Proceedings of the 2016 Winter Simulation Conference
T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, eds.

AN ANALYSIS OF QUESTIONNAIRES AND PERFORMANCE MEASURES FOR A SIMULATION-BASED KINESIC CUE DETECTION TASK

Jonathan Hurter
William Aubrey
Sushunova G. Martinez
Crystal S. Maraj

Institute for Simulation and Training
University of Central Florida
3100 Technology Parkway
Orlando, FL 32826, USA

Irwin Hudson

U.S. Army Research Laboratory – Human Research and Engineering Directorate Advanced Training and Simulation Division
12423 Research Parkway
Orlando, FL 32826, USA

ABSTRACT

The attraction of Simulation-Based Training for unmanned Intelligence, Surveillance, and Reconnaissance tasks has sparked testing for instructional strategies in a kinesic cue detection task. Early evidence of training effectiveness for this task is manifested by performance and self-report measures. The wealth of surveys collected include aspects of users' technology acceptance, immersion, intrinsic motivation, stress, workload, and demographics. This paper reviews these detection task measures in light of an instructional strategy, Kim's Game. A cross-scale analysis of the provided measures indicates strong correlations between several subscales. An investigation of potential predictors of performance indicates weekly computer use is statistically significant in predicting a user's Posttest Median Response Time for behavior cue detection. Recommendations for future initiatives include adding feedback, questioning concern for increasing immersion, and comparing results to other instructional strategies.

1 INTRODUCTION

1.1 Simulation-Based Training

Simulation-Based Training (SBT) affords learners a synthesized representation of an environment or space where instruction is presented (Martin et al. 2010). A simulation is an approximation of a real or hypothetical referent that can ultimately permit a user an active role in creating meanings of real-life concepts. Prime rationales for simulations include the mitigation of risks (e.g., time, costs, and safety) for stakeholders, and the performance aspects of users practicing and assessing skills (Lyons et al. 2002).

The U.S. Department of Defense (2009) has laid out a goal to exacerbate unmanned systems support for warfighters, which raises the demand for solutions that can rapidly meet performance needs. Since a timely transition will partly rely on training effectiveness, simulations are proposed here to alleviate concerns. Specifically, the rise for a need in Intelligence, Surveillance, and Reconnaissance (ISR) competencies suggests a consequent need for suitable training technologies. For ISR tasks, particular interest is placed on researching Combat Profiling, an aspect of irregular warfare where situational awareness and perceptual skills are applied to judge situations and anticipate disruptive events (Fautua et al. 2010).

1.2 Behavior Cue Detection

As an outcome of a training initiative for Combat Profiling (Fautua et al. 2010), SBT was recommended as a potential part of Combat Profiling curricula, where behavioral cue detection provides an array of uses: a method for gauging norms of cultural activity (and conversely, anomalies), determining intent of a threatening nature, and pinpointing influential targets, or High-Value Individuals. Basic detection of an individual's threatening body-related cues is a prerequisite skill for decision-making, placing the signal detection task as one piece of a warfighter's effort to analyze a field situation. Kinesic cues are one form of behavioral identifiers, and are defined as non-verbal movements free of conditioned learning, that conform to the particular cultural influences exerted on an individual (DeVito 1968, Birdwhistell 1970). Kinesic cues do not rely necessarily on verbal output to convey a particular attitude, feeling, or intention.

According to Signal Detection Theory (SDT), all detection tasks involve a signal of interest that an operator is subject to discern among a cloud of distraction, or noise (Wickens 1992). Any effort of detection is made in the face of ambiguous signals, the latter which may vary during tasks.

A behavior cue detection task also involves memory components, where memory is used in the process of observing, analyzing, and comparing elements for pattern recognition. Unfortunately, training for pattern recognition for military applications is wanting (Fischer and Geiwetz 1996). When a scene is viewed, information is kept provisionally in short-term or working memory (Peterson 1966; Baddeley and Hitch 1974; Rensink 2002). The items stored in such brief memory mechanisms are compared to items solidified in long-term memory to decide if a signal (i.e., a match) exists. SDT and memory concepts may be leveraged to fill the gap in ISR kinesic cue education, through the use of Kim's Game.

1.3 Instructional Strategy: Kim's Game

Initial testing for behavior cue detection training has leveraged virtual simulations that employ varying types of instructional strategies for recalling cues in a detection task, including Highlighting and Massed Exposure techniques (Salcedo 2014; Salcedo, Lackey, and Reinerman-Jones 2014; Lackey et al. 2013).

The Kim's Game strategy is borrowed from Kipling's (1901) book, *Kim*. In the book, the character Kim was shown a tray of items, which she studied, before the items were veiled from sight. Then, Kim had to describe each item and the amount of items shown, using memory and pattern recognition techniques. This observational puzzle has been extended for U.S. sniper training (U.S. Department of the Army 1994), coaxing a sniper to improve recall ability of item features through restructuring personal mechanisms of thinking. For behavior cue detection, Kim's Game allows the user to view target and non-target kinesic cues (showed through virtual human models performing the gestures), which are removed from sight to promote memorization of model features (Maraj 2015).

1.4 Purpose

The purpose of the presented research effort is to advance the fulfillment of kinesic cue detection training effectiveness. Participants whom completed an experiment implementing the Kim's Game strategy are analyzed in terms of their signal detection performance and subjective reactions towards the training testbed. Select correlations and regressions between the subjective subscales and performance, as well as between each of the scales, are critiqued.

2 MEASUREMENTS

Various survey measurements (i.e., scales and subscales) were given throughout the Kim's Game training. These measurements are introduced, here, to provide context.

2.1 Demographics

This research used a demographic questionnaire for questions related to age, sex, eyesight, military experience, educational experience, technology experience (i.e., computer and video game usage), current state of health, and current occupation.

2.2 Intrinsic Motivation Inventory

Intrinsic motivation is determined when someone participates in an activity because it is intrinsically interesting; there exists no reward or purpose outside the activity itself. Extrinsic motivation, in contrast, is characterized by motives oriented outside the activity, such as the incentive of fame, fortune, power, or avoiding punishment.

Deci (1975) covered intrinsic motivation in his Cognitive Evaluation Theory, and pointed out two aspects linked with rewards: a controlling aspect and an informational aspect. Controlling aspects, such as monetary rewards or due dates, are thought to raise extrinsic motivation; whereas informational aspects, such as positive feedback on progress towards a goal, are thought to raise a person's intrinsic motivation. The Intrinsic Motivation Inventory (IMI) (Ryan, Mims, and Koestner 1983) subjective self-report was formed in light of these principles, and supplemented free task persistence as a method for measuring intrinsic motivation. Free task time is performance related: participants that felt intrinsically motivated were more likely to continue playing an endless runner video game (Birk et al. in press). Thus, a goal of gaming and education is to foster intrinsic motivation, as a user will return to an activity repeatedly for the sake of experience (Gunter, Kenny, and Vick 2008, Konetes 2010). We expected that high stress levels would be correlated with being overchallenged, and thus connect to a drop in intrinsic motivation.

2.3 National Aeronautics and Space Administration – Task Load Index

Workload involves the amount of perceptual processing, including physical and cognitive demands that a user must afford for a task. The National Aeronautics and Space Administration – Task Load Index (NASA-TLX) uses subjective report to measure six main scales of perceived workload, which consist of three types of demand (Physical Demand, Temporal Demand, and Mental Demand) and three personal reflections on task performance (Performance, Effort, and Frustration) (Hart and Staveland 1988). These six aspects are averaged to provide a single measure of Global Workload. Mental, physical, and temporal demands relate to the workload factors of a specific task and are interpreted as the cognitive and physical resources required for task completion (Hart and Staveland 1988). Performance and effort are behavioral measures of workload that incorporate the participant's perceived effort on the success or failure of the task, and frustration is used to measure the participant's psychological agitation as related to workload factors of the specific task (Hart and Staveland 1988). Intersections of NASA-TLX with training include mapping aviator eye movements to phases of flight and workload (Di Nocera, Camilli, and Terenzi 2007), degrees of signal exposure in ISR tasks to workload (Abich, Taylor, and Reinerman-Jones 2013), and operation of a ground-assisted robot via job-aids with workload (Evans III, Hill, and Pomranky 2015).

In the current study, we ventured that users with higher immersive tendencies (especially those who are expected to lose track of time) will perceive a drop in Temporal Demand (i.e., the task-related pressure of time). As will be elaborated later, a connection between the Dundee Stress State Questionnaire and the NASA-TLX is also expected.

2.4 Dundee Stress State Questionnaire

Stress occurs as a reaction to internal or external stimuli, and a stressor (i.e., an inciter of stress that can signal an organism to adapt to an event) may stem from environmental or cognitive sources. The Dundee Stress State Questionnaire (DSSQ) measures a person's stress using a multidimensional view, where subjective self-report is collected to gain insight into three main factors of stress, including Engagement scales of motivation, concentration, and energy; Distress scales of confidence, low hedonic tone (or discontent mood), confidence-control, and tension; and Worry scales of self-esteem, cognitive interference (or distracting thoughts limiting task execution), and self-consciousness (Matthews et al. 1999; Matthews et al. 2002; Matthews et al. 2006). These scales equate to three domains: affect, cognition, and motivation (Matthews et al. 2002).

The DSSQ is revealing in terms of how tasks vary along state aspects, with vigilance tasks that are relatively boring showing a higher drop in engagement than a working memory task, and working memory tasks showing a high level of distress, but less worry (Matthews et al. 2002, Matthews et al. 2006). Here, strategies to improve performance may be inferred via DSSQ scores (e.g., a task associated with low engagement may require more challenge). Matthews et al. (2006) further demonstrated that tasks with the highest amount of distress were associated with the highest workload, with a vigilance and an anagram task both more demanding than a control group. Users practicing a room-clearing training task while situated in a virtual simulation were assessed with DSSQ, and a significant raise in distress and worry was correlated with task-related inexperience (Lackey et al. 2014). Per these experiments, we expect Kim's Game to show connection between DSSQ and NASA-TLX. The task involves novices performing signal detection over repeated trials (suggesting vigilance), in conjunction with a working memory component (due to remembering kinesic cues). The DSSQ results could show signs of dropped engagement (due to vigilance) and heightened distress (due to working memory) in terms of a compound task.

2.5 Technology Acceptance Model

For determining whether users will adopt a training technology, the Technology Acceptance Model (TAM) offers a glimpse into user perceptions. The TAM framework consists of two main, interdependent elements for predicting adoption behavior: perceived usefulness and perceived ease of use (i.e., usability) (Davis 1989). Traditionally, perceived usefulness is the extent to whether the adoption is expected to raise performance for a desired task (i.e., a goal-oriented purpose) and usability is the extent to whether using the technology will be free from effort (i.e., a process-oriented preference). Over the past few decades, TAM has been repurposed across disciplines, and has been subjected to meta-analyses for its professed value (Legris, Ingham, and Collerette 2003, King and He 2006). Comprehensively, the TAM is qualified as a general, acceptance-forecasting instrument via subjective report measures. At an early stage of technological design and development, the questionnaire involved with TAM can help anticipate downstream issues. Noted as a limitation, the TAM's simplicity has been cited as vague or inadequate at pinpointing particulars of what exactly needs to be fixed to further adoption (Sarker and Wells 2003). However, the TAM has been modified to create a more robust, illustrative view of adoption and continued usage, including expansion by importing elements of self-efficacy (Venkatesh 2000, Levy and Green 2009) trust (Reid and Levy 2008), and intrinsic motivation (Venkatesh 2000).

However, immersive tendencies have yet to have been combined with the TAM during ongoing test and evaluation for evidence of a successful behavior cue detection training application. A user that is more immersed may be expected to be at ease with using the technology.

2.6 Immersive Tendencies Questionnaire

Immersion is an effect whose concept is traced through conventional media to new media. Broadly, immersion occurs during some experience where a user has entered the frame of a synthetic world and has suspended their disbelief. Virtual reality allows for both mental and physical immersion (Craig and Sherman 2003). Immersion has been cited as a desirable trait for entertaining and serious gaming (Gunter, Kenny, and Vick 2008). Story, gameplay, and believability of visual presentation are credited as drivers of immersion (Lu et al. 2012; Adams 2004; Salen and Zimmerman 2003).

For the needs of virtual reality research, Slater, Usoh, and Steed (1995) proposed an objective view of immersion, where a raise in the degree of fidelity (i.e., the degree to how well a system duplicates the physicality and functionality of its intended environment) afforded by interactive display technologies will inevitably designate a raise in immersion. In contrast, Witmer and Singer (1998) support a subjective view of documenting immersion, where a user's psychological impression of being transported into a virtual environment is a key indicator, and where the Immersive Tendencies Questionnaire (ITQ) acts as an operationalization of the phenomena. Immersion and learning are expected to be related, as both rely on a "coherent set of stimuli" (Witmer and Singer 1998, p. 226) for a user to create meaning. However, immersion is only one tool, and does not guarantee transfer: prior evidence suggests immersion is best suited to a task that may rely on immersion benefits (Ortiz et al. 2013, Slater, Linakis et al. 1996). To determine if immersion is needed to improve the training technology, we would expect a system with a lack of ease of use to force a user's attention towards the lack of tool readiness, and trigger a break in immersion. Further, a useful technology would promote attraction to tool operation, raising involvement aspects of immersion.

3 METHOD

3.1 Participants

The participant recruitment process began by using a flyer containing general experiment information. This flyer was circulated through several media outlets within the University of Central Florida (UCF) community (i.e., through the UCF Psychology Department e-mail distribution, verbal means, and the Institute for Simulation and Training (IST) SONA system website). Interested participants were required to sign-up using an online participant system, SONA, through the Psychology Department or through IST. Each participant was offered either class credit or monetary compensation for their time. Those who chose monetary compensation were paid ten U.S dollars an hour, for a maximum of five hours. Those who chose class credit received one credit for every hour of participation, for a maximum of five hours.

Participant demographic analysis indicated a total of 75 participants (n = 75), consisting of 41 females and 34 males. The ages of participants ranged from 18 to 38, with an average age of 22.27 years (SD = 3.75). Each participant was screened using specific inclusion/exclusion criteria. Specifically, each participant was not able to participate in other experiments related to simulation training in virtual environments. Further, he or she must have been at least 18 years of age, have been a U.S. citizen, and have had normal or corrected normal vision (by glasses or contacts). Each participant, given that they met the inclusion and exclusion criteria, was permitted to participate in the study.

3.2 Research Effort

This study focused on the instructional strategy of Kim's Game. The experimental design culled survey data collection for the purpose of making inferences about a particular behavior or attribute found in the general population. Finally, conclusions were drawn from the survey data via statistical analyses.

3.3 Measurement Instruments

All the aforementioned measurements in the Measurements section were administered during the study session, in their original form (except for TAM, which was extended). Additionally, performance measures directly related to Kim's Game were broken down into three specific domains: participant detection accuracy, number of false positives reported, and response time. Detection accuracy was the percentage of correctly identified targets (i.e., virtual actors performing kinesic cues associated with either nervousness or aggressiveness). False positive detection was the number of non-target models depicting a behavior cue. Response time was the amount of time for a participant to react to an event that was depicted on the screen.

Performance was assessed in a pretest and posttest capacity between the Kim's Game condition and a control group.

3.4 Materials

The Kim's Game environment was created using the VBS3 virtual environment platform, which is employed by the U.S Army as the Program of Instruction for Soldiers. The task was presented on a 22-inch standard desktop computer with a 16:10 aspect ratio.

3.5 Procedure

Upon arrival to the designated laboratory location, each participant was checked by the experimenter for enrollment verification. The experimenter administered an informed consent to the participant. The experimenter and participant cosigned the informed consent to acknowledge a willingness to participate. The experimenter then asked a sequence of pre-experimental questions. Next, the participant completed a demographics questionnaire, followed by the DSSQ pretest, and the ITQ.

Each participant was presented with two interface training scenarios. The minimum score for both scenarios was predetermined as 75%. The first training scenario allowed for practice of the navigation and detection techniques to prepare the user for the experimental task. The practice content differed to prevent added bias for the Kim's Game task.

The second training scenario involved a task where the participant observed color changes between two consecutively presented screens. The participant was given a five-minute break after successfully completing the second scenario.

After the break, the experimenter presented a series of kinesic training slides. These training slides portrayed a human model exhibiting target behavior cues. When the slides were completed, the experimenter administered the practice vignette. The vignette allowed for additional practice of the kinesic training slides to prepare for the posttest scenario. The total time for this practice vignette was approximately 17 minutes. Next, the participant received the DSSQ posttest and the NASA-TLX. Upon completion of these questionnaires, the participant was given a five-minute break.

Following the break, the participant was shown a slide deck that detailed the requirements of the posttest scenario. The posttest scenario was subsequently administered, lasting approximately 40 minutes. When the scenario concluded, the participant completed the IMI and the TAM questionnaire. Performance measures were also collected and prepared for statistical analysis. The experimenter then debriefed and dismissed the participant. The actual time for completion of the procedure was approximately three hours.

4 RESULTS

4.1 Correlations

All of the data (i.e., questionnaires, surveys, and performance metrics) were analyzed through applying the Pearson product-moment correlation coefficient. The total number of participants (n = 32) remained the same for all measures. Four anomalous participants were removed before the final analysis, as their scores appeared as outliers consistently across different measures. Violations for the assumption of normality, linearity, and homoscedasticity were not encountered during investigation of the statistics. All correlations were rated based on Cohen (1988) conventions to determine effect size. The ITQ games subscale resulted in a non-normal data distribution and was consequently not analyzed.

Various significant correlations were found between the DSSQ and the IMI, the DSSQ and the TAM, and the TAM and the IMI (Tables 1, 2, and 3, respectively). Further, IMI showed connections to NASA-TLX and performance measures: a moderate, positive correlation appeared between IMI Pressure/Tension and NASA-TLX Temporal demand, $r = .464$, $p < .01$; a strong, positive correlation was found between IMI

Pressure/Tension and NASA-TLX Global Workload, $r = .566, p < .01$; and a moderate, negative correlation was found between IMI Perceived Competence and the performance measure of Pretest Median Response Time, $r = -.375.$, $p < .05$.

Select ITQ subscales were significantly correlated with other questionnaire subscales, as well: a strong, positive correlation was found between ITQ Involvement and NASA-TLX Temporal Demand, $r = .535$, $p < .01$.

Table 1: Correlations between DSSQ and IMI subscales for Kim's Game.

		DSSQ	
		Distress	Engagement
IMI	Perceived Competence	-.358*	-
	Pressure/Tension	.435*	-
	Interest/Enjoyment	-	.447*

*$p < .05$.

Table 2: Correlations between DSSQ and TAM subscales for Kim's Game.

		DSSQ	
		Distress	Engagement
TAM	Temporal Dissociation	-	.427*
	Focused Immersion	-	.504**
	Control	-	.420*
	Curiosity	-.355*	-
	Perceived Usefulness	-	.490**

*$p < .05$. **$p < .01$.

Table 3: Correlations between IMI and TAM subscales for Kim's Game.

		IMI		
		Effort/ Importance	Perceived Competence	Interest/ Enjoyment
TAM	Temporal Dissociation	-	-	.631**
	Focused Immersion	.598**	-	.691**
	Control	.364*	-	.391*
	Curiosity	.393*	.363*	.758**
	Perceived Ease of Use	-	.590**	-
	Perceived Usefulness	-	.500**	.651**

*$p < .05$. **$p < .01$.

4.2 Regressions

The data was analyzed using a standard multiple linear regression to predict if the subjective variables (surveys) influenced the participant performance (response time, detection accuracy, and number of false positives).

A standard multiple linear regression was conducted to predict TAM subscales on the Performance Pre Median Response Time. The results showed that the TAM subscales did not significantly predict Median Response Time (R^2 =.346, adjusted R^2 =.189, p =.076); however, the findings indicated that the TAM Curiosity subscale was the greatest contributor (β = -0.108, F (1, 32) = 2.205, p = .038).

A standard multiple linear regression was conducted to predict Demographics on Performance Post Median Response Time. This resulted in the Demographics (Computer Hours per Week) as a statistically significant predictor of Post Median Response Time (R^2 = .148, adjusted R^2 = .120, p = .030).

5 DISCUSSION

ITQ and NASA-TLX subscales of Involvement and Temporal Demand, respectively, indicate the existence of a strong positive relationship. One would expect that in analyzing Hart and Staveland's (1988) subscale of Temporal Demand, in which a user indicates time pressures on task pace, there exists a relation between perceived temporal demand and the effects of immersion. As the user needs to be fully focused on the task to achieve a high degree of involvement (Witmer and Singer 1998), perhaps more cognitive resources are involved when immersion arises, concordant with a rise in workload. The perception of time contracts as effort accelerates. Also, if a user was not involved with a task, they may in fact feel less constrained by workload due to allocating most of their cognitive resources outside the task itself.

A particular noteworthy relationship can be found between the IMI Interest/Enjoyment subscale and the TAM subscales (Table 3). Four of the TAM subscales (i.e., Perceived Usefulness, Temporal Dissociation, Focused Immersion, and Curiosity) had a strong positive relation to the Interest/Enjoyment subscale of the IMI. However, statistical significance was not found between the TAM Ease of Use subscale and the IMI Interest/Enjoyment subscale. Reeve (1989) found that interest and enjoyment are more effective when combined then when measured separately. Therefore, it is unnecessary to attempt separation of this subscale to investigate whether different statistical effects are found with the TAM Ease of Use subscale. Venkatesh (2000), indicates the Ease of Use subscale is rated according to the effort required of the individual. If the individual had more experience using the equipment (i.e., a computer) for the pattern recognition task, the required effort could become marginalized with respect to the task environment. If the task was very repetitive or uninteresting, the user would no longer be interested, and ease of use would be rendered irrelevant.

As expected, the DSSQ Engagement subscale and the TAM subscales were found to have significant results. When adopting Matthews' (2002) purpose for the DSSQ, the relationships could be explained by the moderate internal stress-free state the participant is experiencing. This moderate stress free-state is likely complimenting the relationship with engagement. Based on our results, it would seem that the less an individual is internally stressed from the task workload, the more likely the individual will be actively involved in the specific task. Also, the technological application that is not perceived as useful, despite the usability of the technology, may instill repulsion and indifference towards the task, and thus disengagement. Therefore, it is not odd to equate engagement with usefulness.

Other areas of results have enabled understandings of some of the subscale relationships. ITQ and TAM subscales were not significantly correlated, suggesting that effort on improving technology acceptance should not focus on improving immersion. It also seems logical that the DSSQ distress subscale would negatively affect TAM curiosity and IMI perceived competence. Negative emotional effects, such as internal distress, would likely inhibit the individual. Further, if a user is not given proper feedback towards current goal progress, they could lose confidence in how capable they feel to complete a task, leading to distress. Understandably, a meta-analysis of digital SBT applications found one of the highest factors for improved transfer and self-efficacy (a concept similar to intrinsic motivation) was a post-simulation assessment, as feedback (Gegenfurtner, Quesada-Pallares, and Knogler 2014).

A factor of adequate challenge has been attributed as a reason for intrinsic motivation (Malone 1981, Csikszentmihalyi 1991), with challenge level providing an indicator of task difficulty, boredom, or the need for meaningful feedback. For performance, users typically need to have perceived competence (i.e., feel confidence in one's capability) for completing an activity (Deci and Ryan 1980). This competency can be lowered if challenge is too high. We expected that high stress levels would be correlated with being overchallenged, and thus connect to intrinsic motivation. This relation was shown between the moderate negative relationship of Distress and Perceived Competence. Further, Interest/Enjoyment positively correlated with Engagement, helping vindicate that optimal experience, or a state of flow where challenge and abilities are in harmony (Csikszentmihalyi 1991), is related to workload and intrinsic motivation.

The lack of significant correlations between DSSQ and NASA-TLX ran counter to expectations, since these constructs appeared related to the working memory and vigilance aspects of the detection task, and since novices were training in the task domain. Although out of the scope of this paper, interest is directed to determine if a part of the compound task, or some other measure, confounds the expected relationship.

One regression showed that the level of computer hours a user accumulated during the week was a significant predictor of Post Median Response time. Although the effect of computer experience is not clear, one main presumption is the advantage of predisposition to a similar task. Higher computer experience may have given users an advantage for pattern recognition, due to a certain preexisting muscle memory. The study task may be compared to general computer use, where symbols are scanned, buttons are pressed, and the rules of a program must be followed. The training task is conceptually similar, giving those participants predisposed with computers a familiar process that could be quickly assimilated, especially after the rules for operation (i.e., training contents) were given.

6 CONCLUSION

For the Kim's Game strategy of kinesic cue detection training, multiple relationships were found between a variety of subjective questionnaires, as well as performance measures. For a more illustrative view of the training and prediction model, other SBT instructional strategies for cue detection should be analyzed and compared. Practical implications from the research include the introduction of task feedback during or after the simulation, which may promote flow; workload manipulation, or difficulty adjustment, to provide an optimal state of enjoyment; and less focus on instilling immersion for improving technology adoption, at least at an early stage of usage.

ACKNOWLEDGMENTS

This research was sponsored by the U.S. Army Research Laboratory – Human Research Engineering Directorate Advanced Training and Simulation Division (ARL - HRED ATSD), in collaboration with the Institute for Simulation and Training at the University of Central Florida. This work is supported in part by ARL HRED STTC contract W911NF-14-2-0021. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of ARL HRED STTC or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government.

REFERENCES

Abich, J., G. Taylor, and L. Reinerman-Jones. 2013. "Establishing Workload Manipulations Utilizing a Simulated Environment." In *VAMR/HCII 2013, Part II, LNCS 8022*, edited by R. Shumaker, 211-220. Heidelberg: Springer.

Adams, E. 2004. "The Designer's Notebook: Postmodernism and the 3 Types of Immersion." *Gamasutra.* http://www.gamasutra.com/view/feature/130531/the_designers_notebook_.php.

Baddeley, A. D., and G. Hitch. 1974. *Working Memory,* vol. 8, in *The Psychology of Learning and Motivation*, edited by G. H. Bower.

Birdwhistell, R. L. 1970. *Kinesics and Context: Essays on Body Motion Communication.* Philadelphia, PA: University of Pennsylvania Press.

Birk, M., C. Atkins, J. T. Bowey, and R. L. Mandryk. in press. "Fostering Intrinsic Motivation Through Avatar Identification in Digital Games." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'2016).* San Jose.

Cohen, J. W. 1988. *Statistical Power Analysis for the Behavioral Sciences.* 2nd ed. Hillsdale, New Jersey: L. Erlbaum Associates.

Craig, A. B., and W. R. Sherman. 2003. *Understanding Virtual Reality: Interface, Application, and Design.* San Francisco: Morgan Kaufmann.

Csikszentmihalyi, M. 1991. *Flow: The Psychology of Optimal Experience.* New York: Harper Perennial.

Davis, F. D. 1989. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology." *MIS Quarterly* 13 (3): 319-340.

Deci, E. L. 1975. *Intrinsic Motivation.* New York: Plenum.

Deci, E. L., and R. M. Ryan. 1980. "The Empirical Exploration of Intrinsic Motivational Processes." In *Advances in Experimental Psychology*, edited by L Berkowitz, 39-80. New York: Academic Press.

DeVito, J. A. 1968. "Kinesics: Other Codes, Other Channels." *Today's Speech* 16 (2): 29.

Di Nocera, F., M. Camilli, and M. Terenzi. 2007. "A Random Glance at the Flight Deck: Pilots' Scanning Strategies and the Real-Time Assessment of Mental Workload." *Journal of Cognitive Engineering and Decision Making* 1 (3): 271-285. doi:10.1518/155534307X255627

Evans III, A. W., S. G. Hill, and R. Pomranky. 2015. "Investigating the Usefulness of Soldier Aids for Autonomous Unmanned Ground Vehicles, Part 2." No. ARL-TR-7240. Aberdeen Proving Ground, MD: Army Research Laboratory.

Fautua, D., S. Schatz, D. Kobus, V. A. Spiker, W. Ross, J. H. Johnston, D. Nicholson, and E. A. Reitz. 2010. "Border Hunter Research Technical Report." ADA571115. Norfolk: U.S. Joint Forces Command.

Fischer, S. C., and J. Geiwetz. 1996. "Training Strategies for Tactical Pattern Recognition" No. ARI-TR-1031. Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences.

Gegenfurtner, A., C. Quesada-Pallares, and M. Knogler. 2014. "Digital Simulation-Based Training: A Meta-Analysis." *British Journal of Educational Technology* 45 (6): 1097-1114. doi: 10.1111/bjet.12188

Gunter, G. A., R. F. Kenny, and E. H. Vick. 2008. "Taking Educational Games Seriously: Using the RETAIN Model to Design Endogenous Fantasy Into Standalone Educational Games." *Educational Technology Research and Development* 56 (5/6): 511-537.

Hart, S. G., and L. E. Staveland. 1988. "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research." In *Human Mental Workload*, edited by P. A. Hancock and N. Meshkati, 139-184. Amsterdam: Elsevier Science Publishers.

King, W. R., and J. He. 2006. "A Meta-Analysis of the Technology Acceptance Model." *Information & Management* 43 (6): 740-755. doi:10.1016/j.im.2006.05.003

Kipling, R. 1901. *Kim.* London: Macmillan & Co.

Konetes, G. D. 2010. "The Function of Intrinsic and Extrinsic Motivation in Educational Virtual Games and Simulations." *Journal of Emerging Technologies in Web Intelligence* 2 (1): 23-26. doi:10.4304/jetwi.2.1.23-36

Lackey, S. J., J. N. Salcedo, G. Matthews, and D. B. Maxwell. 2014. "Virtual World Room Clearing: A Study in Training Effectiveness." In *Proceedings of the 2014 Interservice/Industry Training, Simulation & Education Conference (I/ITSEC).* Orlando, FL.

Lackey, S., C. Maraj, J. Salcedo, E. Ortiz, and I. Hudson. 2013. "Assessing Performance of Kinesic Cue Analysis in Simulation-Based Training Environments." In *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC 2013).* Orlando, FL.

Legris, P., J. Ingham, and P. Collerette. 2003. "Why Do People Use Information Technology? A Critical Review of the Technology Acceptance Model." *Information and Management* 40 (3): 191-204. doi:10.1016/S0378-7206(01)00143-4

Levy, Y., and B. D. Green. 2009. "An Empirical Study of Computer Self-Efficacy and the Technology Acceptance Model in the Military: A Case of a U.S. Navy Combat Information System." *Journal of Organizational and End User Computing* 21 (3): 1-23.

Lu, A. S., D. Thompson, J. Baranowski, R. Buday, and T. Baranowski. 2012. "Story Immersion in a Health Videogame for Childhood Obesity Prevention." *Games for Health: Research, Development, and Clinical Applications* 1 (1): 37-44.

Lyons, D. M., D. Schmorrow, J. C. Cohn, and S. J. Lackey. 2002. "Scenario Based Training With Virtual Technologies and Environments." In *Proceedings of the Image 2002 Conference.*

Malone, T. W. 1981. "What Makes Computer Games Fun?" *Byte* 6 (12): 258-277.

Maraj, C. 2015. “Investigating Simulation-Based Pattern Recognition Training for Behavior Cue Detection.” Ph.D. Thesis, College of Engineering and Computer Science, University of Central Florida, Orlando, Florida.

Martin, G. A., C. E. Hughes, S. Schatz, and D. Nicholson. 2010. "The Use of Functional L-systems for Scenario Generation in Serious Games." In *Proceedings of the 2010 Workshop: Procedural Content Generation in Games.* 1-5.

Matthews, G., A. K. Emo, G. Funke, M. Zeidner, R. D. Roberts, P. T., Jr. Costa, and R. Schulze. 2006. "Emotional Intelligence, Personality, and Task-Induced Stress." *Journal of Experimental Psychology* 12 (2): 96-107. doi:10.1037/1076-898X.12.2.96

Matthews, G., L. Joyner, K. Gilliland, S. E. Campbell, S. Falconer, and J. Huggins. 1999. "Validation of a Comprehensive Stress State Questionnaire: Towards a State ‘Big Three?’" In *Proceedings of the European Conference on Personality.* Tilburg University.

Matthews, G., S. E. Campbell, S. Falconer, L. A. Joyner, J. Huggins, K. Gilliland, R. Grier, and J. S. Warm. 2002. "Fundamental Dimensions of Subjective State in Performance Settings: Task Engagement, Distress, and Worry." *Emotion* 2 (4): 315-340. doi:10.1037/1528-3542.2.4.315

Ortiz, E., C. Maraj, J. Salcedo, S. Lackey, and I. Hudson. 2013. "Assessing Engagement in Simulation-Based Training." In *Virtual augmented and mixed reality 5th international conference, VAMR 2013, part of HCI International 2013, Las Vegas, NV, USA, July 21-26, 2013: Proceedings*, edited by R. Shumaker, 211-220. Heidelberg: Springer.

Peterson, L. R. 1966. "Short-Term Memory." *Scientific American.* http://www.indiana.edu/~p1013447/dictionary/stm.htm.

Reeve, J. 1989. "The Interest-Enjoyment Distinction in Intrinsic Motivation." *Motivation and Emotion* 83-103.

Reid, M., and Y. Levy. 2008. "Integrating Trust and Computer Self-Efficacy with TAM: An Empirical Assessment of Customers’ Acceptance of Banking Information Systems (BIS) in Jamaica." *Journal of Internet Banking and Commerce* 13 (3): 1-18.

Rensink, R. A. 2002. "Change Detection." *Annual Review of Psychology* 53 (1): 245-277.

Ryan, R. M., V. Mims, and R. Koestner. 1983. "Relation of Reward Contingency and Interpersonal Context to Intrinsic Motivation: A Review and Test Using Cognitive Evaluation Theory." *Journal of Personality and Social Psychology* 45: 736-750. doi:10.1037/0022-3514.45.4.736

Salcedo, J. N. 2014. “Instructional Strategies for Scenario-Based Training of Human Behavior Cue Detection and Classification With Robot-Aided Intelligence, Surveillance, And Reconnaissance.” Ph.D. Thesis, College of Engineering and Computer Science, University of Central Florida, Orlando, Florida.

Salcedo, J. N., S. J. Lackey, and L. Reinerman-Jones. 2014. "Massed Exposure Improves Response Time for Detection of Nervous Human Behaviors in Robot-Aided ISR Missions." *Proceedings of MODSIM World.* Hampton, VA.

Salen, K., and E. Zimmerman. 2003. *Rules of Play: Game Design Fundamentals.* Cambridge, Massachusetts: MIT.

Sarker, S., and J. D. Wells. 2003. "Understanding Mobile Handheld Device Use and Adoption" *Communications of the ACM* 46 (12): 35-40. doi: 10.1145/953460.953484

Slater, M., M. Usoh, and A. Steed. 1995. "Taking Steps: The Influence of a Walking Technique on Presence in Virtual Reality." *ACM Transactions On Computer-Human Interaction (TOCHI)* 2 (3): 201-219.

Slater, M., V. Linakis, M. Usoh, R. Kooper, and G. Street. 1996. "Immersion, Presence, and Performance in Virtual Environments: An Experiment With Tri-Dimensional Chess." *ACM Virtual Reality Software and Technology* 163-172.

U.S. Department of Defense. 2009. "FY2009-2034 Unmanned Systems Integrated Roadmap." ADA522247. Washington, DC: Department of Defense.

U.S. Department of the Army. 1994. "Sniper Training." FM 23-10. Washington, DC: U.S. Department of the Army.

Venkatesh, V. 2000. "Determinants of Perceived Ease of Use: Integrating Control, Intrinsic Motivation, and Emotion into the Technology Acceptance Model." *Information Systems Research* 11 (4): 342-365.

Wickens, C. D. 1992. *Engineering Psychology and Human Performance.* New York: HarperCollins.

Witmer, B., and M. Singer. 1998. "Measuring Presence in Virtual Environments: A Presence Questionnaire." *Presence* 7 (3): 225-240. doi:10.1162/105474698565686

AUTHOR BIOGRAPHIES

JONATHAN HURTER earned his master's degree in Modeling & Simulation (M&S) from the University of Central Florida (UCF) in 2015. Jon is a Research Technician for UCF's Institute for Simulation and Training (IST). His main interests lie in instructional design for simulations. His email address is jhurter@ist.ucf.edu.

WILLIAM AUBREY joined IST in the fall of 2015 as an Undergraduate Research Assistant. William is working toward a B.S. in Psychology, and a minor in Cognitive Sciences. William's ultimate goal is to complete a Ph.D. in Neuropsychology. His email address is waubrey@ist.ucf.edu.

SUSHUNOVA G. MARTINEZ earned her B.A in Mechanical Engineering at UCF in 2014, and is pursuing her M.A. in Industrial Engineering. She is interested in integrating the fields of mental health and emerging technology to create innovative designs for state-of-the-art simulations. Her email address is smartinez8@knights.ucf.edu.

CRYSTAL S. MARAJ attained her Ph.D. in M&S from UCF. Currently, Dr. Maraj is a Research Associate at UCF IST, and also works for the Army Research Lab (ARL) in Orlando, FL. She focuses on the training effectiveness of SBT applications through empirically-based research experiments. Her email address is cmaraj@ist.ucf.edu.

IRWIN HUDSON is the Science & Technology Manager at the U.S. ARL Advanced Training and Simulation Division; and the Cooperative Agreement Manager for the UCF Infrastructure Cooperative Agreement (CA), the Ground Robot Aided Intelligence, Surveillance & Reconnaissance CA, and the Evaluating and Informing Army Training CA. He is currently pursuing his Ph.D. in M&S from UCF. His email address is irwin.l.hudson.civ@mail.mil.