Perceptually-Informed Virtual Environment (PerceiVE) Design Tool

Anna Skinner[1], Jack Vice[1], Corinna Lathan[1], Cali Fidopiastis[2], Chris Berka[3], and Marc Sebrechts[4]

[1] AnthroTronix, Inc. 8737 Colesville Rd., L203 Silver Spring, MD 20910
[2] Institute for Simulation and Training, University of Central Florida 3100 Technology Parkway, Orlando, FL 32826
[3] Advanced Brain Monitoring, Inc. 2237 Faraday Ave., Ste 100 Carlsbad, CA 92008
[4] Catholic University of America, Department of Psychology, 620 Michigan Ave, NE Washington, DC 20064
{askinner,jvice,clathan}@atinc.com, cfidiopia@ist.ucf.edu, chris@b-alert.com, sebrechts@cua.edu

Abstract. Virtual environments (VE's) are becoming more and more prevalent as training tools for both military and civilian applications. The common assumption is that the more realistic the VE, the better the transfer of training to real world tasks. However, some aspects of task content and fidelity may result in stronger transfer of training than even the most high fidelity simulations. This research effort seeks to demonstrate the technical feasibility of a Perceptually-informed Virtual Environment (PerceiVE) Design Tool, capable of dynamically detecting changes in operator behavior and physiology throughout a VE experience and comparing those changes to operator behavior and physiology in real-world tasks. This approach could potentially determine which aspects of VE fidelity will have the highest impact on transfer of training. A preliminary study was conducted in which psychophysiological and performance data were compared for a visual search tasks with low and high fidelity conditions. While no significant performance effects were found across conditions, event-related potential (ERP) data revealed significant differences between the low and high fidelity stimulus conditions. These results suggest that psychophysiological measures may provide a more sensitive and objective measure for determining VE fidelity requirements.

Keywords: Psychophysiological Measures, Virtual Environments, Fidelity, Transfer of Training, Simulation Design.

1 Introduction

Virtual environments (VE's) and simulations are being employed for training applications in a wide variety of disciplines, both military and civilian. Technological advances are enhancing the ability of developers to create VE's with visual, auditory, haptic, and even olfactory realism. Such VE's allow the military to train skills that are too costly, too dangerous, or are otherwise impossible to practice. While a significant

D.D. Schmorrow et al. (Eds.): Augmented Cognition, HCII 2009, LNAI 5638, pp. 650–657, 2009.

research has been conducted examining the transfer of training from VEs (for example, [1, 2]), only a limited number of efforts have used psychophysiological measures to do so. The common assumption is that the more realistic the VE, the better the transfer of training to real world tasks. However, some fidelity components (e.g., display resolution, frame rate, texture mapping, physics modeling, etc.) may result in stronger transfer of training than others for a given task or domain. This has traditionally been determined by performance measurements compared before and after design iterations. With each design modification, end users are tested using the VE and their performance is compared to performance on the prior VE design. Improved performance is often assumed to be related to improved design and fidelity. However, it is difficult to identify the specific design components that directly relate to transfer of training improvements. Furthermore, this method of design focuses on trial and error, and is therefore time consuming, undirected, and may result in false associations between performance and VE characteristics. For example, unless each component of the new simulator design is introduced separately, it will not be known which fidelity design improvements bear the strongest significance to performance improvements. Thus, a more sensitive, objective, and comprehensive assessment of the quality of interaction with a simulation is needed to effectively identify the specific components of simulation that bare relevance to real world operational tasks.

One of the major questions simulation designers must address is "what components of fidelity have the greatest impact on transfer of training?" Fidelity is defined as the degree to which features (e.g., visual, auditory, etc) in the Virtual Environment (VE) match features in the real environment. Following this premise, one can argue that a VE with maximum fidelity would result in transfer of training equivalent to real-world training since the two environments would be impossible to differentiate [3; Martin, 1981). However, developers are limited by practical restrictions such as cost, time, and development resources. Thus, trade-offs are necessary. There is currently a limited understanding of the specific trade-offs between increases in simulation fidelity and operator behavior, and essentially no guarantee to developers that a particular level/area of simulation fidelity is sufficient to provide effective transfer of training.

Under an Office of Naval Research-funded Small Business Technology Transfer (STTR) effort the authors proposed to develop a Perceptually-informed Virtual Environment (PerceiVE) Design Tool, which utilizes physiological measures to determine fidelity requirements with the goal of optimizing transfer of training between simulated and real world tasks. We hypothesized that a physiologically-based system capable of dynamically detecting changes in operator behavior and physiology throughout a VE experience, and comparing those changes to operator behavior and physiology in real-world tasks, could potentially determine which aspects of VE fidelity will have the highest impact on transfer of training.

EEG and event related potential (ERP) approaches offer excellent temporal resolution for tracking of neural activity representing the flow of information from sensory processing, detection and identification of relevant objects, and decision-making. ERP signature components associated with the identification of target stimuli were first reported in 1965 and named "P300s or P3b or Late Positivity" [4, 5], (Squires, Squires, & Hillyard, 1975) because target stimulus presentations are associated with large positive potentials maximal over parietal cortex with peak latency ranging from 300-800 ms after presentation of the target stimulus. The P300 is generally accepted to be a post-sensory signal elicited when subjects attend and respond to target stimuli and is believed to be related to higher cognitive processes including updating working memory [5]. Several reports suggest that when target stimuli are degraded, obscured or difficult to recognize, the amplitude of the P300 is decreased (Kok, 1985, Kok, 1980), [6].

In addition to the extensive work on describing the P300, a growing body of ERP evidence reveals ERP neural signatures of target recognition and discrimination as early as 150-200 milliseconds post-stimulus (Hopf, 2002, Vogel and Luck, 2000), [7, 8]. Johnson and Olshausen [9] demonstrated an early object recognition arising around 135 ms when low-level feature discrimination was present. These studies suggest that basic discriminative processing (e.g. differentiating faces from words, animals from non-animals, shape or color distinctions) is performed so rapidly that it must be accomplished in one feed-forward sweep of activity propagated through the visual system integrating basic visual processing with top-down template models [10]. These recent investigations quantify the difference between target and non-target ERPs to reveal distinctive ERP signatures occurring as early as 150-200 ms post-stimulus and maintained for up to 800-1000msec. post-stimulus. The differences have been identified following the presentation of objects that vary only in their target status. These target-related neural signatures provide an index of the time when object recognition is sufficiently complete for the brain to initially discriminate "targetness". These early target-related differences may reflect facilitated sensory processing (i.e., enhanced neural responses associated with matching to a top-down target template) or to decision-related post-sensory processing and recognition.

2 Method

An extensive literature review was conducted to assist in determining the appropriate classification of VE fidelity components, the trade-offs between VE fidelity components and overall VE fidelity, ways in which fidelity components can be objectively measured, and which components are most likely to have a significant impact on an observation task. This review included an investigation of human information processing (HIP) and visual perceptual skills; as well as prior research relating performance differences to various levels of VE fidelity, physiological assessment during VE-based tasks, and the effects of photorealism on task performance.

A study was then designed to determine whether physiological measures could be used to detect simulation fidelity. The experimental design and VE task environment were developed based on the literature review and resulting targeted objectives. A static, VE-based visual search task consisting of militarily-relevant vehicles in low and high fidelity conditions was developed using computer-aided drafting (CAD) software.

The stimuli consisted of a series of images containing 4 objects, one in each corner of the screen. At least 3 of the objects were distractors; the fourth was either a distractor or the target object. The target object was identified prior to the trials and remained consistent throughout the trials. In the low fidelity (LoFi) condition, minimal polygon count was used, with each object ranging from 9-14 traingles depending on its inherent complexity, and no contrast existed within each object. A sample LoFi

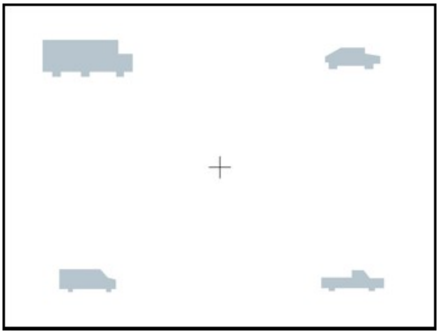

Fig. 1. Sample low fidelity (LoFi) stimulus

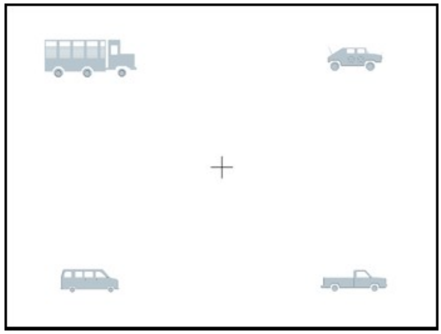

Fig. 2. Sample high fidelity (HiFi) stimulus

stimulus is shown in Figure 1. In the high fidelity condition (HiFi) the polygon count of each object was increased by 30x (+/-3%) and contrast was added, emphasizing depth and contours within each object. In both conditions the object color and background color remained constant (gray and white, respectively). A sample HiFi stimulus is shown in Figure 2.

At the start of the experiment, subjects were informed that they would be shown a series of graphics of varying detail containing a cross in the center of the screen and four objects selected from the following: battle tank, commercial truck, pick-up truck, humvee, and van. An instruction screen (shown in Figure 3) was then displayed, providing the subjects with a likeness of target objects for the low and high fidelity conditions, as well as instructions to keep their eyes focused on the cross throughout the search task.

Fig. 3. Task Instruction Screen

Low and high fidelity images were then displayed in random order on a 19-inch monitor. Subjects were positioned 30 inches from the display. Approximately 50% of the presentations contained a target object.

A total of 12 participants each performed two 10-minute consecutive trials consisting of 200 stimuli presentations per trial. Each image was displayed for 2 seconds, with a 1 second inter-stimulus interval (ISI). Trials in which the participant did not provide a response within the 2-second display period were considered missed trials; these trials were reported separately from incorrect and correct response trials.

Physiological measures collected during the trials included electroencephalogram (EEG), heart rate, galvanic skin response, and eye tracking. The B-Alert® wireless Sensor Headset from Advanced Brian Monitoring (ABM) was used to acquire EEG data from 9 sites (F3, F4, C3, C4, P3, P4, Fz, Cz, and POz), referenced to linked-mastoids. The Wearable Arousal Meter (WAM) collected heart rate data, which was used to calculate arousal. Galvanic Skin Response (GSR) was assessed using the Thought Technologies Procomp System; and Eye tracking was measured via an Arrington system. A DLL was implemented to allow the EEG signal to be synchronized with the other physiological measures and the task stimuli, which were presented within a custom program using E-prime experiment management software. In addition to the task performance and physiological data, a post-task questionnaire was given to the participants.

The independent variable for this task was the fidelity condition (low or high). The dependent variables included the physiological response data, as well as the task performance data (i.e., reaction time and accuracy).

3 Results

3.1 Physiological Results

Eye Tracking. For this preliminary study, the eye-tracking data was used for the purpose of identifying if and when subjects looked away from the cross in the center of the screen during the task. Of the 12 subjects, 6 consistently looked around the screen at the stimulus objects, while the remaining 6 kept their eyes fixated on the center of the screen.

EEG. Initial data analysis was conducted for the 6 subjects that completed the task as instructed, without moving their eyes from the cross in the center of the screen, and included only the midline electrode sites (Fz, Cz and POz) as a preliminary assessment. Absolute/relative power spectral density (PSD) variables were computed for each 1-second epoch. Metrics for "engagement" and "workload" were calculated using quadratic and linear discriminant function analyses of model-selected PSD variables (1-Hz bins,1-40Hz). Event Related Potentials (ERPs) were derived based on time-locking to the presentation of the stimuli (1-second post-stimulus) or to the 1-second prior to the response. acquired from 9 scalp sites at 256 samples/sec.

ERP waveforms were combined into grand averages. All ERP waveforms were computed using only trials on which the subject correctly identified the test stimulus as either a target or a nontarget, and all were time-locked to the presentation of the test stimulus. Before averaging, all data were artifact rejected on a trial-by-trial basis for eyeblinks, excursions and excessive muscle activity using automated in-house software [11]. Trials with predominant alpha activity (present in two of the six participants) were not eliminated to allow for sufficient numbers of trials in each

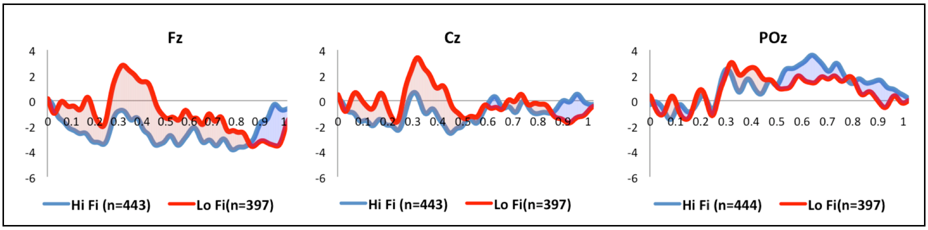

Fig. 4. Averaged ERP data for midline electrodes

average. No additional smoothing or filtering was applied on this preliminary investigational analysis. Figure 4 presents the averaged ERP waveforms for the Fz, Cz, and POz electrodes.

ERP templates for high fidelity/low fidelity (HiFi/LoFi) and targets/non-targets (T/NT) were consistently identified across the six participants. Distinctions between the HiFi/LoFi T/NT ERP templates were evident as early as 200ms post-stimulus onset and were sustained for windows in excess of 900ms post-stimulus. Maximal T/NT differences varied from 400-900msec No attempt was made in this pilot study to sort ERPs based on reaction times, which varied significantly within- and between participants (mean reaction times ranging from 550 – 1450 ms.).

Preliminary data analysis in the present study suggests that an early (onset of 200-250msec. post-stimulus) frontal-central positivity is present for all correctly identified targets and non-targets with an increase in amplitude for degraded (or low fidelity) stimuli across all 6 subjects. A much later parietal positivity (peaking between 500-700ms post-stimulus) which is likely to be a true P300 or P3b type component is evident for correct targets and is of higher amplitude for high fidelity stimuli when compared to the low fidelity. This late P300 component confirms previous reports (Kok, 1985) of degraded stimuli eliciting reduced amplitude P300.

Other Results. Arousal levels were averaged and a significant fidelity effect was found for 1 of the 6 subjects who performed the task without moving their eyes. GSR cannot be mapped to individual trials due to latency.

3.2 Performance Results

No significant effect was found as measured by reaction time and accuracy of responses for fidelity or fixation conditions; however, the task was quite simple, and thus a ceiling effect was evident.

4 Discussion

While no significant performance effects were found across conditions, consistent and detectable differences in ERP data were observed for subjects performing the visual search task in low and high fidelity conditions. Accurate identification of HiFi vs. LoFi targets was shown to elicit distinctive ERP components. Two components distinguish LoFi from HiFi: early frontal-central (250-500ms) and late parietal

(500-800ms). The early frontal-central positivity clearly distinguished the LoFi from HiFi ERPs for all participants. Though preliminary, these data suggest evidence of an early feature extraction process. Based on studies of ERPs during visual search, where individuals scan through a set of visual stimuli for a particular target, Luck and Hillyard, (1994) proposed that a spatial filtering process conducts a preliminary analysis of the stimulus array containing relevant features. They identified a component of the visual ERP in the range of 200-250msec. post-stimulus elicited by visual search arrays that varied in accordance with the filtering of distractors. When a target was distinguished by a salient feature, spatial filtering began approximately 175 ms after search onset. Second, the filtering process is dependent on the outcome of a preliminary stimulus analysis capable of rejecting non-targets on the basis of simple feature information. Alternatively the early positivity may be a reflection of working memory processes in prefrontal cortex.

The arousal data results demonstrated that some measures are more sensitive to fidelity variations than others. Some sensors can be considered as global measures, such as the skin conductance and arousal. The EEG is more specific and localized. In future studies, the eyetracking data will also be used to compare pupilometry during low and high fidelity conditions.

These results suggest that psychophysiological measures, specifically ERP, may provide a more sensitive and objective measure than traditional metrics for determining VE fidelity requirements. This research is currently being leveraged within a perceptual skills VE task in which performance is impacted significantly by fidelity degradation. Future research will compare physiological data collected in equivalent real world (RW) and VE tasks to further determine the impact of various fidelity components on task performance and training transfer.

References

1. Lathan, C.E., Tracey, M.R., Sebrechts, M.M., Clawson, D.M., Higgins, G.: Using virtual environments as training simulators: Measuring transfer. In: Stanney, K. (ed.) Handbook of Virtual Environments: Design, Implementation, and Applications, pp. 403–414. Lawrence Erlbaum Associates, Mahwah (2002)
2. Sebrechts, M.M., Lathan, C.E., Clawson, D.M., Miller, M.S., Trepagnier, C.: Transfer of training in virtual environments: issues for human performance. In: Hettinger, L.J., Haas, M.W. (eds.) Virtual and Adaptive Environments; Applications, Implications, and Human Performance Issues, pp. 67–90. Lawrence Erlbaum Associates, Mahwah (2003)
3. Waller, D., Hunt, E., Knapp, D.: The transfer of spatial knowledge in virtual environment training. Presence: Teleoperators and Virtual environments 7(2), 129–143 (1998)
4. Sutton, S., Braren, M., Zubin, J., John, E.R.: Evoked-potential correlates of stimulus uncertainty. Science 150, 1187–1188 (1965)
5. Donchin, E., Coles, M.G.H.: Is the P300 component a manifestation of context updating? Behavioral and Brain Sciences 11, 357–374 (1988)
6. Verleger, R.: Event-related potentials and cognition: A critique of the context updating hypothesis and an alternative interpretation of P3. Behavioral and Brain Sciences 11, 343–427 (1988)
7. Thorpe, S., Fize, D., Marlot, C.: Speed of processing in the human visual system. Nature 381, 520–522 (1996)

8. Fabre-Thorpe, M., Delorme, A., Marlot, C., Thorpe, S.: A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. Journal of Cognitive Neuroscience 13, 171–180 (2001)
9. Johnson, J.S., Olshausen, B.A.: Timecourse of neural signatures of object recognition. Journal of Vision, 3(7), 499-512 (2003), `http://journalofvision.org/3/7/4/`
10. VanRullen, R., Thorpe, S.J.: The time course of visual processing: From early perception to decision-making. Journal of Cognitive Neuroscience 13, 454–461 (2001b)
11. Berka, C., et al.: Real-time Analysis of EEG Indices of Alertness, Cognition and Memory with Wireless EEG Headset. International Journal of Human-Computer Interaction 17(2), 151–170 (2004)