Undergraduate Boot Camp: Getting Experimental Populations Up To Speed

Roberto Champney, Laura M. Milham, Meredith Bell Carroll, Kay M. Stanney, David Jones
Design Interactive, Inc.
Oviedo, FL
Kay@designinteractive.net, Roberto@designinteractive.net, Laura@designinteractive.net, Meredith@designinteractive.net, David@designinteractive.net

KC Pfluger
Trestles, Inc.
Washington D.C.
kc@trestlesinc.com

Joseph Cohn
Naval Research Laboratory
Washington D.C.
Cohn@itd.nrl.navy.mil

ABSTRACT

A struggle common across many researchers is the definition of a process to systematically train undergraduates to a degree of competence on experimental tasks. This process becomes more critical when paired with the notion that much data resultant from academic research is utilized to either generalize or apply findings to a much larger population. Many researchers have criticized the practice of using undergraduates as experimental participants, claiming that there are characteristics of the undergraduate population which are different than those found in the general population (Ward, 1993). This is of particular concern in military training research in which the goal is to generalize experimental results to military populations. Further, in these settings, it is challenging to make conclusive recommendations regarding training effectiveness of systems when the experimental population may not have the requisite knowledge, skills and attitudes (KSAs) necessary for task accomplishment, or an understanding of the requisite tasks and the motivations behind mission success.

This paper presents a method designed to increase the KSAs of undergraduate participants such that they more closely parallel the target population; this is accomplished through the design and implementation of an "Undergraduate Boot Camp". This methodology was applied within the Military Operations on Urbanized Terrain (MOUT) domain, and focused on providing both classroom and practical application training in which a subject matter expert (SME) instructed, demonstrated and evaluated trainee performance. As a next step, training was supplemented by detailed scenario based feedback targeted at developing self and team awareness of task accomplishment. A key aspect of the methodology is to develop a central set of competencies, incorporate them into a pre training advance organizer, developed to support rapid development of a mental model of the target domain. Given this approach, the goal is to begin to define a methodology for systematically providing quality pre experimental instruction to trainees, resulting in an experimental group that has basic knowledge, skills, and strategies for interacting in the targeted domain.

ABOUT THE AUTHORS

Kay Stanney, Ph.D. is president of Design Interactive, Inc. and Professor and Trustee Chair with UCF's Industrial Engineering & Management Systems Department, where she joined in 1992. She conducts research in the areas of multimodal interaction, human-systems integration, and augmented cognition. She is currently conducting user-centered design research for ONR's Virtual Technologies and Environments Program.

Roberto Champney is a Senior Research Associate at Design Interactive and has significant experience in systems design and evaluation having performed several system evaluations in industry and academia. Roberto holds an MS Human Engineering and Ergonomics, and is an ABD PhD candidate at the University of Central Florida.

Laura Milham, Ph.D. is the Director of Training Systems at Design Interactive, Inc. She has served as a Lead Scientist within the Navy's VIRTE program, designing the human computer interaction aspects of building VE training systems, designed transfer of training experiments, and training evaluation experiments.

Meredith Bell Carroll is a Senior Research Associate at Design Interactive, Inc. and is currently supporting virtual training system design and evaluation for the Office of Naval Research's Virtual Technologies and Environments (VIRTE) Program. Her research efforts have been directed towards human/team performance and training in complex systems in aviation and military domains, with focuses on performance measurement and virtual training technology. She holds a B.S in Aerospace Engineering from the University of Virginia, an M.S. in Aviation Science from Florida Institute of Technology, and is currently working toward her doctorate in Human Factors and Experimental Psychology at the University of Central Florida.

David L. Jones is a Research Associate at Design Interactive, Inc. His primary research focus is on designing multimodal interactive systems. He has a Masters in Industrial Engineering from the University of Central Florida, with a focus on human-computer interaction (HCI) and usability. He has a Bachelors in Human Factors Psychology from Embry-Riddle Aeronautical University, with a focus on human factors and computer science.

LCDR Joseph Cohn is an Aerospace Experimental Psychologist in the U.S. Navy's Medical Service Corps. Dr. Cohn received his commission in 1999 as a Lieutenant and completed naval flight training in February 2000. He has overseen the validation and transition of three combat simulators to the Fleet and Force and was the Founder and Director of the Naval Research Lab's War fighter Human Systems Integration Laboratory.

K.C. Pfluger is the President of Trestles Inc. and a Marine Reservist. He serves as subject matter and technical expert on USMC tactics, techniques, procedures, and training for infantry units. He is a graduate of the Naval Academy and has 10 years experience as an Infantry Officer. He has spent the last 7 years as a Marine Liaison to the Naval Research Laboratory and industry, supporting developing technologies for the Fleet Marine Force.

Undergraduate Boot Camp: Getting Experimental Populations Up To Speed

Kay M. Stanney, Roberto Champney, Laura M. Milham, Meredith Bell Carroll, David Jones
Design Interactive, Inc.
Oviedo, FL
Kay@designinteractive.net, Roberto@designinteractive.net, Laura@designinteractive.net, Meredith@designinteractive.net, David@designinteractive.net

KC Pfluger
Trestles, Inc.
Washington D.C.
kc@trestlesinc.com

Joseph Cohn
Naval Research Laboratory
Washington D.C.
Cohn@itd.nrl.navy.mil

INTRODUCTION

A controversial issue that has been the center of much debate in the past revolves around using undergraduates in laboratory research and the generalizability of these research findings to other populations (Wintre, North & Sugar, 2001). Many researchers have criticized the practice of using undergraduates as experimental participants, claiming that there are characteristics of the undergraduate population which are different than those found in the general population (Ward, 1993) and therefore findings with respect to undergraduates are not necessarily true for other populations. This is of particular concern in military training research in which the goal is to generalize experimental results to military populations, populations which usually have very specifically trained skills and knowledge. In these instances, it is challenging to make conclusive recommendations regarding training effectiveness of systems when the experimental population may not have the requisite knowledge, skills and attitudes (KSAs) necessary for task accomplishment, or an understanding of the essential tasks and the motivations behind mission success.

It is proposed herein that there is a feasible and cost effective means of addressing this issue without resorting to the use of actual military personnel who many times have limited availability and whose incorporation into a study can impose stringent limitations. This paper presents a method that identifies the task constructs and context that are used to increase the KSAs of undergraduate participants such that they more closely parallel the target population; this was accomplished through the design and implementation of an "Undergraduate Boot Camp." The goal of this method is to begin to systematically provide quality pre-experimental instruction to trainees, resulting in an experimental group that has basic knowledge, skills, and strategies for interacting in the targeted domain.

Problem

Applied human factors is inevitably called upon to produce effective solutions for complex operational issues, while being limited by such resource constraints as restricted access to relevant samples (Aidman, Galanis, Manton, Vozzo and Bonner, 2002). Most studies strive to draw conclusions about specific theories or constructs, with the end goal of generalizing these findings to a particular domain. Undergraduate students have been the primary population utilized as experimental participants due to availability.

The populations to which researchers have attempted to generalize research findings based on undergraduate participants are countless, ranging from upper class business executives, to military personnel, to minimum wage blue collar workers. Over the past several decades, there has been a great deal of criticism of this practice. Gordon, Slade, & Schmitt (1986) cite Gloria Gordon as emphasizing that the behavioral sciences are one of the few sciences that do not take great pains to demonstrate the experimental sample representativeness of target populations. In the biological sciences, for instance, species selection for experimental drug testing involves careful consideration with respect to similarity to humans in relevant metabolic processes and chemical tolerance (Gordon et al., 1986). Researchers in the behavioral sciences may target a specific subset of the undergraduate population (e.g., males vs. females), but it is rare that researchers take great efforts to ensure participants are similar to the target population with respect to relevant behavioral characteristics. Some scientists allege that this lack of representativeness leads to ungeneralizable findings (Gordon, et al., 1986). Gordon, et al. (1986) analyzed several studies in which students and non-students participated as subjects under the same conditions. The researchers found there to be statistical differences between the two groups and that despite this, no major difference between the two groups were reported in the majority of the studies. Several researchers have criticized this study claiming that the evidence presented does not support calls for practitioners to be weary of studies involving undergraduate participants (Greenberg, 1987; Dobbins, Lane & Steiner, 1988). For example, Greenberg (1987) argued that representative samples do not necessarily ensure generalizability and that there is a place for undergraduate research in understanding specific constructs. Greenberg (1987) also argued that differences between sample groups does not necessarily mean non-generalizability, but may constitute valuable boundary conditions for the construct of study.

There are many researchers who herald the benefits of laboratory studies. Ward (1993) found some behavioral constructs generalizable between undergraduates and working adults. Greenberg (1987) highlighted that generalizability does not result from one study regardless of the participant sample, but that each study has the potential to provide valuable insight. "Good laboratory research can provide us with an essential element in external validity: an understanding of the processes which underlie behavior in work settings." (Dobbins, Lane, & Steiner, 1988, p. 282).

Another issue of concern is that of task meaningfulness. For instance, Stewart (2006) found a positive relationship between task meaningfulness and team performance. Similarly Morris, Hancock and Shirkey (2004) found significantly higher stress-indicator scores for a "Context Enhanced" pre-training condition versus a low context condition in a military task training experiment. Of interest in this last study, is the fact that while their group who received an emotively filled pre-training video ("Context Enhanced" group) scored higher in their mission completion measure, although they did not differ in their learned doctrine measures. This is of importance because this latter measure is one associated with specific training content, which as suggested by the results, was not influenced by enhancing the context alone. As the authors describe it "Because initially there were no significant differences between the groups on items such as individual's gaming habits or initial stress, an important conclusion is that the experimental manipulation (stress context-relevant materials) produce more 'motivation to succeed' in game training, but may not affect specific skill acquisition per se" (Morris et al., 2004, p 144). Thus the motivational impact alone of such interactive training is insufficient to produce increased learning (Ricci, Salas, Cannon-Bowers, 1996). This is very similar to the ad-hoc observation of Stanney, et. al, (study in progress) during early pilot sessions of test environment described herein. These authors observed how participants would display competitive behaviors that led to "good" mission scoring, but disregarded the doctrine for adequate execution (e.g. safety, techniques, risk management). This behavior calls into question the external validity of the results observed in this type of system assessment and strive to design studies which compensate for participants' lack of knowledge and skills which are believed to facilitate learning of experimental tasks.

As is evident by the arguments presented above, there are unresolved concerns with respect to the generalizability of undergraduate research findings. Gordon, Schmitt, and Schneider (1984) propose that the generalizability of results obtained with undergraduate research participants are confounded by several experimental aspects in addition to population difference, including artificiality of experimental setting, tasking meaningfulness and duration of project.

Based on the above arguments, it is clear that several of these methods should be combined to facilitate an optimal solution to increasing the generalizability of such research to student populations. Given the arguments of Greenberg (1987), Stewart (2006), and Morris et al., (2004), it the contention of this paper that generalizability may be optimized using a combination of two primary factors. The first requirement is for the sample to represent the target population with respect to the constructs of interest or those that may influence the constructs of interest which may be achieved through pre-training. Gordon et al. (1986) discuss several studies in which differences between student and non-student sample results were attributed to sample differences on a specific construct thought to influence the variable of study. The studies were organizational studies and the constructs evaluated ranged from personality traits to social values to socioeconomic status (Gordon et al., 1986). In the realm of human performance, there are also participant constructs which may quite obviously affect the variable of choice. For example, if a study is exploring team performance on a military task, it is important for the participant sample to have the task relevant knowledge and skills that the target population possesses, not simply to allow completion of the task, but to facilitate completion of the task as the target population would be expected. The second requirement is for the participant to gain an understanding of the task in its context to give the task meaningfulness.

The methodology proposed herein attempts to bridge the gap in external validity by providing participants with enough training to allow them to not only immerse themselves in the context, but to also achieve the knowledge and skills necessary to accomplish target tasks and more closely represent the target population as well as providing context through training to give meaning to the tasks they are performing. As discussed above, many researchers do provide participants with task training, however, the training is usually only to a degree which allows successful completion of the task involved in the study (e.g. Morris, et. al., 2004, Lampton & Larson, 2001). This paper discusses extending this approach to provide both classroom and practical training on domain-defined requisite knowledge and skills to support participant learning and experiment external validity.

Domain Background

The need for the development of this methodology arose while performing a study which compared the training efficacy of various virtual environment training technologies in transferring training to live training environments for a Military Operations on Urban Terrain (MOUT) Room Clearing task. For these purposes all participants were treated with the same conditions but assigned different task training environments (Low Fidelity desktop VE with game pad and headset, or High Fidelity fully immersive VE with Head-Mounted Display, Headset, Haptic Vest and Airsoft Gun). In order to facilitate optimal learning of the task and to support the external validity of the results, contextual and task relevant information needed to be instilled into participants. This lead to the development of a carefully devised pre-training instruction methodology and curriculum, discussed below.

Undergraduate Boot Camp

A method was developed to increase the KSAs of undergraduate participants such that they more closely parallel the target population; this was accomplished through the design and implementation of an "Undergraduate Boot Camp." This methodology was developed based on MOUT domain documentation and in conjunction with a Subject Matter Expert (SME). The training included both classroom and practical application training in which the SME instructed, demonstrated and evaluated the trainees' performance. This was supplemented by detailed scenario based feedback targeted at developing self and team awareness of task accomplishment. A key aspect of the methodology was to develop a central set of constructs, incorporate them into a mnemonic, and wrap instruction around the mnemonic such that it supported rapid development of a mental model of the target domain. The goal of this method was to begin to provide a methodology for systematically providing quality pre-experimental instruction to trainees, resulting in an experimental group that has basic knowledge, skills, and strategies for interacting in the targeted domain.

METHODOLOGY

The methodology for development of the Undergraduate Boot Camp included two primary elements: 1) Task Analysis and Training Objective Identification, 2) Training Plan, Procedure and Material Development all of which are detailed in the sections below.

1. Task Analysis and Training Objectives Identification

As discussed previously, the experimental task focused on room clearing, one critical component of MOUT operations. A task analysis for the room clearing task was conducted (Milham, Gledhill-Holmes, Jones, Hale, & Stanney, 2004) through interviews, collaborations with SMEs, observation of task demonstrations, and reviews of military doctrine. The task analysis included a breakdown of the room clearing task into Tasks and Subtask, KSAs necessary to complete these tasks, and metrics and criterion performance standards.

Based on this, underlying competencies required for performance, Training Objectives, were identified. Specifically, the Knowledge, Skills, and Attitudes for successful room clearing were utilized to create training materials, performance metrics and standards (see Tables 1 and 4).

Table 1: Training Objectives and Performance Metrics

1. Technical Skills in MOUT Performance
- Engagement/Acknowledgement
  - Enemies Neutralized
  - Noncombatants Acknowledged
  - Reaction Times
  - Decision Times
  - Missed Shots
- Room clearing
  - Percentage of room scanned
  - Time to clear room
- Survivability
  - Shots taken
- Exposure (doorways, windows, entryways)
  - Danger areas
  - Enemy line of sight

2. Higher Order Skill Sets
- Spatial/Relational Knowledge
  - Danger areas
  - Room clearing order
  - Weapon Muzzle Awareness/Masking
- Strategic Knowledge/ Decision Making
  - Noncombatants Shot
  - Enemies Acknowledged
  - Entry Method
- Procedural Knowledge/Skill
  - Noncombatants Shot
  - Time spent in entry way danger area
  - Time spent in open door danger area
  - Time spent in window danger area
  - Entry Method
  - Proper Stacking before room entry
- Team Coordination/Performance/Communication
  - Separation Threshold
  - Room Entry Time Difference
- Affective and Attitudinal
  - Stress
  - Team Efficacy

2. Training Plan, Procedure and Material

After identification of the training objectives, a training curriculum was developed that included the following components: Classroom Instruction, Practical Instruction with SME Feedback, Rehearsal Material, and Training Feedback.

A training plan was developed, which consisted of (5) phases: 1) Initial Classroom Instruction: Each team proceeded through an initial instruction to familiarize the participants with the task context, and KSAs. This was the core of the contextual training, necessary to immerse the participants in the MOUT context and develop the appropriate KSAs by having direct instruction with an SME. 2) Rehearsal Review: Given that there would be a period of several weeks between when participants completed the initial instruction and the actual experimental task, a review was given just before engaging in the experimental training task. 3) Familiarization Practice: Before engaging in the training task participants familiarized themselves with the training equipment and environment. This was necessary to ensure that the participants were familiar with and felt comfortable interacting with the test environment. 4) VE MOUT Training: During the training task, through the VE scenarios, participants were subjectively assessed and provided feedback on their performance. This portion of the training was conducted in either of the two VE conditions. Feedback at this stage was imperative as a means to correct performance and maintain task context. Phases 2, 3 and 4 occurred in order in the same day. 5) MOUT Testing: This was the last phase in the experimental design in which the outcomes of the different training means (High and Low Fidelity) would be evaluated (this last phase took place within one week of phase 4. These five phases are illustrated in Figure 1 and explained in detail below.

Using the information gathered through the task analysis, the KSAs and continuous input from a SME, a training curriculum was developed that included the following components: Classroom instruction, Practical Training with SME feedback, Rehearsal Material, and Training Feedback.

Phase 1: Initial Classroom Instruction

Potential participants were brought in for initial training and screening. This phase served to provide training on the MOUT task. In an attempt to elevate participants to a MOUT comprehension level closer to that of actual Marines coming out of the school house, participants listened to a lecture given by a live MOUT instructor whose training was tailored to match the identified KSAs, followed by a question and answer session. Later the same instructor observed, rated and provided feedback to each participant as they practiced the material just learned (e.g. guided practice).

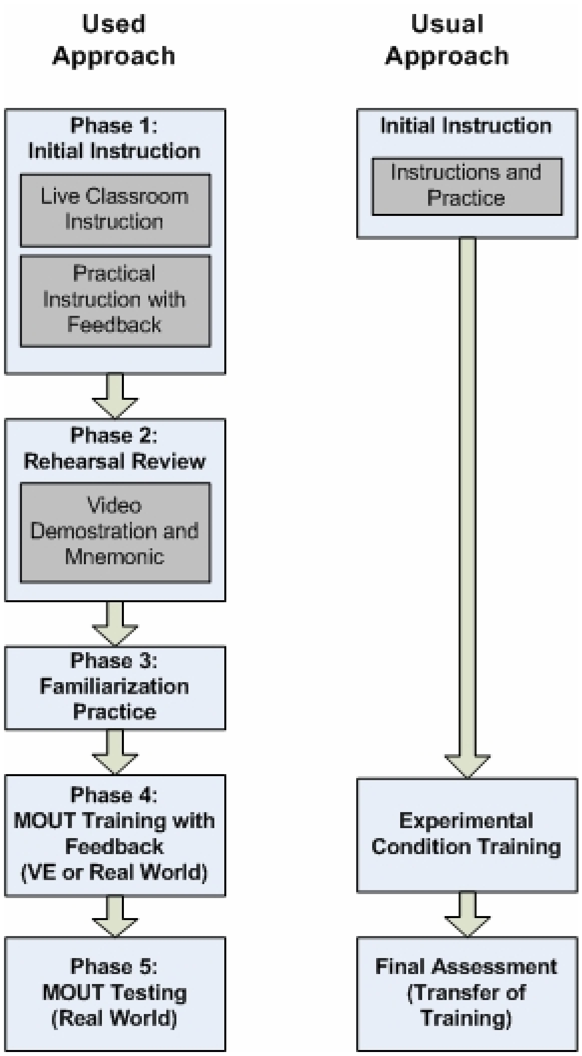

Figure 1. Procedure Process

Given that participants would be required to learn physical skills such as postures, movement techniques, and communications, while coordinating with other participants, participants were required to recreate and practice these under observation from the instructor. The instructor observed and then provided feedback on their performance as they practiced these concepts individually and in team formations, past research (Cohn et al., in press) has indicated that the target insertion point for VE training technology is after classroom training. The VE provides an opportunity for trainees to consolidate declarative knowledge they obtained in the school house and move on to acquiring procedural and strategic knowledge. Thus, for this study initial instruction was used to approximate knowledge that would be conveyed in the classroom. Specifically, a MOUT instruction session was conducted prior to actual experimentation. The goals of the initial instruction were to provide: 1) individual training; 2) team training; and 3) practical application exercises to the participants. The structure of this 2-hour training session is illustrated in Table 2. The lecture was developed by an (SME) and was based on the operational MOUT 45 minute course. In addition, the lecture included best practice strategies and the use of a mnemonic (STAC DECS see Table 3.) to assist participants in remembering important goals, techniques, and strategies. A mnemonic was used to encourage participants to organize, rehearse and recall the new information just learned, and to highlight the important aspects that were primordial in the task. Note that the mnemonic was developed as a by-product of the Task Analysis and the resulting KSAs.

Table 2. Pre-Training Session Structure

Topic	Method of Instruction	Time
Individual and Team Training	Lecture	60 minutes
Practical Application	Hands-on Practice and feedback with MOUT expert	1 hr (Each participant will practice each technique 3 times)

Table 3. Instruction Mnemonic

Mnemonic: STAC DECS
Survivability **T**echniques **A**wareness of Muzzle **C**learing **D**iscrimination **E**xposure **C**ommunication **S**tacking

Phase 2: Review Instruction

Prior to experimentation, all participants received MOUT task refresher instruction. The goal of the Review Instruction was to review the materials learned in the initial instruction using a summary video that highlighted key portions of the material. This material included examples of both correct and incorrect performance, and provided illustrations of all the key points utilizing the STAC DECS mnemonic. This phase was necessary to prime participants to the important aspects of the task, reconstruct the task context, and assist participants in the recall of learned concepts with the aid of the mnemonic.

Phase 3: Familiarization Practice

Before engaging in the training task participants familiarized themselves with the training equipment and environment. Participants proceeded through an equipment/environment training session to familiarize themselves with the equipment and to ensure that they understood how to perform basic tasks within the training system. Specifically, they watched a presentation that introduced the equipment and the cues in the environment which they would encounter. They then explored the equipment before donning it, completing a practice scenario and were given the opportunity to ask any questions they may have about the procedures or experiment in general.

Phase 4: Training Trials

Next, participants, in four-member teams, proceeded through the experimental training trials of scenarios in their respective environments. The experimenters prescribed who among the participants would be the front-man for each trial, rotating everyone through an equal amount of times. During the trials, trainees were given feedback after every trial for the first four trials, then after every fourth trial until they reached 20 trials.

Corrective feedback was used during training to reinforce appropriate behavior, correct misconceptions, and extend learning opportunities (Werts et al., 1995). Feedback is what differentiates training from mere practice of already learned skills. Thus, in this study, where the experimental training trials represented “training” that would have been received after school house instruction, feedback was provided. In an operational context, this feedback represented the After Action Review (AAR) that is provided during MOUT field training operations. Feedback was an important component of this phase given that it provided participants with the needed evaluation of their performance. The procedure allowed participants to engage in the environment and then correct their performance based on the received evaluation. This not only ensured that no “bad habits” crept into the task, but also assisted in maintaining the task context in the participants mind. The details of the feedback procedure are discussed further below.

Feedback

Feedback is a long accepted training component (Bilodeau, Bilodeau & Schumsky, 1959) necessary to ensure progressive improvement and guard from undesired deterioration of performance. Furthermore it has been observed that feedback may produce long lasting effects in some domains (e.g. differences observed 5 years post-training, for interviewing skills during a psychiatry clerkship, Maguire, Fairbairn and Fletcher, 1986). In training, assessment feedback encourages retention through encoding by allowing a trainee to clarify what has been learned and what remains to be learned (Ricci, Salas, Cannon-Bowers, 1996)

As it is critical to have feedback that is both useful and consistent across teams, a feedback methodology was devised. This methodology encompassed a subjective assessment of mistakes made by participants along a set of pre-established criteria. For this an assessment form and feedback guide was developed and tested to ensure reliability and efficiency in use across multiple experiment observers. Feedback was provided on the targeted performance outcomes (based on the mnemonic STAC DECS) listed in Table 4 using a structured feedback methodology (Smith-Jentsch et al. 1997). Steps in the feedback methodology include: identification of critical team behaviors, observation of performance, and providing information to allow for self-correction. For this study the method was instantiated with the following steps:

1. During training, trained experimenters observed each trial and used the MOUT Mnemonic Feedback Form to record the number and specific types of errors committed. The errors were based on the performance categories summarized in Table 4.
2. Following a training trial the experimenter provided the team with the Mnemonic Feedback Form with their performance data and those categories in which there were mistakes circled.
3. The team was then given 2 minutes to discuss among themselves how they would make needed corrections.

While feedback should be as consistent as possible across all groups, it is recognized that there was to be some variability/inconsistency in assessing performance; however, some inconsistency exists in the operational world (during AAR of MOUT operational training), as well. To minimize the inconsistency, assessors received extensive feedback training and their inter-rater reliability was assessed to determine proficiency in assessing performance. Furthermore event-based measurement, as a structured method of collecting performance data, has been found to increase agreement. Non-expert trained raters can use event-based measures and achieve agreement with expert trained raters. Fowlkes and Milham (2000) found that correlation between an expert and a non-expert rater was high (r = .89), when both evaluated a student navigator with an event-based measurement tool during a low-level mission in an aircraft.

Table 4. Performance Categories for Feedback

- Survivability
  - Avoided enemy and friendly fire/Number of Shots Received
- Techniques
  - Used correct Entry Techniques
  - Used correct Room Clear Techniques (L-Shape)
  - Used correct Hallway Techniques (T-Shape)
- Awareness of Muzzle
  - Avoided Flagging Muzzle
  - Avoid Pointing Muzzle At Teammate
- Clearing
  - Engaged and Acknowledged All Entities,
  - Cleared All Spaces
- Discrimination
  - Engaged /Acknowledged Entities Correctly
  - Adequate Reaction Time
- Exposure
  - Silhouetting avoided
  - Exited Fatal Funnel Quickly
  - Moved To Points of Domination (Corners)
- Communication
  - Used Correct Executions Calls
  - Used Correct Confirmatory Calls
- Stacking
  - Stacked Properly

Phase 5: Training Trials

During the last phase the teams were brought to a real world environment and asked to perform a room clearing task under new scenarios to assess their proficiency in the MOUT task.

DISCUSSION

There were several pieces of evidence pointing to the success of the methodology, including SME testimony that participants actually looked more like their Marine counterparts after completing the pre-training sessions. Additionally, ad-hoc observations from pilot testing (Stanney, et. al, study in progress) revealed some readily observable distinctions in participant behavior. For instance a pair of participants that was pre-trained using the described material was consistently aware of the task context and showed restraint in the display of "game behavior" (e.g. ignore dangers, disregard of doctrine, indiscriminant shooting). While a test pair who received context training through a video and presentation lost context and displayed "game behavior" within a few trials. One additional anecdotal observation is that of the observed differences between the performance measures of these two pairs of participants. The team that in addition to having context rich pre-training also had the training described herein showed a percentage increase in performance for a "Team Coordination" measure that was four times superior to the other group. This supports observations made by Morris, et. al. (2004), whose context enhanced group performed better in completion type metrics, while doing poorly in doctrine type measures. In contrast, both teams had context enhanced pre-training, with one having KSA and SME training showing superior doctrine learning. While caution should be utilized while interpreting these results given the small sample, this methodology provides an initial approach to addressing the shortcomings observed by Morris, et. al. (2004), and providing greater external validity to the observed results in this type of research.

CONCLUSION

As the use of student populations in research is likely to continue to be utilized, it is critical to define and address issues related to their generalizability. As such, steps should be taken to ensure participants 1) possess the task relevant KSAs that the target population possesses to facilitate completion of the task as the target population would be expected, and 2) gain an understanding of the task in its context to give the task meaningfulness.

The methodology presented above attempts to bridge gaps in the generalizability of research finding using undergraduate participants by providing pre-training and during-training feedback which facilitates the learning of requisite KSAs necessary for task accomplishment, in addition to an understanding of the requisite tasks and the motivations behind mission success.

REFERENCES

Aidman, E., Galanis, G., Manton, J., Vozzo, A., and Bonner, M., 2002. Evaluating human systems in military training. *Australian Journal of Psychology*. Vol. 54, No. 3. pp 168-173

Bilodeau, E.A., Bilodeau, I.M., and Schumsky, D.A., 1959. Some effects of introducing and withdrawing results early and late in practice. *Journal of Experimental Psychology*. Vol. 58, pp 142-144

Cohn, J., Schmorrow, D., Nicholson, D., Muller, P., Stanney, K., & Milham, L. (2005). Trainee-Centric Design for Developing Military-Based Virtual Environment Training Systems. *Proceedings of the 2005 VR International Conference*. Las Vegas, NV

Dobbins, G.H., Lane, I.M., and Steiner, D.D. (1988). A note on the role of laboratory methodologies in applied behavioral research: Don't throw out the baby with the bath water. *Journal of Organizational Behavior 9*, 281-286.

Fowlkes, J. & Milham, L. (2000). *Team member perspective in defining situation awareness.* Paper presented at APS. Miami, FL.

Gordon, M.E., Schmitt, N., and Schneider, W. (1984). An evaluation of laboratory research bargaining and negotiations. *Industrial Relations*, 23, 218 – 233.

Gordon, M.E., Slade, L.A., and Schmitt, N. (1986). The "Science of the sophomore" revisited: From conjecture to empiricism. *Academy of Management Review, 11*, 1, 191-207

Lampton, D.R., and Parsons, J.B., 2001. The fully immersive team training (FITT) research system: design and implementation. *Presence*, Vol. 10, No. 2, pp 129-141

Maguire, P., Fairbairn, S. and Fletcher, C., 1986. Consultation skills of young doctors: Benefits of feedback training in interviewing as residents persist. *Medical Education*, Vol. 14, No. 292, pp 573-576

Morris, C.S., Hancock, P.A., and Shirkey, E.C., 2004. Motivational Effect of Adding Context Relevant Stress in PC-Based Game Training. *Military Psychology*. Vol. 16, No. 1, 135-147.

Ricci, K.E., Salas, E., and Cannon-Bowers, J.A., 1996. Do computer-based games facilitate knowledge acquisition and retention? Military Psychology, Vol. 8, No. 4, pp 295-307

Smith-Jentsch, K. A., Cannon-Bowers, J. A. & Salas, E. (1997). New wine in old bottles: A framework for theoretically-based OJT. In K. A. Smith-Jentsch, Toward a continuous learning environment. Paper presented at the 12th annual meeting of the Society for Industrial and Organizational Psychology, St. Louis, MO.

Ward, E. A. (1993). Generalizability of psychological research from undergraduates to employed adults. *Journal of Social Psychology, 133*, 4, 513-519.