Evaluating Mission Training Fidelity Requirements: Examining Key Issues in Deployability and Trainability

Brian T. Schreiber
Lumir Research Institute
195 Bluff Avenue
Grayslake, IL 60030
USA

brian.schreiber@mesa.afmc.af.mil

Winston Bennett Jr
Air Force Research Laboratory
6030 South Kent Street
Mesa, AZ 85212
USA

wink.bennett@mesa.afmc.af.mil

Robert Rickard
Rickard Consulting Group, Inc.
2642 North 138th Avenue
Goodyear, AZ 85338
USA

robert.rickard@mesa.afmc.af.mil

Jeffry Bell
Anteon
6030 South Kent Street
Mesa, AZ 85212
USA

LABell67@aol.com

Michael France
Anteon
6030 South Kent Street
Mesa, AZ 85212
USA

michael.france@mesa.afmc.af.mil

David Greschke
Anteon
6030 South Kent Street
Mesa, AZ 85212
USA

david.greschke@mesa.afmc.af.mil

EVALUATING MISSION TRAINING FIDELITY REQUIREMENTS

Due to infrequent training while deployed, warfighters' skills can decay, thus creating a training gap. Deployed training has historically been hindered by restrictions to live-fly training opportunities due to factors such as ops-tempo, airspace/range restrictions, security issues, alert requirements, and wartime rules of engagement. In order to maintain high proficiency and readiness levels, changes are needed in standard training programs while warfighters are deployed. The United States Air Force's Distributed Mission Operations (DMO) concept has become critical to warfighter training across all mission areas. Despite an increased reliance on DMO training, a deployable DMO training capability does not exist to provide the critical training opportunities previously unavailable during extended deployments. Moreover, while researchers have discussed and described the tradespace associated with varying levels of physical and functional fidelity, opportunities to conduct controlled studies to examine and quantify "how much of what kind of fidelity" for skills improvement and maintenance are few and far between. This presentation will describe a series of research studies underway at the Warfighter Readiness Research Division in Mesa

Schreiber, B.T.; Bennett, Jr., W.; Rickard, R.; Bell, J.; France, M.; Greschke, D. (2006) Evaluating Mission Training Fidelity Requirements: Examining Key Issues in Deployability and Trainability. In *Virtual Media for Military Applications* (pp. 15-1 – 15-8). Meeting Proceedings RTO-MP-HFM-136, Paper 15. Neuilly-sur-Seine, France: RTO. Available from: http://www.rto.nato.int/abstracts.asp.

Arizona designed to assess within-training and training transfer to live operations for both individual and tactical teams, with varying levels of fidelity in the training environment. These studies are examining a variety of specific parameters associated with achieving and maintaining combat readiness and proficiency. We conclude with discussion of our longer-term plan to demonstrate and quantify the impact of deployed training and rehearsal environment on transfer to the operational environment and to subsequent retraining intervals.

1.0 CHANGING TRAINING NEED

A shift in available resources and changes in the strategic environment demand a transformation in how we prepare the warfighter. The financial cost of actual training events has increased, while the opportunities to conduct actual training events have decreased. Operational tempo, security issues, alert requirements, environmental restrictions, encroachment on training areas, and the reduced tolerance of the general population for close proximity of military training have all combined to make it more difficult to carry out the type of live training activities common 20 or even 10 years ago [1]. Additionally, the emphasis has shifted towards smaller deployed forces capable of operating anywhere, anytime, day or night, and in all weather. Training must, therefore, transform to support the needs of the force. Often, our warfighters are deployed for extended periods of time, but without sufficient opportunity to train/maintain their skills. This insufficient deployed training results in skill proficiency loss, reduced readiness, and additional training time later to reconstitute those skills. Unfortunately, deployable systems for use in-theatre do not currently exist to fill this training gap. In order to maintain high proficiency and readiness levels, changes are needed in standard training programs while warfighters are deployed.

Meeting the training need of the deployed warfighter encompasses a range of requirements. Ideally, a deployable system would allow the warfighter to train critical aspects of the "job", i.e., the transfer environment. The "job" is, of course, more than what is learned in the classroom and more than just checklists and emergency procedures. A warfighter must be prepared for Mission Essential Competencies (MECs), which are the higher-order individual, team, and inter-team competencies that a fully prepared pilot, crew, flight, operator, or team requires for successful mission completion under adverse conditions and in a non-permissive environment [2]. The conditions of performance in this MEC definition are most important. The focus here is on *combat conditions* to define standards of warfighter performance [3].

1.1 Distributed Mission Operations Training as a Solution

Distributed Mission Operations (DMO) training capabilities can be generally defined as affording the ability to bring a number of warfighters together to train complex individual and/or team tasks during the course of larger-scale, realistic combat missions [4]. The very definition of DMO appears to align closely with the combat-relevant conditions of performance necessary as part of maintaining MEC proficiency. Interestingly, in 1991, a time period just preceding DMO, researchers [5] surveyed 94 F-15 air combat pilots and discovered higher order experiential areas as receiving less than adequate training in their current unit. The experiences identified were: bogey, reaction to SAMs, dissimilar air combat tactics, all-weather employment, electronic countermeasure employment, communications jamming, low altitude tactics, chaff/flare employment, escort tactics, track-while-scan assessment, and work with the Air Weapons Controller. The F-15 pilots also felt all of these experiences were better suited to the simulator, this opinion provided even during an era in which much of the benefits of networking and the emerging DMO environment were still absent! The early evidence for simulator advocacy was not isolated. Other practitioners in 1993 [6] interviewed 99 F-16 pilots, who reported that the highest "difficult aspects" of attaining/maintaining mission-ready status were weapons delivery, radar interpretation, electronic combat, cockpit switchology, and air-to-air combat. When the F-16 pilots were asked which ground-based medium they would like to see used more, the most preferred option was the simulator.

Another large sample pilot opinion study was conducted recently, specifically on contemporary DMO training. In that study [7] the authors surveyed 327 F-16 pilots and 49 Airborne Warning and Control Station controllers, who based their opinions on experiences obtained during a five-day DMO training research program containing over 40 total air combat scenarios. "I would recommend this training experience to other pilots/controllers" was rated with the highest rating possible of "Strongly Agree" by all but one of 49 controllers and all but 16 of 327 pilots. When asked to rate to what extent the DMO environment provided the 45 different F-16 critical air-to-air experiences (defined by the MEC process), the pilots rated that 38 experiences (84%) could be obtained "to a moderate extent" or higher, more than all seven other training environments surveyed.

These large sample, opinion-based studies reveal warfighter advocacy for DMO, but how effective is it as a training device? Recently, a large-scale study was done specifically to examine DMOs training effectiveness [8]. 76 F-16 teams (384 pilots) flew for one week in DMO and were "tested" at the beginning of the week and again at the end of the week on mirror-image point defence scenarios. On the post-test, the teams, on average per scenario, allowed 58.33% fewer enemy strikers to target, killed 9.20% more enemy aircraft, reduced F-16 mortalities by 54.77%, and registered 55.20% less time allowing hostiles into vulnerable ranges. Also part of this study, expert observer ratings—both those taken in real-time and those done according to a scientifically blind protocol—revealed statistically significant improvements as a function of the DMO training. The improvements in expert ratings were found both for briefs/debriefs and also for mission execution, corroborating the objective results. DMO training is obviously an effective training paradigm, but no DMO solution currently exists to fulfill the deployed warfighter's training need. So, what are the issues involved in researching and building a deployable DMO solution?

1.2 Issues in Researching and Building a Deployable Distributed Mission Operations Tactical Trainer

Designers and users of training devices often attempt to replicate as many physical and functional stimuli as possible in a training device [9]. The goal here is often to increase physical fidelity (i.e., look & feel) and functional fidelity (i.e., the dynamics, or "actions"), two arguably interdependent system aspects [10]. Prior research addresses familiar themes in building the "right" simulation training environment, with many of the issues revolving around the level of fidelity needed. In general, the literature tends to support the notion that higher fidelity will translate into a better training device (e.g., [11]), but this does not mean that effective training devices must all be high fidelity. Indeed, even low fidelity pilot trainers absent of stick and throttle have shown to positively transfer compared to a control group [12]. Further, some studies have shown that added fidelity does not significantly contribute to performance [13]. As a rule-of-thumb, however, increasing fidelity tends to increase the training utility and value, perceived and/or real. But, increases in fidelity are coincident with costs [14]. Increasing costs almost inevitably leads to fidelity compromises. One solution is to build a reconfigurable system capable of training different platforms (e.g., [15]). However, adopting a reconfiguarable system does not ameliorate, and may exacerbate, the difficulty of choices needed to be made when it comes to cost versus fidelity.

To build a capable deployable DMO tactical trainer (in our case an F-16 tactical trainer), a number of objectives are desirable, including the ability: (a) to train forces with a realistic mix of experiences from strategic to tactical levels; (b) to integrate live, virtual, and constructive elements within a common synthetic environment; (c) to integrate low density/high demand assets through distributed simulation; (d) to practice missions requiring a "total team" environment to meet national objectives; (e) to meet institutional training goals and needs for individual, collective, and joint/coalition training; (f) to tailor training for individuals and teams; (g) to support Air Expeditionary Force training and preparatory spin-up and sustain wartime readiness; and, (h) to provide joint training with command linkages. To work towards these objectives, scores of decisions must be made regarding physical fidelity (e.g., display field of view, stick and throttle, switches, etc.), and functional fidelity (e.g., flight model, weapons models, threat generation system, etc.). Of course, the same central question presents itself: In building a DMO environment, how do we know which fidelity components, when interconnected, provide the best training value for the money spent? And, once the system is built, additional questions must be addressed, including those that relate to assessing, tracking, and comparing warfighter performance both within a deployable DMO system and at other locations. That is, what approach needs to be taken to assess combat readiness and proficiency across sites? Such questions, properly addressed, would provide answers for the degree of training transfer between environments, skill acquisition rates, and recommended DMO retraining intervals.

2.0 APPROACH

To build and complete sound empirical evaluations of all possible, practical, variations of deployable DMO tactical trainers is cost-prohibitive, especially given the rapid change in technology progress and alternatives. Therefore, the approach of choice should be appropriately aggregated so as to avoid assessing all possible technology interaction effects, yet still provide meaningful information. A simple approach yielding data-based recommendations is needed. Likewise, an approach to assessment, tracking, and comparing performance needed to be simple if it is to be implemented successfully across sites. These common assessments would provide the proficiency-based data needed to help answer questions regarding skill acquisition, decay, training transfer, etc.

2.1 Building and Evaluating the System

Our goal for the deployable system is to meet the needs of the warfighter with a level of fidelity that best prepare him/her for *combat experience*. To this end, we are able to leverage outputs from the MEC process, in which operational warfighters are brought together and define a critical list of experiences they believe would fully prepare a warfighter for combat. In the case of the F-16, the MEC process identified 171 total experiences across a variety of mission types (e.g., air-to-air, air-to-ground, suppression of enemy air defense). Selected examples of the F-16 experiences are provided in Table 1. Note that many of the experiences require the simulation to be a *system of technologies* to include sophisticated capabilities (e.g., night vision, various threat capacities) and multi-player connectivity (e.g., buddy lasing, large force employment).

Table 1: Selected Mission Essential Competency Experiences for the F-16

Performing buddy lasing	Unlocated or pop-up threats
Employing weapons with degraded systems	Using chaff/flare for threat reaction
Operating in mountainous versus flat terrain	Seeing surface-to-air threat launches under NVG
Large force employment	Degraded aircraft avionics

These 171 experiences provided us with a comprehensive set of mission/tactical level ideals. To augment this exhaustive tactical level list, we chose to add the emergency procedures (EP) list from the United States Air

Force Dash One document for the F-16. The MEC experiences and the EPs, if fully and faithfully realized in any DMO environment, would represent the highest trainability/usability standards we could hope to achieve. Hence, the warfighter-centered list therefore became our system evaluation criteria.

Since scientifically evaluating all technological permutations was cost-prohibitive and impractical, our initial deployable DMO build/configuration was decided upon by examining the spectrum of MEC experiences for *common* underlying functionalities/requirements. Examples of these included F-16 flight dynamic model, threat generation system, avionic displays, visual display system, simulator connectivity, weapons models, terrain database, etc. For each common area, an initial low-cost solution was identified, procured, and integrated, thus creating an initial, baseline deployable DMO configuration. This configuration was documented in a survey and is in the process of being evaluated by F-16 Subject Matter Experts (SMEs), who separately rate the ability of the system to provide each of the MEC/EP experiences. Analyses will reveal the system's trainability/usability by each experience and EP, by aggregated mission area (e.g., air-to-ground), and overall. Examining commonalities amongst the lowest rated MEC/EP experiences will reveal the technological deficiencies that need to be addressed (e.g., insufficient field of view). The deficient functionalities will form a follow-up survey. In this survey SMEs will rate each deficiency against each of the MEC experiences/EPs for the degree to which each detracts from the ability to gain/train that experience/EP. The results will provide a priority order of fidelity/technology areas to be upgraded, constrained by cost considerations.

Once this evaluation/development cycle is complete, a number of scientific opportunities become available. First, we can re-evaluate the (upgraded) system to determine the training value added for the cost incurred. Second, we can implement the same survey evaluation process on other, higher fidelity environments for additional comparison of trainability/cost. And third, in conjunction with data from the first two efforts, we can re-use the same two-cycle evaluation/deficiency survey process iteratively in building the deployable DMO system to determine the point of diminishing returns, in terms of added trainability/usability for additional costs. Though this approach cannot delineate all possible interaction effects between technology subsystem components, it does provide us cost-effective data-driven assessments of the overall system and the fidelity/cost trade-offs involved.

2.2 Need for Standardized Assessment Suite (SAS) and protocol.

Once a configuration of the deployable DMO trainer is settled and deployed for training use, a number of research questions arise. Research will need to be undertaken in the deployable trainer to examine skill acquisition rates needed to meet proficiency, skill decay rates (for new and experienced pilots), training effectiveness, and degree of transfer between the deployable trainer and live-fly and other DMO training environments. To examine these issues, a central, critical need is to quantitatively assess *proficiency* using a standard, controlled protocol. Doing so permits comparisons of warfighter skill levels both longitudinally and across sites. To meet this assessment criterion for addressing the research objectives, a standardized assessment suite (SAS) will need to be developed to evaluate warfighter performance within not only the deployable trainer, but also any other DMO location used as part of the investigation.

Conceptually, the SAS approach provides a common method to compare warfighter proficiency across all DMO environments (e.g., simulator, live-fly training, exercises, and potentially combat locations)—similar to the method of using Graduate Record Exams (GRE) across universities to assess student proficiency. At periodic intervals within the deployable trainers (and other DMO locations), a set of missions would be performed under strict, standardized data collection protocols. The set of missions to be performed would be a sample of the suite, representing a spectrum of difficulty. The scenarios comprising the mission "set" to be periodically flown would be a sample of the overall library of available "test" missions. The assessments for all scenarios would be standardized, leveraging final product versions from three projects currently under development: (1) MEC vignette knowledge test [16], (2) subjective assessment of skill proficiency [17], and (3) objective data collected directly from the simulation systems [18].

At each DMO location, experiment control of the standardized assessment suite and a protocol for managing the training conditions and for data collection will be required. However, no restriction is needed regarding who participates and when, only that demographic data is captured on each participant when he/she performs a SAS set of scenarios (e.g., dates of participation at each location and relevant moderator training events such as weapons school). By utilizing this approach, data can be collected on all participants through each DMO environment, thereby maximizing sample size. By controlling the SAS tests and tracking other moderating demographic variables of interest, sufficient data can be collected at each site to allow for the application of multiple regression and/or mixed models of analysis to subsets of the database that address specific research questions of interest.

These mixed models of analysis will use the standardized assessment suite as a multivariate outcome space and use other variables (site/fidelity, pilot, important potential moderators) as predictors to address the research questions of skill acquisition rates to meet proficiency, skill decay rates, training effectiveness, fidelity to meet proficiency criterion, and degree of transfer between DMO environments. In addition to the benefits of collecting data observationally at each DMO location (i.e., not having predefined experimental conditions) and therefore maximizing sample size, the SAS approach was also chosen because it: (a) minimizes necessary DMO environment controls, this method is likely to introduce the lowest amount of risk; (b) requires fewer controls to implement, track and enforce, leading to the lowest research labor cost possible; and (c) ensures data in the alternative environments will be tracked according to their operational uses, resulting in an ability to compare DMO training strategy implementations and relative effectiveness of each environment. This SAS approach, if successful within the deployable trainers and a few other DMO locations, would then become a key component in the USAFs training transformation plan to augment the Ready Aircrew Program with competency-based assessment.

[1] Schank, J. F., Thie, H. J, Graf, C. M., Beel, J., & Sollinger, J. (2002). *Finding the right balance: Simulator and Live Training for Navy Units*. Santa Monica: RAND Corporation.

[2] Colegrove, C. M. & Alliger, G. M. (2002, April). *Mission Essential Competencies: Defining Combat Mission Readiness in a Novel Way*. Paper presented at the NATO RTO Studies, Analysis and Simulation Panel (SAS) Symposium. Brussels, Belgium.

[3] Alliger, G. M., Beard, R., Bennett, W. Jr., Symons, S., & Colegrove C. (in press). A Psychometric Examination of Mission Essential Competency (MEC) Measures Used in Air Force Distributed Mission Operations Training Needs Analysis. *Journal of Military Psychology.*

[4] Chapman, R., Colegrove, C. M., & Greschke, D. A. (in press). Transforming Operational Training in the Combat Air Forces. *Journal of Military Psychology.*

[5] Houck, M. R., Thomas, G. S., & Bell, H. H. (1991). *Training evaluation of the F-15 advanced air combat simulation* (Rep. No. AL-TP-1991-0047, AD A241 675). Williams AFB, AZ: Armstrong Laboratory, Aircrew Training Research Division.

[6] Gray, T. H., Edwards, B. J., & Andrews, D. H. (1993). *A Survey of F-16 Squadron-level Pilot Training in PACAF* (Rep. No. AL-TR-1993-0041). Williams AFB, AZ: Human Resources Directorate, Aircrew Training Research Division.

[7] Schreiber, B. T., Silverberg-Rowe, L., & Bennett, W. Jr. (2006). *Distributed Mission Operations Training Effectiveness Baseline Study*. Volume IV: Participant Utility and Effectiveness Opinions and Ratings.

[8] Schreiber, B. T. & Bennett, W. Jr. (2006). *Distributed Mission Operations Training Effectiveness Baseline Study*. Volume I: Summary Report.

[9] Andrews, D. H., Carroll, L. A., & Bell, H. H. (1994). *The Future of Selective Fidelity in Training Devices* (ADA316902). Wright-Patterson Air Force Base, OH: Armstrong Lab Human Resources Directorate.

[10] Allen, J.A., Hays, R. T., & Buffardi, L. C. (1986). Maintenance training simulator fidelity and individual differences in transfer of training. *Human Factors, 28(5),* 497-509.

[11] Gebman, J. R., Stanley, W. L., Barbour, A. A., Berg, R. T. & Birler, J. L. (1986). *Assessing the Benefits and Costs of Motion for the C-17 Flight Simulators* (ADA174822). Washington DC: Department of the Air Force.

[12] Dennis, K. A. & Harris, D. (1998). Computer-based simulation as an adjunct to ab initio flight training. *International Journal of Aviation Psychology,* 8, 261-276.

[13] Evans, K. L. (1989). *Effects of Recoil on Rifle Marksmanship Simulator Performance* (ADA 206984). Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences.

[14] Hughes, T. & Rolek, E. (2003). Fidelity and Validity: Issues of Human Behavioral Representation Requirements Development. In *Proceedings of the 2003 Winter Simulation Conference*. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

[15] Williams, G. W., Lawrence, K., & Weeks, R. (2004). Modeling and Simulation Technologies: Reconfigurable Flight Simulators in Modeling and Simulation (AD56624). In *Proceedings of AIAA Modeling and Simulation Technologies Conference*. Reston, VA: American Institute of Aeronautics and Astronautics.

[16] Gehr, S., Schreiber, B. T., Metz, C., & Bennett, W. Jr. (2005). Air Superiority Knowledge Assessment System. In *2005 Interservice/Industry Training, Simulation and Education Conference (I/ITSEC) Proceedings*. Orlando, FL: National Security Industrial Association.

[17] MacMillan, J., Entin, E., Morley, R., & Bennett, W. Jr. (in press). Measuring Team Performance in Complex and Dynamic Military Environments: The SPOTLITE Method. *The Journal of Military Psychology.*

[18] Schreiber, B. T., Watz, E., Bennett, W. Jr., & Portrey, A. M. (2003). Development of a Distributed Mission Training Automated Performance Tracking System. In *Proceedings of the Behavioral Representations in Modeling and Simulation (BRIMS) Conference*. Scottsdale, AZ.