A Novel Approach to Determine Integrated Training Environment Effectiveness

LTC Glenn A. Hodges
Naval Postgraduate School
Monterey, California
glenn.a.hodges.mil@mail.mil

ABSTRACT

This paper discusses the development and use of an analytical assessment methodology anchored in systems engineering principles, affordance theory, and human abilities, to measure the potential of integrated training environments (ITE) to effectively support training. An integrated training environment is defined here as any human in-the-loop training system that includes live, virtual, constructive or game-based training aids, devices, simulators, or simulations (TADSS) alone or in combination, used to support the deliberate practice of skills for defined mission tasks. Empirical investigation of ITE is costly, lacks formal guidance, and is therefore often unreliable. Ad hoc studies, commissioned by individual organizations, constitute the current state of Army ITE evaluation. These assessments are often entirely based on subjective opinions gained through surveys, which produce results that are linked indirectly and loosely to the ITE. What is required is a repeatable, inexpensive, analytical approach to ITE assessment that bounds the potential of a given system to the support it provides to the deliberate practice of specific tasks. The results of this research include the development and use of the integrated training environment assessment methodology (ITEAM). ITEAM was used to evaluate the ability of several ITE to support the deliberate practice of specific tasks during training. During application, ITEAM consistently predicted where training was supported by an ITE and generally how well. ITEAM is offered as a tool to be used early in the material acquisition process to affordably define and verify the requirements of candidate ITE solutions for Department of Defense needs.

ABOUT THE AUTHOR

LTC GLENN A. HODGES U.S.A. is a recent graduate of the Modeling Virtual Environments and Simulation (MOVES) curriculum at the Naval Postgraduate School. During his career, he has served in a variety of command and staff positions within the U.S. Army. In 2001 he was selected for and designated as a Simulation Operations Officer (FA57). Before earning his Ph.D., LTC Hodges served as the Proponent Officer for the Army's Simulation Operations personnel functional area at the Center for Army Analysis (CAA). In this capacity he was responsible for the training and education of all officers designated as FA57's. LTC Hodges holds a Bachelor of Science degree in Business Administration from Old Dominion University, a Master of Science degree and Ph.D. in Modeling, Virtual Environments and Simulation (MOVES) from the Naval Postgraduate School. His interests include the application of systems engineering principles and processes to the development of modeling and simulation capabilities and the use of modeling and simulation to improve human performance. His current duty assignment is at the Army Capabilities Integration Center (ARCIC) working as the Future Technology Branch Chief in the Joint and Army Models and Simulations Division (JAMSD).

A Novel Approach to Determine Integrated Training Environment Effectiveness

LTC Glenn A. Hodges
Naval Postgraduate School
Monterey, California
glenn.a.hodges.mil@mail.mil

INTRODUCTION

The value of human-in-the-loop (HITL) simulation primarily comes from its ability to offer practice opportunities in environments that replicate important features of the real world (Salas, Rosen, Held, & Weissmuller, 2008). While this is true, at some point the focus of requirements determination, definition and solution development for military training systems shifted focus away from human performance and skill acquisition towards advanced technology. Operational and system requirements documents (ORD/SRD) have driven the increasing focus on the technological aspects of possible training solutions while marginalizing the importance of the front-end human analysis. This situation has resulted in the common practice of providing technical requirements specifications for training systems to defense contractors (the what) and then requiring the defense contractor to provide the government with a detailed explanation of how the training system will support the user (Klein, Johns, Perez, & Mirabella, 1985).

Between conflicts, the Armed Forces rely heavily on integrated training environments (ITE) to maintain warfighting skills. ITE are comprised of various live, virtual, constructive and game based training aids, devices, simulators and simulations, which allow Soldiers, Sailors, Airmen and Marines to practice the skills and engrain the knowledge necessary to execute their combat missions successfully on the battlefield (Hodges, Darken, & McCauley, 2014). ITE are extremely resource intensive and are rarely described as lightweight or turnkey. They require verification, validation and accreditation (VVA) just as their analytical counterparts that support budgetary and force structure decisions. A major difference between ITE that support training and other types of simulation is how they are evaluated.

The most common method of determining ITE effectiveness it through the use of empirical transfer of training (TOT) studies that are expensive and often provide limited or misleading insight into ITE utility. Some researchers have attempted to use non-empirical means to evaluate ITE in an effort to reduce costs and accurately capture positive system attributes (Tufano & Evans, 1982; Keesling et al., 1999; Sticha, Campbell, & Knerr, 2002, Gilligan, Elder, & Sticha, 1990). Despite their best efforts only a handful of researchers have had their techniques successfully implemented outside of the research arena and of those few have been used more than a handful of times. Most of the techniques developed have not been extensible, user friendly or well documented to facilitate reuse. Additionally, many have used mathematical equations that have not been validated with empirical data. Many have been automated due to their extreme complexity without concern for program documentation making them nearly impossible to implement by others and their focus has been similar to that of empirical attempts.

Until 2012, the Unites States Army (USA) had a system to provide analysis of training programs called the Training Effectiveness Analysis (TEA) system. The TEA system, established in 1975, was a Training and Doctrine Command (TRADOC) program focused on the impacts associated with training and hardware costs, hardware development cycles and complexity, training resources, and the overall effectiveness of Army programs to prepare Soldiers for battlefield conditions (Neal, 1982). Prior to 2012, the TRADOC Analysis Center at White Sands Missile Range was the Army's lead agency for providing technical assistance and conducting TEA for training systems. TRADOC Regulation 350-32 governed the TEA program. At the time, Simpson (1995) offered that the Army TEA system was the most robustly defined training analysis system that existed. Several system analysts have described the use of TEA studies for the benefit of their respective programs and offered examples of how they conducted TEA studies (Carter, 1982; Maitland, 1982). Despite this, in the summer of 2012, the USA officially concluded its last TEA study, eliminating both the office responsible for the conduct and oversight of TEA, and the regulation that governed the TEA system (Drillings, 2013).

HUMAN ABILITIES AND AFFORDANCES

Human ability (HA) research has been ongoing since the 1960's and has been used as a tool for empirical work investigating training system design and fidelity (Hays & Singer, 1988; Napoletano, 2013). Initially, the Defense Advanced Research Projects Agency (DARPA) sponsored research into HA to assist the military with job placement and training (Cockayne, 1998). The HA body of research has been developed as part of an umbrella taxonomic effort attempting to standardize the way human performance is described. The objective of the ability requirements approach was to identify and define the fewest number of independent ability categories that would be useful and meaningful for describing performance in the widest variety of tasks (Fleishman & Quaintance, 1984). HA development is an iterative process intended to produce a list of verified abilities that are empirically derived from patterns of responses to different tasks. The assumption is that specific tasks require certain abilities and that tasks requiring the same types of abilities can be categorized similarly. This assumption allows researchers to discuss task performance in relative terms. The HA project, through experimentation and collaboration with multiple subject matter experts, derived 52 HA with the possibility of adding more. Examples of HA are oral comprehension, deductive reasoning, dynamic strength, peripheral vision, and sound localization. HA are grouped into one of four categories (i.e., physical, sensory, psychomotor and cognitive). The United States Department of Labor uses HA as the basis for their O*NET (http://www.onetonline.org) program that provides information about jobs based on the HA needed to execute them.

Through years of research, Fleishman and his colleagues analyzed various jobs and tasks to ascertain and develop the list of 52 human abilities that can be found throughout various human activities. During this process, they executed numerous task analyses (TA). Through their process of defining ability requirements, they linked information dealing with task characteristics to HA (Fleishman & Mumford, 1991; Fleishman & Quaintance, 1984; Fleishman & Bartlett, 1969). The results of their efforts led to means of description, understanding and categorization of human activity (i.e., work) based on HA instead of through the use of TA. HA are viewed as enduring attributes of the human being (i.e., they are the same in the real or virtual world) and they play an important role in the methodology discussed here.

Affordance theory comes from ecological psychology and James J. Gibson. Gibson (1986) coined the term "affordance" to capture the essence of what an environment offers or provides an animal in either a positive or negative fashion. Affordance theory provides a context for discussing the qualities of the human-environment relationship within an ITE. Precedent exists for the use of affordance theory in supporting computer science and human factors research (Bærentsen & Trettvik, 2002; Chemero & Turvey, 2007; Lintern, 2000; Rome, Paletta, Şahin, Dorffner, Hertzberg, Breithaupt, Fritz et al., 2008). Affordance theory is naturally associated with HA, most notably with human perception. Gibson's theory of affordances has been met with varying degrees of enthusiasm and criticism over the years (Jones, 2003). As initially described, the concept of affordances was simple, clear and appealing (Michaels, 2003). However, Gibson's later attempts to describe affordances in more detail, resulted in a situation that "makes them seem like impossible, ghostly entities, entities that no respectable scientist (or science worshiping analytic philosopher) could have as part of their ontology" (Chemero, 2003, p. 182). Attempts at providing clarity and concrete definitions for affordances have been offered and debated (Stoffregen, 2003; Turvey, 1992).

Affordances are used in this research as a means of identifying the qualities and characteristics of the ITE that are absent or present in relation to the HA associated with specific tasks. We have elected to use affordances as part of our methodology because they provide context and allow us the opportunity to view an ITE unlike any other approach. Using affordances we are not only able to identify the characteristics of an ITE that support deliberate practice; but also why those identified characteristics are important to the trainee's execution of the tasks. Through the use of affordances, we are able to determine specific task elements with the highest likelihood of positive TOT.

INTEGRATED TRAINING ENVIRONMENT ASSESSMENT METHODOLOGY (ITEAM)

Figure 1 depicts the integrated training environment assessment methodology (ITEAM), a human-centered systems engineering approach to ITE analysis. ITEAM was developed based on the lessons learned from the literature and based on the recognition that front-end human analysis is critically important to training system development. Of the pieces of a training program (e.g. technology, requirements, humans) it has been established that computer

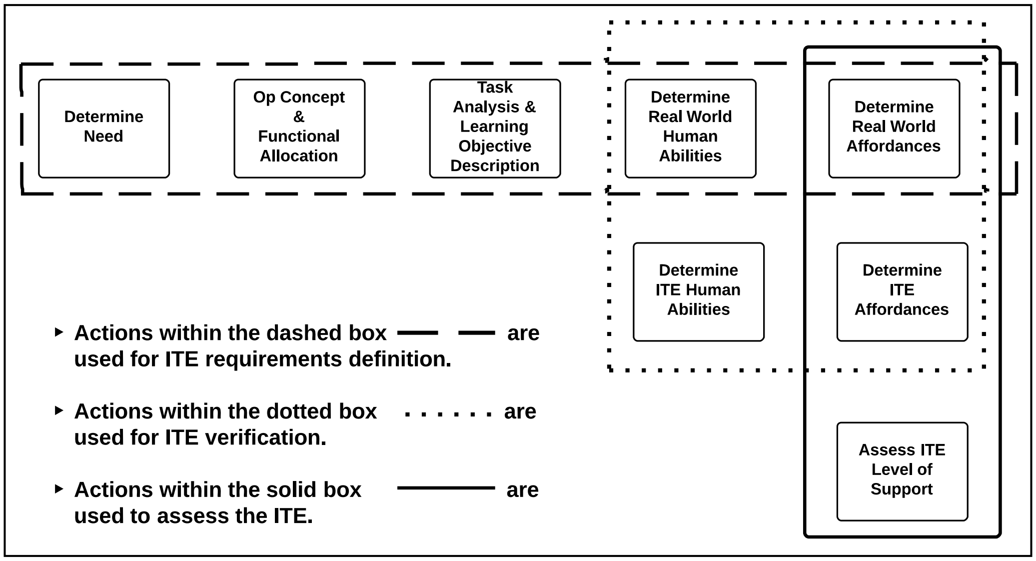

Figure 1. Integrated Training Environment Assessment Methodology

technology evolves the fastest (i.e. Moore's Law). Requirements determination occurs more slowly. Human beings evolve the slowest yet their evolutionary stability is ignored in ITE development in favor of an emphasis on advanced technology. ITEAM takes this into account by focusing on the support provided by an ITE to the deliberate practice of specific tasks and not on any specific technology. ITEAM was developed as a set of three main processes each containing multiple sub-processes. All of the sub-processes are iterative in nature and steps may be abbreviated or skipped depending on the time available and level of detail required. Requirements definition occurs first and proceeds from left to right beginning with determining the need and ending with determining the real world (RW) affordances. Verification follows and builds on requirements definition by determining the ITE HA and ITE affordances. Assessment of ITE support to training happens last and only after the RW and ITE affordances have been identified for comparison.

Requirements Definition

Proper problem description and analysis are critical to the ITE development process. ITEAM groups the activities of determining the need for the ITE, how it will be used, which functions will be performed by the ITE and the human, description of the tasks to be executed during training and the desired learning outcomes, within the boundary of requirements definition. Also included is a list of RW HA and RW affordance requirements that are necessary to accomplish the training tasks. HA are used to help illuminate the critical aspects (environmental affordances) required of the ITE. Affordances are used to describe the attributes of the ITE that are necessary to support the execution of the desired training.

Verification

Verification is defined as "the process of determining that a model or simulation implementation and its associated data accurately represent the developer's conceptual description and specifications" (Under Secretary of Defense, 2009). The sub-processes of ITEAM considered to be useful for verification consist of compiling the identified RW and system-supported HA as well as the RW and system-provided affordances. During this process, the evaluator uses the TA to determine the RW HA and affordance requirements associated with the tasks to be trained. Then, the ITE is investigated to determine what HA it supports and what affordances are available. Comparison of these items provides the basis for an initial judgment on whether or not the ITE will support the execution of the desired training.

Assessment

The final process of ITEAM assesses ITE ability to support desired training by quantifying ITE affordance resources based on ITE affordance requirements. The quantification of resources provides the customer/stakeholder/user with an estimate of the level of support that the ITE provides. ITE scoring is based on a subject matter expert (SME) evaluator's judgment on the absence or presence of specified affordances using the scale seen in Figure 2. The design of the scale was purposely skewed to ensure that ratings of excellent would not be common. The scale was set up so that the first two scoring levels (Poor and Fair) each contain 25 percent. The next two scoring levels (Good and Very Good) each contain an additional 20 percent and the final scoring level (Excellent) contains only 10 percent. Constructing the scale in this manner provides a progressive level of difficulty in reaching a rating of excellent, which requires that an ITE contain 90 percent or better of the affordances identified as being required to support the deliberate practice of specific tasks.

Subtask Affordance Scoring in Detail

Subtask Affordances are Unique. A unique affordance is one that has not been previously evaluated or accounted for as part of another subtask evaluation. If a subtask's affordances are unique, then a simple average of the number of affordances present divided by the total number required provides the percentage of affordances available for the subtask. This percentage is compared to the rating scale (Figure 2) and results in a rating of 1–5 Poor to Excellent.

Scale Definition

5–Excellent – the ITE contains all but a few (90–100%) of the affordances determined during the analysis
4–Very Good – the ITE contains a significant portion (70–89%) of the affordances determined during the analysis
3–Good – the ITE contains a good portion (50–69%) of the affordances determined during the analysis
2–Fair – the ITE contains some (25–49%) of the affordances determined during the analysis
1–Poor – the ITE contains very few (0–24%) of the affordances determined during the analysis

Figure 2. ITEAM scoring scale definition

Subtask Affordances are Previously Accounted for. If a subtask's affordances are completely accounted for in other analyses (referred to as affordance rollups), then those analyses are consulted and the ratings for their subtask affordances are obtained and averaged together to compile a numerical score for the current subtask under assessment. If more than one rollup is listed, then this process is executed for each of those rollups. Once all of the subtask affordance scores are collected they are summed and then divided by the total number of subtasks involved to obtain an average score. The average score now represents a number on the rating scale of 1–5 (see Figure 2). Raw scores containing 0.50 or less are rounded down to the nearest whole number for scoring purposes. Scores containing 0.51 or greater are rounded up.

Subtask Affordances are Partially Unique. If the affordances for a subtask are partially unique and partially accounted for in other analyses then the calculation is conducted in three steps. **Step one**—Treat each affordance rollup as an individual affordance that is present and unique. **Step two**—Evaluate and account for the presence of any unique affordances associated with the subtask. Once every affordance is accounted for, the calculation for determining the percentage present is conducted as described in (subtask affordances are unique). The result (rating of 1–5) is temporarily assigned as the subtask score. **Step three**—Obtain the values (scores) for the subtask affordances from the previous analyses (see subtask affordances previously accounted for) and sum them. Add the temporary value for the subtask currently under evaluation. Average this value by the total number of subtasks (including the current one). The derived number represents a number on the scale between 1 and 5 (see Figure 2) that when rounded appropriately (0.50 and lower round down) provides the qualitative rating for this subtask.

Subtask Affordances Contained in Multiple Analyses. In the case where a task's affordances are accounted for in multiple nested layers of sub-analyses, we have elected to stop the decomposition at the top of the second nested level. In such a case the top-level raw score of the high-level task at the second nested level is used in the value calculation for the current subtask. By our estimation, conducting further decomposition during the analysis leads to inflated results.

High-level Task Scoring

High-level tasks are also scored using the scale seen in Figure 2. The procedure to score a high-level task consists of summing all of the subtask scores and dividing them by the total number of subtasks. The result is a numerical value that is associated with a level of support provided (Poor to Excellent) by the ITE to the deliberate practice of the task.

APPLYING ITEAM TO THE RE-EVALUATION OF FULL SPECTRUM COMMAND

The remainder of this paper will discuss the use of ITEAM to evaluate the utility of the game Full Spectrum Command (FSC) to support the enhancement of adaptive decision-making abilities of officers attending the Infantry Captain's Career Course (ICCC) at Fort Benning, Georgia.

The Assessment in a Nutshell

The first step in our assessment of FSC was to obtain the original TEA study and sanitize it by removing any results, data, or analysis that might contaminate our assessment. The TEA was parsed to identify learning objectives, training task requirements and any other useful information. We reviewed the FSC user manual and spoke with the system developer at the Institute for Creative Technologies (ICT) and SMEs who were involved in the original TEA study to gain an understanding of the games capabilities and the intent for developing it. A TA was conducted on the mission and task data mined from the TEA to the level of detail necessary to support the identification of RW HA and RW affordances associated with the tasks. The RW HA were identified by reviewing the TA and the 52 HA from Fleishman and Quaintance (1984) and listing all HA believed to be associated with the tasks. Affordances necessary to stimulate the HA in support of the tasks were then described. A similar process was used to identify the system supported HA and system affordances. Examination of the user manual, focused game play and scenario development, discussions with developers and SMEs, all assisted in identifying the system supported HA and affordances. Once the RW and FSC HA and affordances were identified they were compared and scored. The resulting scores provide our prediction about how well FSC could be expected to support the deliberate practice of the tasks under investigation. For FSC this entire process took one person four months to complete. Readers interested in viewing the full analysis of FSC are encouraged to visit Hodges (2014).

Initial Assessment

Little was yielded from the FSC TEA to support the first two sub-processes of the ITEAM requirements definition process aside from a statement of intent to provide trainees with experience in developing plans and reacting to changes in their plans during contact (Beal & Christ, 2004). Our assessment gained traction at the point of conducting the TA. In the original TEA, subjects were asked to subjectively evaluate 22 separate action items using a survey. 19 of the 22 action items were used for our ITEAM TA. Three of the 22 original action items were not re-evaluated because they were viewed as redundant or unclear. Since FSC was intended for use at the company commander level, we reviewed a broad range of Army doctrine from the Army level down to the company level (i.e., FMs 3-0, 3-21.8, 3-21.10 and 5-0), including any doctrine specific to a warfighting function (e.g., breaching and engineers) to ensure that we had a good understanding of how infantry officers and their attached assets operate together in tactical, operational and strategic environments. The TA approach employed is one commonly used in the design of graphical user interfaces (GUI) (Gieskens & Foley 1991). It begins by identifying task pre-conditions, then describes and defines the tasks and ends by listing any post conditions from task execution. Given the complex nature of military operations we found that this approach works nicely.

The following provides a description of how ITEAM was used to assess FSC's ability to support the task of perform terrain analysis. A single subtask was identified for this high-level task: conduct a terrain analysis from a map and materials provided by the higher headquarters using the acronym OAKOC (Obstacles, Avenues of Approach, Key

Terrain, Observation and Fields of Fire, Cover and Concealment) as a guide. The pre-conditions required for this task were specified as (1) trainee has been issued a mission order and understands his area of operations and area of interest (2) trainee is familiar with troop leading procedures, mission analysis, and the mission, enemy, terrain, troops available, time, and civilian considerations (METT-TC) and (3) trainee has conducted steps 1 and 2 of the troop leading procedures. The desired post conditions were specified as (1) the terrain analysis answers questions about the terrain's effect on the operation and (2) the task results in a graphical display of the terrain (GDOT). Determining the RW HA associated with the task required that all 52 HA be reviewed and those believed to be associated with the task assigned. The following RW HA were listed as being required to conduct the task of perform terrain analysis: **Cognitive**—written comprehension, deductive reasoning, inductive reasoning, problem sensitivity, visualization, spatial orientation, and memorization; **Sensory**—near vision and visual color discrimination; **Psychomotor**—arm-hand steadiness, manual dexterity, and finger dexterity; **Physical**—None. The RW affordances corresponding to the HA and task were described as: (1) Paragraph 1 of an operations order that provides information from higher headquarters about the terrain and weather. (2) Maneuver Combined Obstacle Overlay (MCOO) from higher headquarters (3) Representation of the terrain that the trainee will maneuver over (e.g., map, aerial photographs) and (4) Intelligence information about enemy emplaced, natural or man-made obstacles known by the higher headquarters. The process for determining the system supported HA included investigating the scenario editor and developing and playing multiple game scenarios. System supported HA were identified as: **Cognitive**—None; **Sensory**—Near Vision and Visual Color Discrimination; **Psychomotor**—Manual Dexterity and Finger Dexterity; **Physical**—None. The resident system affordances were identified as (1) a scenario editor that allows for as much or as little detail as desired in paragraph 1 of the OPORD (2) Scenario editor allows maps and simulated photography to be provided in the scenario and (3) No MCOO functionality exists in the game. All of the affordances for this particular subtask were unique and of the four identified as being required, three were determined to be present. FSC's ability to support the practice of this task was rated as very good based on the presence of 75 percent of the affordances required. Figure 3 graphically depicts all of the results of the initial analysis of FSC using ITEAM.

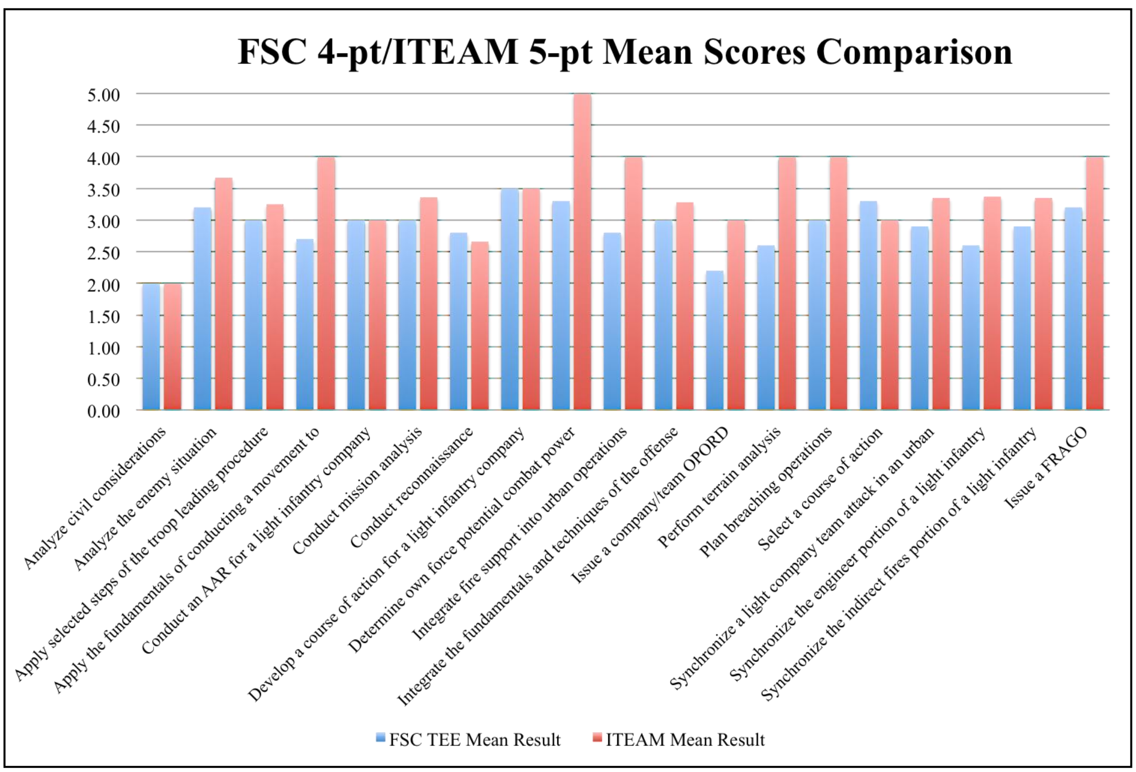

Figure 3. Initial FSC/ITEAM analysis comparison

After conducting this same procedure on each of the 19 action items we recognized that the ITEAM values of FSC support tended in some cases to be significantly high as can be seen in Figure 3. One possible explanation for this was the difference in scales used in the original TEA and by ITEAM. The TEA used a four-point scale where 1 = Not at all; 2 = Not very well; 3 = Moderately well; and 4 = Very well whereas ITEAM uses the scale seen in Figure 2. To determine if the difference in scales affected the ratings, the scores of each of the 19 action items were recalculated using a redesigned four-point scale seen in Figure 4.

Scale Definition

4–Very Good – the ITE contains a significant portion (75–100%) of the affordances determined during the analysis
3–Good – the ITE contains a good portion (50–74%) of the affordances determined during the analysis
2–Fair – the ITE contains some (25–49%) of the affordances determined during the analysis
1–Poor – the ITE contains very few (0–24%) of the affordances determined during the analysis

Figure 4. Redesigned four-point ITEAM scoring scale

The results of the recalculation of raw mean scores are graphically depicted in Figure 5. The use of the four-point scale did lower the ITEAM results in some cases but we still did not have full confidence that our comparison values were similar. One final adjustment to the scoring analysis finally allowed us to fairly (in our minds) compare the

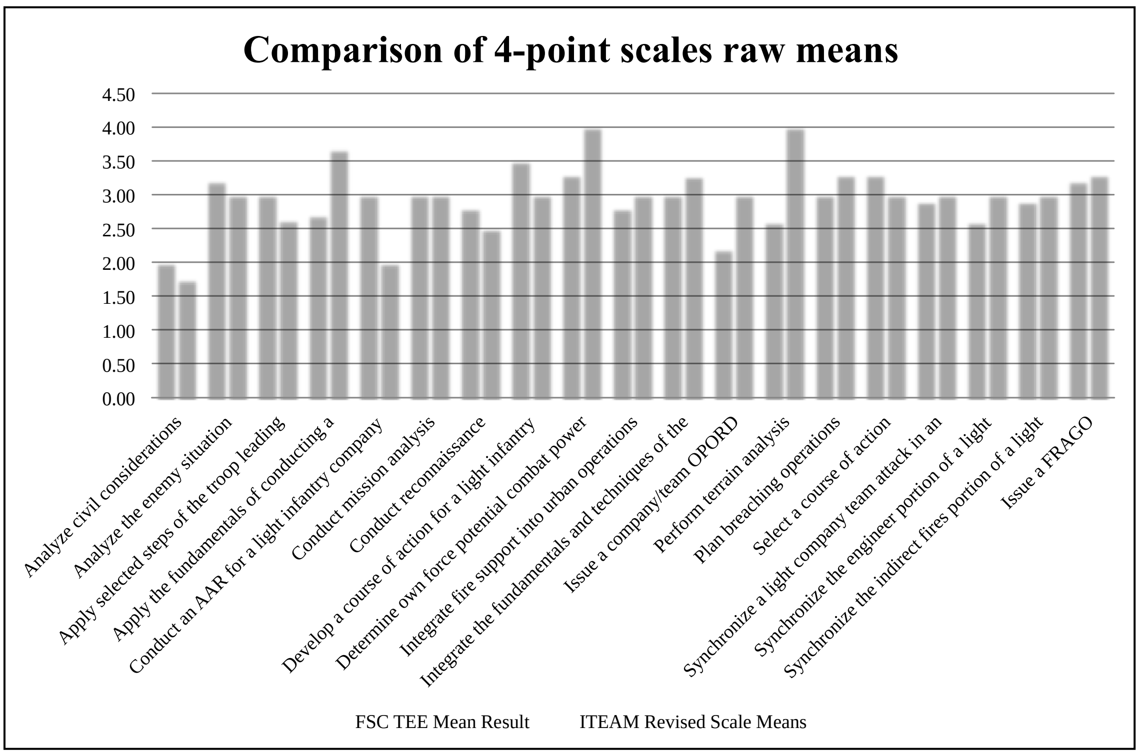

Figure 5. Four-point scale raw mean comparison

TEA results with the FSC student assessments of the games capability. The final adjustment consisted of rounding both TEA and ITEAM results similarly (scores containing .51 or higher rounded up) to produce values that were more easily comparable. The results of the rounding and the final comparison scores may be seen in Figure 6.

DISCUSSION

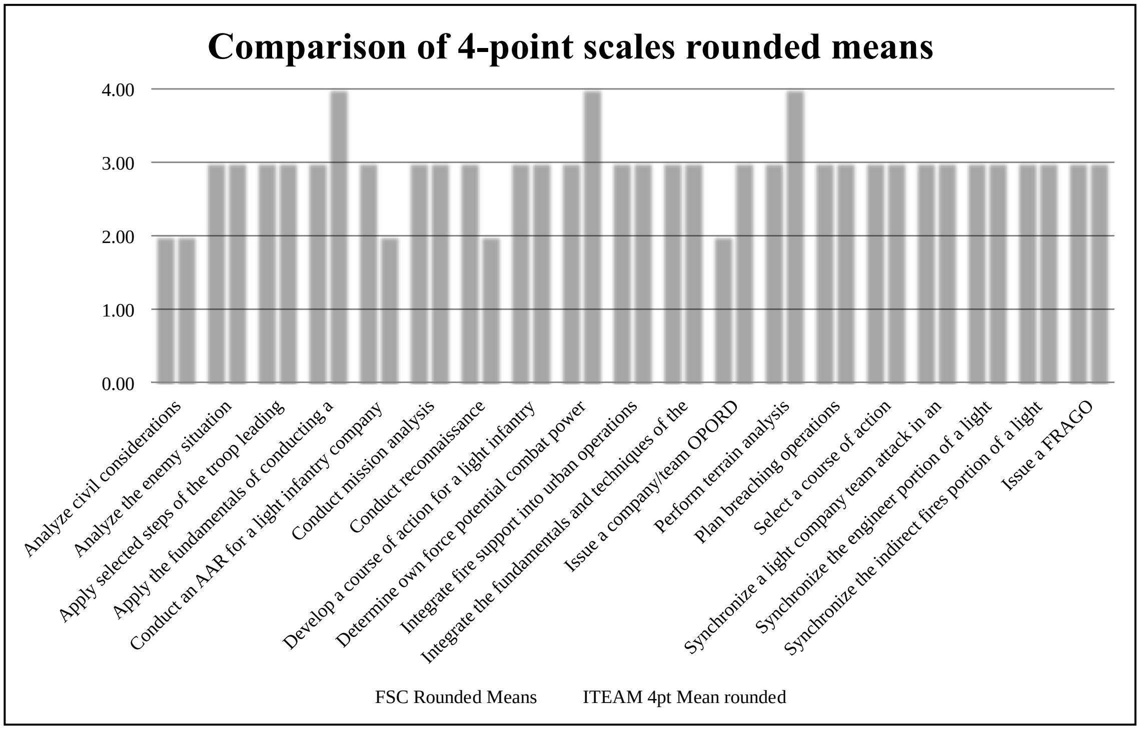

Figure 6. Final comparison of scored between FSC TEA and ITEAM

The original FSC TEA collected data from the participants involved in the study through the use of various survey items. An empirical test using a robust simulation was designed and executed to collect data pertaining to student performance. Considerable statistical analysis was executed on the data collected to draw conclusions. In the end, as the TEA states, the results were inconclusive, meaning evaluators were unable to determine the training effectiveness of FSC. Reasons given for this were the result of unforeseen obstacles to the empirical protocol as well as the lack of any identified improvement of decision-making by the trainees. We postulate that the result was due to the lack of focus on FSC and too much focus on Soldier performance. While improvement in performance is a good measurement to indicate the value of an ITE, it only works if the capabilities of the system are known beforehand and applied appropriately.

As we investigated FSC and applied our analytical methodology, several things came to light about the original TEA. First, FSC was never the main focus of the study; rather it was an enabler to a larger investigation into adaptive decision-making. Second, opinions of users who were unfamiliar with the full capability of FSC, instructional systems design, the systems approach to training (ISD/SAT), and the tasks, were relied upon to judge the utility of FSC. These same users based their opinions on scenarios that we believe were incomplete, possibly by design. So instead of providing a full picture of the capabilities of FSC and requesting feedback, users were only exposed to some of the capability of FSC and then asked to offer their opinions about the games effectiveness.

During our initial review of FSC, we felt that the game was not suitable for military training at all (evaluator bias). However, after we conducted our analysis using ITEAM, that opinion was drastically changed. Merely passing judgment based on opinions is not as effective or reliable as basing opinions on structured analysis. Users were never asked whether FSC supported the accomplishment of any specific tasks. They were asked whether they felt that the game allowed them to conduct certain action items. If asked to decompose those actions and then consider the human abilities needed to conduct them, we believe that students would have adjusted their assessments.

CONCLUSIONS

ITEAM was developed based on the information and lessons learned from the literature and a systematic approach to problem solving taught within the systems engineering discipline. Several lessons were learned during this process about ITEAM and its application. First, domain experience is necessary in order to use ITEAM effectively. Using properly focused SMEs with domain expertise strengthens the validity and reliability of the ITEAM data and reduces the risk of a false positive with respect to ITE capability. Without domain knowledge and experience, an essential understanding of the necessary ITE affordances does not exist and cannot be determined appropriately. The consistent application of ITEAM mitigates the effects of evaluator bias. With FSC, we recognized that in at least one instance we treated two similar affordances in two different analyses inconsistently. This resulted in an inflated rating for one of the analyses and a conflict. Taking the time to draft a study plan, rules for ITE examination and the handling of unique and similar situations all help to mitigate SME bias and strengthen the reliability of the ITEAM results. An approach that focuses attention on the tasks needing practice, the HA involved with those tasks, and the necessary ITE affordances, mitigates an overemphasis on the application of advanced technology in the design, development and implementation of ITEs. Taking time to determine the true need for an ITE and a concept for human/ITE interaction is time well spent.

This research focused on the development of a methodology that can be used to determine ITE utility to support the deliberate practice of specific tasks. We demonstrated the reliability and consistency of the process using three distinct case studies employing three different ITEs. We did not attempt to investigate or measure the inter-rater reliability (IRR) of using the methodology due to time constraints and the belief that consistency and reliability in the application of the methodology should come first. We recognize that establishing IRR for ITEAM is of foremost importance and a logical next step in ITEAMs evolution.

The debate over the value of analytical assessment of ITEs will continue but we believe that our efforts have shed light on a new way to approach the issue. Implementing a methodical process in assessment efforts forces an accounting of things that the current acquisition process ignores or bypasses. Each process and sub-process of ITEAM unlocks information about the stakeholder’s needs and ITE requirements that otherwise might be missed if the methodology was not followed. Furthermore, given that the cost of using the methodology is so small, it will result in savings of time and money in the areas of design, development and manufacturing of ITEs.

ACKNOWLEDGEMENTS

I would like to thank Rudolph Darken, Ph.D., Michael McCauley, Ph.D., Eugene Paulo, Ph.D., Jeffrey Appleget, Ph.D, and CDR Joseph Sullivan, Ph.D., for their support on this research effort.

REFERENCES

Bærentsen, K., & Trettvik, J. (2002). An activity theory approach to affordance. In *Proceedings of the second Nordic conference of Human Computer Interaction* (pp. 51–60). New York: Association for Computing Machinery.

Beal, S., & Christ, R. (2004). *Training effectiveness evaluation of the full spectrum command game* (TR 1140). FT Benning: U. S. Army Research Institute for the Behavioral and Social Sciences.

Carter, R. J. (1982). Methodologies for evaluating training products and processes. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 26,* (pp. 258–262).

Chemero, A. (2003). An outline of a theory of affordances. *Ecological Psychology*, *15*, 181–195. doi:10.1207/S15326969ECO1502_5

Chemero, A., & Turvey, M. (2007). Gibsonian affordances for roboticists. *Adaptive Behavior*, *15,* 473–480. doi:10.1177/1059712307085098

Cockayne, W. (1998). *Two-handed, whole-hand interaction* (Masters thesis, Naval Postgraduate School). Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA356098

Drillings, M., (2013, Summertime). Director's corner. *MANPRINT*. Retrieved from http://www.manprint.army.mil

Fleishman, E., & Mumford, M. (1991). Evaluating classifications of job behavior: A construct validation of the ability requirement scales. *Personnel Psychology, 44,* 523–575. doi:10.1111/j.1744-6570.1991.tb02403.x

Fleishman, E., & Quaintance, M. (1984). *Taxonomies of human performance: The description of human tasks*. Orlando, FL: Academic Press.

Fleishman, E., & Bartlett, C. (1969). Human abilities. *Annual Review of Psychology*, *20*, 349 to eoa.

Gibson, James. J. (1986). The theory of affordances. In *The ecological approach to visual perception* (pp. 127–143). Hillsdale: Lawrence Erlbaum Associates, Inc.

Gieskens, D., & Foley, J. (1991). Controlling user interface objects through pre- and postconditions. In *Proceedings of SIGCHI Conference on Human Factors in Computing Systems* (pp. 189–194). New York: Computer Human Interaction (CHI)

Gilligan, E., Elder, B., & Sticha, P. (1990). *Optimization of simulation-based training systems: User's guide* (RN 91-04). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences.

Hays, R., & Singer, M. (1983). Research issues in training device design. In *Proceedings of the Human Factors Society 27th Annual Meeting* (pp. 147–150).

Hodges, G. (2014). *Identifying the Limits of an Integrated Training Environment using Human Abilities and Affordance Theory*. (Doctoral dissertation, Naval Postgraduate School).

Hodges, G., Darken, R., & McCauley, M. (2014). An analytical method for assessing the effectiveness of human in the loop simulation environments : A work in progress. In *Proceedings of the 2014 Spring Simulation Multi-conference*. Tampa: The Society for Modeling and Simulation International.

Jones, K. (2003). What is an affordance? *Ecological Psychology*, *15*, 107–114. doi:10.1207/S15326969ECO1502_1

Keesling, J., King, J., & Mullen, W. (1999). *Simulation training strategies for force XXI final technical report* (Report 2000-02). Alexandria: U. S. Army Research Institute for the Behavioral and Social Sciences. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA370982

Klein, G., Johns, P., Perez, R., & Mirabella, A. (1985). *Comparison-based prediction of cost and effectiveness of training devices: A guidebook* (Report 85-29). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA170941

Lintern, G. (2000). An affordance-based perspective on human-machine interface design. *Ecological Psychology*, *12*, 65–69. doi:10.1207/S15326969ECO1201_7

Maitland, A (1982). Training effectiveness analysis: Where the operator meets the equipment. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 26, (pp. 255–257). doi:10.1177/154193128202600316

Michaels, C. (2003). Affordances: Four points of debate. Ecological Psychology, 15, 135–148. doi:10.1207/S15326969ECO1502_3

Napoletano, N. (2013). The eyes have it: Simulated sound visualization for testing. In Proceedings of the Interservice/Industry Training, Simulation & Education Conference. Orlando, FL: NTSA. Retrieved from http://www.iitsec.org/about/PublicationsProceedings/Pages/BestPapers.aspx

Neal, G. (1982). Overview of training effectiveness analysis. In Proceedings of the Human Factors and Ergonomics Society 26th Annual Meeting (pp. 244–248). doi:10.1177/154193128202600313

Rome, E., Paletta, L., Sahin, E., Dorffner, G., Hertzberg, J., Breithaupt, R., Fritz, G.,...& Uğur, E. (2008). The MACS project: An approach to affordance-inspired robot control. In E. Rome, J. Hertzberg & G. Dorffner (Eds.) *Towards affordance-based robot control* (pp. 173– 210). Berlin: Springer

Salas, E., Rosen, M., Held, J., & Weissmuller, J. (2008). Performance measurement in simulation-based training: A review and best practices. *Simulation & Gaming*, *40*, 328–376. doi:10.1177/1046878108326734

Simpson, H. (1995). *Cost-effectiveness analysis of training in the department of defense.* (TR 95-004). Monterey, CA: Defense Manpower Data Center. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA302985

Sticha, P., Campbell, R., & Knerr, M. (2002). *Individual and collective training in live, virtual and constructive environments. Training concepts for virtual environments* (SR 2002-05). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA402315

Stoffregen, T. (2003). Affordances as properties of the animal-environment system. *Ecological Psychology*, *15*, 115–134. doi:10.1207/S15326969ECO1502_2

Tufano, D., & Evans, R. (1982). *The prediction of training device effectiveness : A review of Army models* (TR 613). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA146937

Turvey, M. T. (1992). Affordances and prospective control: An outline of the ontology. *Ecological Psychology*, *4*, 173–187. Retrieved from http://www.tandfonline.com/doi/abs/10.1207/s15326969eco0403_3

Under Secretary of Defense. (2009). Modeling and simulation (M&S) verification, validation, and accreditation (VV&A) instruction (DoD Instruction 5000-61). Washington, DC: Author. Retrieved from http://www.dtic.mil/whs/directives/corres/pdf/500061p.pdf