustering, multiple message stores, and the ability to use any database [30], the neurocognitive and psychophysiological profile data from the Profiling Module are analyzed by the Controller Module and transferred to the Command Module as "command" signals. Values of the command correspond to changes in position and appearance of objects in the user's view of the VE, the instantiation or inhibition of immersive stimuli, and changes in the difficult of the primary task. The signals are fed to the "Command Module" to be processed by the VE.

C. Command Module

We developed a C++ plugin for the VBS2 engine that polls incoming messages once per game frame. When a command signal is detected, the environment can then be adapted via scripting calls within the VBS2 engine. For our AVE, we scripted a number of stimuli for the high stress zones:

1) Improvised explosive device blasts near the in-game avatar of the user.

2) Insurgent AI characters who fire at the in-game avatar of the user.

3) Changing the in-game weather to be cloudy and rainy.

4) Activation of a scent machine to release smells similar to gunpowder in the environment.

The severity of the stimuli can be tailored to how far the psychophysiological features exceed the mean for arousal state. For example, if the features are within one standard deviation of a hypo-aroused state then the explosion would be minor. If the features are two standard deviations below the mean for a hypo-aroused state, then the explosion would be severe and cause side effects such as briefly impaired vision. Stimuli can also be inhibited for several seconds to prevent the user from reaching a hyperaroused state.

In addition to increasing the stress of the AVE, the complexity of the Humvee follow task can be adjusted through variance of the lead vehicle's acceleration and deceleration. As in the arousal assessment above, cognitive performance metrics (distance headway between the LV and the following vehicle) are computed across a fifteen second sliding window. Herein, the user is said to be performing at expectation or deviating (above or below) from expected performance relative to the user's ability to maintain the 23 m distance from the LV. If the user is unable to maintain the 23 m distance relative to the LV, then the LV decelerates and maintains speed. If the user is consistently able to maintain the 23 m distance, then the LV's speed can become more variable.

The types of stimuli presented to the user and task difficulty adjustments are logged along with the user's neurocognitive performance and their raw psychophysiological readings. These logs are then parsed to score the neurocognitive functioning of the test subject.

The VRCPAT 2.0's "Command Module" has been designed to offer an adaptive virtual environment that can be explored by Military Service members under the supervision of a military psychologist. This virtual adaptive assessment and rehabilitation system aims to place the injured Military Service member into a state of optimal experience defined as "flow" to trigger a broad recovery process [31]. According to Csikszentmihalyi [32], "flow" is best understood as an optimal state of consciousness that is characterized by a state of concentration so focused that it results in complete immersion and absorption within an activity. Following the work of Fairclough [33], we partition the "flow" state of the Military Service member into four quadrants or "zones" (see Figure 2).

Our "Command Module" is being developed to place the patient (e.g., Military Service member that has experienced a blast injury and/or combat stress symptoms) in VRCPAT 2.0 at the optimal starting point for that Military Service member; Zone A. We do not conceptualize the flow of rehabilitation/treatment to be a static experience. A Military Service member's skill level tends to be low when first immersed in VRCPAT. As the patient's experience of the program increases, skills increase and s/he may become bored if the challenge remains constant (Zone C). Within VRCPAT2.0, the challenge will increase, but usually at a different rate than the Military Service member's ability level. Hence, the Military Service member is constantly in a state of flux between the four points shown in Figure 5.
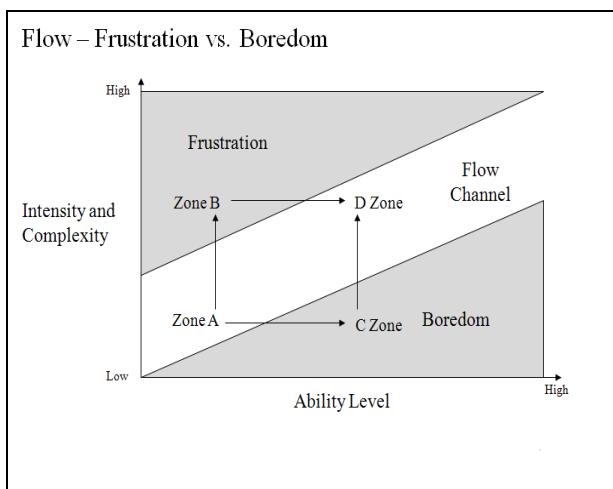

Figure 5: Two-dimensional representation of state of the user: Note, this is an adaptation of graphs found in Csikszentmihalyi [32] and Fairclough [33].

At times the user may begin to disengage (start to experience boredom and move toward Zone C) when the challenge does not increase in pace with his or her skills. At other times, the user may move towards frustration (Zone B) when s/he is slow to learn the necessary skills. Particularly relevant to Csikszentimihalyi concept of flow states is Zone B because it represents a "stretch" zone, in which the Military Service member is engaged and ability levels are being increased as they are pushed toward frustration. Fairclough [33] has explained that this state may be tolerated for short periods (e.g., learning phases and/or a demanding but

rewarding period of performance). Overall, the goal of VRCPAT 2.0 is to keep the Military Service member in Zone D—continually adapting the intensity and difficulty of the environment to have the Military Service member in a flow state with improved skills and being able to function at a higher level of challenge. This conceptualization allows the adaptive virtual environment to make a distinction between two states of low performance, both of which require different categories of adaptive response. For example, in Zone B, the intensity and complexity of the stimuli should be reduced until the ability level has been optimized. Further, if the Military Service member's results indicate that s/he is heading to Zone C, the VRCPAT should adapt so that task demands be increased. This complex representation of the Military Service member provides the "Controller Module" with greater specificity in order to target the adaptive response.

V. DISCUSSION

While advances in military relevant simulations provide potential for increasing assessment of Soldier readiness to Return-to-Duty, little has been done to develop these simulations into AVEs for assessing cognitive workload. Although there are other attempts to apply adaptive virtual environments to posttraumatic stress disorder [34], [35] and neurocognitive rehabilitation [36], this work is in its earliest stages. We have presented a framework to do enhanced neurocogntive assessment using an AVE that is being validated at West Point, the University of Southern California, Madigan Army Medical Center, and Tripler Army Medical Center. The serious game described herein offers a dynamically adapting functional assessment of patients. Additionally, the described project is novel in that it includes both psychophysiological and neurocognitive profiles to enhance its adaptive capabilities. Further, the neuropsychological tests found in the Normative Database include a set of neuropsychological tests that have been validated against standard paper-and-pencil, as well as computerized neuropsychological measures (e.g., Automated Neuropsychological Assessment Metrics). To the best of our knowledge, there is no other project that does both psychophysiological and neurocognitive profiling using the level of validation provided by the VRCPAT 2.0.

A real-time adaptive virtual environment that is sensitive to cognitive and emotional aspects of user experience, as delineated in this manuscript, is considered to be the future alternative for devising cognitive assessment and training measures that will have better ecological/predictive validity for real-world performance.

Further, the flexibility of stimulus delivery and response capture that are fundamental characteristics of such adaptive virtual environments is viewed as a way for military psychology objectives to be addressed in a more efficient fashion for long term needs. Such flexibility would allow for this system to be viewed as an open platform on which a wide range of research questions could be addressed that would have significance to military psychologists.

REFERENCES

[1] T. Parsons, & C. Courtney, "Neurocognitive and psychophysiological interfaces for adaptive virtual environments," in Human Centered Design of E-Health Technologies, C. Röcker, T. & M. Ziefle, Eds. Hershey: IGI Global, pp. 208-233, 2011.

[2] M. Ambinder, "Biofeedback in gameplay: how Valve measures physiology to enhance gaming experience," presented at the Game Developer's Conference, 2011.

[3] C. Liu, P. Agrawal, & S. Chen, "Dynamic difficulty adjustment in computer games through real-time anxiety-based affective feedback," Int'l J. Human-Computer Interaction, vol. 25, pp. 506-529, 2009.

[4] R. Yerkes, & J. Dodson, "The Relation of Strength of Stimulus to Rapidity of Habit-Formation", Journal of Comparative Neurology and Psychology, vol. 18, pp. 459-482, 1908.

[5] T. Parsons, A. Iyer, L. Cosand, C. Courtney, & A. Rizzo, "Neurocognitive and psychophysiological analysis of human performance within virtual reality environments," Studies in Health Technology and Informatics, vol. 142, pp. 247-252, 2009.

[6] M. Barlow, M. Luck, E. Lewis, M. Ford, R. Cox. "Factors in Team Performance in a Virtual Squad Environment", Proceedings of SimTecT04, Canberra, May, pp. 94-99, 2004.

[7] T. Parsons, C. Courtney, L. Cosand, A. Iyer, A. Rizzo, & K. Oie, "Assessment of Psychophysiological Differences of West Point Cadets and Civilian Controls Immersed within a Virtual Environment," Lecture Notes in Artificial Intelligence, vol. 5638, pp. 514–523, 2009.

[8] T. Parsons, C. Courtney, A. Rizzo, J. Edwards, & G. Reger. "Virtual reality paced serial assessment test for neuropsychological assessment of a military cohort." Studies in Health Technology and Informatics, 173. pp. 331-337, 2011.

[9] L. Williams, L. Phillips, M. Brammer, D. Skerrett, J. Lagopoulos, C. Rennie, H. Bahramali, G. Olivieri, A. David, A. Peduto, & E. Gordon. "Arousal dissociates amygdala and hippocampal fear responses: Evidence from simultaneous fMRI and skin conductance recording." NeuroImage, 14, pp. 1070–1079. 2001.

[10] N. Kobayashi, A. Yoshino, Y. Takahashi, & S. Nomura. "Autonomic arousal in cognitive conflict resolution." Autonomic Neuroscience: Basic and Clinical, 132, pp. 70–75, 2007.

[11] G. Berntson, S. Boyson, & J. Cacioppo. "Cardiac orienting and defensive responses: Potential origins in autonomic space." In B.A. Campbell, H. Hayne, & R. Richardson (Eds.), Attention and information processing in infants and adults: Perspectives from human and animal research (pp. 163–200). Hillsdale, NJ: Erlbaum, 1992.

[12] B. Mehler, B. Reimer, J. Coughlin, & J. Dusek. "Impact of incremental increases in cognitive workload on physiological arousal and performance in young adult drivers." Journal of the Transportation Research Board, 2138, pp. 6–12, 2009.

[13] Y. Hoshikawa and Y. Yamamoto, "Effects of Stroop Color-Word Conflict Test on the Autonomic Nervous System Responses," AJP—Heart and Circulatory Physiology, vol. 272, no. 3, pp. 1113-1121, 1997.

[14] F. Boiten, N. Frijda, & C. Wientjes. "Emotions and respiratory patterns: Review and critical analysis." International Journal of Psychophysiology, 17, pp103–128, 1994.

[15] T. Parsons, A. Rizzo, J. Bamattre, & J. Brennan. "Virtual Reality Cognitive Performance Assessment Test." Annual Review of CyberTherapy and Telemedicine, 5, pp. 163-171, 2007.

[16] T. Parsons, & A. Rizzo. "Initial Validation of a Virtual Environment for Assessment of Memory Functioning: Virtual Reality Cognitive Performance Assessment Test." Cyberpsychology and Behavior, 11, pp. 17-25, 2008.

[17] T. Parsons, T. Silva, J. Pair, & A. Rizzo. "A Virtual Environment for Assessment of Neurocognitive Functioning: Virtual Reality Cognitive Performance Assessment Test." Studies in Health Technology and Informatics, 132, pp. 351-356, 2008.

[18] T. Parsons, C. Courtney, B. Arizmendi, & M. Dawson, "Virtual reality Stroop task for neurocognitive assessment," Studies in Health Technology and Informatics, vol. 143, pp. 433-439, 2011.

[19] D. Wu, C. Courtney, B. Lance, S. Narayanan, M. Dawson, K. Oie, & T. Parsons, "Optimal arousal identification and classification for affective computing: virtual reality Stroop task," IEEE Transactions on Affective Computing, vol. 1, pp. 109-118, 2010.

[20] S. Kreibig, F. Wilhelm, W. Roth, & J. Gross. "Cardiovascular, electrodermal, and respiratory response patterns to fear- and sadness-inducing films." Psychophysiology. 44, pp. 787–806, 2007.

[21] P. Rainville, A. Bechara, N. Naqvi, & A. Damasio. "Basic emotions are associated with distinct patterns of cardiorespiratory activity." International Journal of Psychophysiology, 61, pp. 5–18, 2006.

[22] R. Backs, & K. Seljos. "Metabolic and cardiorespiratory measures of mental effort: The effects of level of difficulty in a working memory task. International Journal of Psychophysiology", 16, pp. 57–68, 1994.

[23] J. Brookings, G. Wilson, & C. Swain, C.R. "Psychophysiological responses to changes in workload during simulated air traffic control." Biological Psychology, 42, pp. 361–377, 1996.

[24] M. Frigo, & S. Johnson. "The design and implementation of FFTW3". Proceedings of the IEEE, 93, pp. 216–231, 2005.

[25] A. Haarmann, B. Boucsein, F. Schaefer, "Combining electrodermal responses and cardiovascular measures for probing adaptive automation during simulated flight," Applied Ergonomics, vol. 40, pp. 1026-1040, 2009.

[26] S. Pan, & Q. Yang. "A survey on transfer learning." IEEE Transactions on Knowledge and Data Engineering 22, pp. 1345–1359, 2010.

[27] D. Wu, & T. Parsons. "Inductive Transfer Learning for Handling Individual Differences in Affective Computing." Lecture Notes in Computer Science, 6975, pp. 142-151, 2011.

[28] R. Lomasky, C. Brodley, M. Aernecke, D. Walt, M. & Friedl. "Active class selection." In: Proceedings of the 18th European Conference on Machine Learning, Warsaw, Poland, pp. 640–647, 2007.

[29] D. Wu., & T. Parsons. "Active Class Selection for Arousal Classification." Lecture Notes in Computer Science, 6975, pp. 132-141, 2011.

[30] B. Snyder, D. Bosanac, R. Davies, (2010), ActiveMQ in Action (1st ed.), Manning Publications.

[31] G. Riva, F. Mantovani, & A. Gaggioli, "Presence and rehabilitation: toward second-generation virtual reality applications in neuropsychology," Journal of NeuroEngineering and Rehabilitation, vol. 1:9, 2004.

[32] M. Csikszentmihalyi. Flow: The psychology of optimal experience. New York, NY: HarperCollins, 1990.

[33] S. Fairclough, "Fundamentals of physiological computing," Interacting with Computers, vol. 21, pp. 133-145, 2009.

[34] K. Cosic´, S. Popovic´, & T. Jovanovic, "Physiology-driven adaptive VR system: Technology and rationale for PTSD treatment," Annual Review of CyberTherapy & Telemedicine, vol. 5, pp. 179–91, 2007.

[35] S. Popovic´, M. Horvat, & D. Kukolja, "Stress inoculation training supported by physiology-driven adaptive virtual reality stimulation," Studies in Health Technology & Informatics, vol. 144, pp. 50–54, 2009.

[36] L. Hettinger, & M. Haas, Virtual and adaptive environments: Applications, implications, and human performance issues. Lawrence Erlbaum Associates, 2003.

BIOGRAPHIES

Thomas D. Parsons, PhD is Director of the Neuropsychology and Simulation Laboratory, Assistant Research Professor, and Research Scientist at the University of Southern California's Institute for Creative Technologies. He developed the Virtual Reality for Cognitive Performance and Adaptive Treatment system. In addition to his patents (with eHarmony.com), he has over 175 publications in peer-reviewed journals and other fora.

James Reinebold, M.S. earned a B.S. in Computer Science from the Georgia Institute of Technology in 2007. He is also recent graduate of the Master's of Computer Science program at the University of Southern California, where he completed a thesis on evaluating the sensing capabilities of mobile phones. His research interests include adaptive virtual environments and applications of artificial intelligence to games.