

Performance-Based Learning in Operational Training Environments

Major General Melvin Spiese, USMC (ret.)

And

Camilla Knott, Ph.D.

The Problem

Warfighting no longer means just going into combat. There was a time not long ago, where success was achieved with technical and tactical proficiency, supported by sound military decision making, both individually and organizationally. Reflected in Operation Iraqi Freedom-I (OIF), that model changed dramatically as counterinsurgency (COIN) replaced conventional military operations. Success in Operation Iraqi Freedom-II, in its many iterations, was achieved through the combination of combat proficiencies and the addition of complex cognitive and social skills. More than additive to a menu of abilities for individuals and organizations, this evolution in skills changes the dynamic of individual and collaborative thinking, interaction, and decision making. These are measurably advanced skills, and not only valuable in successfully navigating the complex insurgency battlespace; they are at the heart of the flexible, adaptive, learning service member, leader and unit needed to achieve success in increasingly complex operating environments.

Training for the former set of skills had been refined and arguably perfected in the latter stages of the Cold War, reflected in the advanced military training regimes established at the US military combat training centers and their professional training organizations, and copied throughout the world. That training was successfully displayed in the First Gulf War and a decade later in OIF-I. Training transformed dramatically as the insurgency raced out of control and the needs of OIF-II saw the construction of physically and culturally immersive training environments populated with ethnically-accurate role players, supported by complex scenarios, all built on a progression of training that provided individuals and organizations the foundations of cognitive and social skills for the challenges of counterinsurgency. Adjustments in training were continuous, reflective of changes on the battlefield, particularly as the enemy morphed and altered actions and behaviors.

Service members, leaders and units were thrust into environments where success depended on effective civil engagement and actions. Intelligence was no longer a function of only higher headquarters, but generated and used at ever-lower echelons in counter network operations. Small unit leaders operated more independently in tight decision cycles and complex situations, requiring very refined and discreet application of military force¹. Success depended upon continual learning and adapting during ongoing operations in ways not considered as the force entered this decade and a half of combat. Many of these lessons learned have been incorporated into the current leader development and training strategy for the U.S. Army² and the Marine Corps³. However, a challenge remains in the implementation of the strategy that could be addressed with a shift in the approach to performance assessment.

The existing training paradigm is an experienced-based learning model, focused on task accomplishment. Typically framed in a task-condition-standard construct, task completion is the primary objective measure and assessment of the more complex cognitive and social skills is difficult and often ignored. Although After Action Reviews (AARs) probe decisions and aspects of how a task was accomplished, task completion generally measures training success, with subjective feedback based on observer-evaluator observations. Objective measures that are applied are those that are easily-determined, (e.g., speed), rather than those necessary to indicate understanding. Because of this task-oriented experience-based learning, true *understanding* of a task (i.e., how and why to do it) may be incomplete.

¹ Brown, R. (2011). The Infantry Squad: Decisive Force Now and in the Future. Military Review: Mission Command, Fort Leavenworth, KS.

² Department of the Army (2014). The Army Operating Concept. *TRADOC Pamphlet 525-3-1*. Ft. Monroe, VA: Training and Doctrine Command

³ Ross, Karol G., Jennifer K. Phillips, Jennifer J. Vogel-Walcutt, Iris D. Rivera, Tegan F. Brown, and Katelynn M. Smith. Final Report for Base Contract: Development of the Initial Small Unit Decision Making (SUDM) Assessment Battery. Rep. no. CPG-0001AE-21Dec12. Orlando: Cognitive Performance Group, 2012. Print.

The Solution

A shift to performance-based learning orients on competencies and behaviors displayed in training, with task and task accomplishment providing context. As a result, the focus is on assessment of all aspects of performance from process (e.g., decision-making, social interaction) to outcome (e.g., task completion). To achieve more comprehensive success in future training, assessment of performance must (a) move to greater objectivity, and drive an effort to assess behaviors that indicate *understanding* of a task, and (b) focus on competency over task accomplishment. For example, during a Key Leader Engagement (KLE), social interaction skills are crucial. Competencies such as self-awareness⁴ and perspective taking should be assessed in addition to the ability to control a battle space.

The U.S. Army, in particular, has identified the need to optimize human performance in a holistic manner, and is deliberately seeking technologies to achieve that goal. The pursuit indicates the Army no longer sees human performance as a byproduct or outcome of activities and events, but requires that as a specific focus of effort in the development of individuals, teams and organizations.⁵

There are several key components to re-orienting to performance-based learning. One component is related to *designing better observer-evaluator assessments*. Although these assessments are subjective by nature, greater objectivity can be designed into the measures to encourage accuracy and reliability (i.e., standardization) across assessments. For example, assessments can be designed so that observer-evaluators know what to look for based on a clear description of performance requirements, and developing performance rubrics that specify behaviors linked to various levels of performance. This approach has been shown to be effective in helping to build this standardization across observer-evaluators and across training audiences⁶. This is particularly important for designing assessments of attributes and competencies that are not traditionally evaluated as part of the training of tactics and techniques, and may be tied to behaviors that are difficult to observe e.g., perspective taking, and self-awareness.

Another component is related to *designing assessments for training objectives and outcomes that are not readily observable (i.e., beyond subjective assessments)*, either because of the size of the organization, or the large number of actions required for a task. This can be achieved in a variety of ways either with or without the observer-evaluator in the loop. However, in a world in which an observer-evaluator is often overwhelmed with data, even in an exclusively-observed environment where an observer-evaluator is required, such as in field training exercises, deliberate action by the observer-evaluator or instructor is required to either filter through data to find those necessary to answer the right questions, or recognize that data must be generated where they are missing. In both cases, effort must be made in first understanding the attributes and competencies required for good performance, defining measures that can accurately and reliably assess those attributes and competencies, and then identifying how those measures can be captured. Likewise, the goal in both cases should be to support the development of actionable and comprehensive feedback.

Yet another component is related to *defining actionable feedback*. Technical and tactical skills, though well-understood by the military services, training organizations, and training audiences, are in fact very challenging to

⁴ Hubal, R., van Lent, M., Wender, J., Lande, B., Flanagan, S., & Quinn, S. (2015). What does it take to train a good stranger? Proceedings of the International Conference on Cross-Cultural Decision Making, July 26-30, 2015, Las Vegas, NV

⁵ United States. The United States Army. U.S. Army Training and Doctrine Command. The Warfighter's Science and Technology Needs. Comp. LTC Eric J. Van De Hey and Michael A. Meneghini. Fort Eustis, VA: Science, Technology, Research and Accelerated Capabilities Division, 2016. Print.

⁶ Ratwani, K. L., Dean, C. R., Knott, C., Diedrich, F., Tucker, J. S., Flanagan, S., & Walker, K. (2015). Measuring leader attributes in the Army Reconnaissance Course. *ARI Research Product Number Pending*. Arlington, VA.

assess to an institutionally-consistent, objective standard. Feedback to participants is direct and effective, but dominated by observer-evaluator subjectivity, and inconsistent over time between evaluators and training audiences alike. Valuable in context of the event and in the moment, the results are difficult to (a) objectively absorb, (b) extrapolate to apply to all situations, and (c) leverage institutionally. The addition of these advanced skills has added great complexity to training methodologies and environments. They have increased the difficulty of accurately and reliably assessing the performance of individuals and organizations and providing actionable feedback for maximum training effectiveness.

Benefits to the development of competency and behavior-based performance measures in training include data-driven assessments and feedback that support deeper learning. When data are purposefully gathered and properly analyzed, particularly if incorporated with predictive analytics, they can enable precise recommendations for corrective actions and learning. Additionally, understanding the actual proficiency level of individuals, teams/cells, and organizations, both discreetly within organizations and for organizations as a whole, can enable tailored training in resource-constrained environments and provide objective, comprehensive unit readiness assessments that can support institutional decisions, especially related to mission-oriented expectations in operational settings.

Likewise, competency and behavior-based performance measurement can benefit the determination of unit readiness which has similar challenges. Current readiness determinations are based on easily-quantifiable factors such as manning levels, required skill matches, and equipment availability, that may not necessarily have strong predictive validity for operational effectiveness. Mission-oriented proficiency tends to be based on events accomplished in a training calendar and the subjective assessment of the commander, rather than institutionally-developed measures that give greater understanding and assurance of the unit's operational abilities. Even when outside assessments are considered, such as rotations to service-level training programs, subjectivity dominates the determination of unit readiness. Additionally, unit readiness is not typically linked to individual billet proficiency, and may in fact, mask areas of needed attention to ensure optimal performance in conducting operations.

The future promises even greater complexity. We return to a world where near-peer competition has given rise to sophisticated, high intensity threats. We see the advent of a new form of hybrid warfare and the incorporation of advanced non-lethal effects being integrated into complex military operations. The battle against an ever-changing terror threat appears to have no discernable end, requiring both direct military action and more advanced military engagements such as capacity building with partners and countries at risk, and active combat operations.

Although we have seen remarkable changes in training, our paradigm in many ways remains the same. *Training is dictated and judged by completion of activities or tasks on a training calendar or event checklist, rather than actual performance of those activities or tasks or achieving levels of competency and proficiency for required skills.* Training progressions are sequenced a priori, where subsequent trials are not adjusted or tailored based on performance or proficiency level. Individual performance tends to get masked in collaborative, collective environments, and is generally not discernable in unit assessments, unless the individual is a key performer, e.g., the unit commander or a squad leader. Most unit readiness determinations related to proficiency remain subjective and activity-based, not objectively measured against quantifiable criteria for either individual competencies or unit capabilities.

The demands on the force have never been greater. The capacity and speed of threats and range of possible mission sets cover a spectrum of reality that was once theoretical. The dynamics of the battlefield environment, and the expanse and complexities of capabilities in an increasingly dispersed force require levels of knowledge and skills that were once the purview of more senior leadership. Consequently, the definition of readiness must evolve to encompass individual and collective competencies. The development of individual and organizational proficiency, anchored in knowledge and understanding, require similar changes and precision in measurement.

Key Challenges

There are three main challenges in achieving the vision of performance-based learning: realism or training fidelity, comprehensive assessment, and cohesiveness of training. With the first two challenges, a key question involves the quality vs the quantity of the training and what is measured. The challenge related to cohesiveness of training applies to not only the intra-event training, but also the training that occurs across various domains (e.g., operational vs. institutional), and platforms (e.g., simulation vs. live).

Realism or training fidelity. One of the great successes in the training transformation of the past decade was the ability to enable troops to see and experience everything they would likely face during training. The transition to immersive environments for the counter-insurgency fight included accurately-constructed mock villages, complete with atmospherics and populated with ethnically-accurate role players having personal histories and stories. The threats to be encountered provided the context for mission rehearsal training. The actions, decisions and behaviors of leaders and service members necessary for success in such complex scenarios built experience prior to combat deployment.

This was not dissimilar to the former Cold War training construct, which enabled replication of combat actions in a more conventional construct, and was the platform for the success seen in OIF-I. The difference was the complete disruption of the warfare model as conventional operations in Iraq transitioned to counter insurgency operations. COIN drove the changes in requirements of the force and its members, and the associated growth in complexity in training, matching both the operating environment service members and units were to see, and more advanced skills (e.g., cognitive and social) necessary for success to be demonstrated.

We continue to seek training environments reflective of all the likely scenarios and challenges we expect our forces are to face in contingency and combat operations, but a key challenge for this approach is balancing cost and training effectiveness in determining fidelity. So, although we can ensure that service members and units see and experience everything they would likely face in combat, there is a tradeoff and balance for the appropriate level of fidelity or realism required for the training to impact performance. For this to be properly achieved, fidelity must be determined by impact on learning, rather than be determined by intuition.

Measurement. As crucial as providing realistic scenarios and conditions in training, is the challenge of gaining accurate and comprehensive assessments of performance to understand the training audience's ability to operate effectively and succeed, particularly in areas of collaborative interaction, leadership, and decision making- both individual and collective. In other words, while the assessment of key tactical and technical skills in an operationally-relevant environment is crucial, there are other important aspects of performance that must be considered. For example, long-term team performance depends on processes that build team trust, effectiveness, commitment, retention, and communication⁷. Hence, assessment of these hard-to-measure factors must be included for training to be effective.

To add to the challenge of measuring beyond tactical and technical skills, most training assessments are input-focused, where assessments are derived post-hoc from data sources that are easily available, rather than from measures defined, a priori, to assess factors necessary to understand the actual abilities of a training audience. Most assessments today are extrapolations of those data (i.e., measure what is available from the data or the sensor), and interpolations of those measures against experience and intuition, not necessarily focused on what must be evaluated against established standards to provide an accurate and comprehensive picture of actual ability. This is particularly true when assessments must be made on factors and capabilities that are deep

⁷ Ma, H., Karri, R., & Chittipeddi, K. (2004). The paradox of managerial tyranny. *Business Horizons*, 4(4), 33-40.

within people (e.g., cognitive skills) or organizations (e.g., team processes which lead to outcomes), or are across a broad, diverse, multi-tiered organization/training audience.

Cohesiveness of training. The question of where (training domain) and how (training platform) to train should be dictated by the skill or competency being developed and not merely the availability of resources. Like the challenge with assessing performance, the where and how of training is often dictated by what is available when and not necessarily how well the platform or the domain can train the skill. For example, communication within a newly-formed team may be initially trained best in a simulator to allow for more controlled communications while students are learning, and targeted assessment of key processes that are difficult to assess out in the field, e.g., who is talking to whom, frequency of communications, flow of communications. Once the newly formed team is proficient in key competencies and skills (e.g., proper use of terms), training may progress to field exercises where newly learned skills can be applied in a more operationally relevant environment. Additionally, as training progresses, rather than generally executing a planned sequence of events, subsequent training should adjust to address areas requiring improvement, and avoiding unnecessary repetition.

Conclusion

Properly structured, sound measures guide observer-evaluators to those key attributes and behaviors that denote demonstrated performance against an established standard of good performance. Augmented by training system generated data that provide both additional measures within actions and activities, and evidence of what actually occurs in events, feedback to training participants become anchored in objectivity. Accordingly, remediation can target those areas in need of improvement that will deliver improved proficiency, as well as enable adjustments in subsequent training for demonstrated high proficiency, vastly improving efficiency and effectiveness in force generation.

The compilation and analysis of data driven objective assessments provide organizational and individual readiness determination based on demonstrated performance, enabling understanding of actual proficiency in operational decision making at any point in force generation. Further data analysis can support institutional understanding of the entire force generation process, inclusive of individual skill development through collective training phases, material and content, methodologies, systems and venues. A cycle of reinforcing assessments based on well-defined measures from data, in contrast to available data dictating what is assessed, can support institutional decision making in seeking optimization of readiness generation and sustainment.

Recommendations and Future Directions

Regardless of methodology, venue or domain, the ever-increasing complexity of military systems and operations point to a need for more objective competency-based construct in training⁸. That goal requires the elevation of measurable learning, both individually and organizationally, over activity accomplishment or a focus on task performance outcomes. Learning should be assessed as an outcome of training, e.g., through assessment of training transfer or retention, and tied directly to required competencies and capabilities that are assessed rather than to tasks completed during training.

The current training paradigm focuses on inputs and resources such as systems and venues, with assessments dominated by vast amounts of data generated from training systems and observer-evaluator subjectivity. A shift

⁸ Colegrove, C. M., & Alliger, G. M. (2002). Mission Essential CompetenciesSM: Defining Combat Mission Readiness in a Novel Way. Paper presented at: NATO Research & Technology Organization, Studies, Analysis, and Simulation Panel, Conference on Mission Training via Distributed Simulation (SAS 38), Brussels, Belgium.

to performance-oriented outcomes based on well-defined measures from both system and observer-evaluator generated data that enables a more comprehensive understanding of individual and organizational proficiency is warranted.

A focus on outputs – what questions need to be answered to properly assess abilities, both individual (e.g., are we deliberately developing leader attributes as well as tactical and technical skills) and collective (e.g., are collaborative decisions sound, does the unit have shared situational awareness regarding Commander’s Intent and expectations, is the unit communicating effectively?) will change the paradigm to orient on unit proficiency. Subsequently, measures needed to answer those questions should be defined a priori, and only then should the inputs or sources of data required to enable measurement and hence assessment of key abilities, be identified. While collecting the right kinds of data, or having the right inputs remains a challenge, it should not drive what is assessed and hence what drives learning. Instead, the learning outcomes should drive what data or inputs are made relevant.

Training systems, venues, and methodologies need major reorientation to accommodate the clearly identified demand in developing and enhancing human performance in every respect, and building adaptability and resilience into individuals, teams and organizations. Optimization of performance is not achievable in a world where objectivity, standardization, sound measures and analysis are absent or are inconsistent and random in training. Optimal performance will only be attained when it is required as an outcome in all dimensions of training, and measured and assessed accordingly.

About the Authors

Major General Melvin Spiese, USMC (ret.)

Major General Spiese was commissioned via the NROTC program after receiving a Bachelor of Science degree in Civil Engineering from the University of Illinois, Urbana-Champaign in May of 1976.

His final assignment on active duty was Deputy Commanding General, I Marine Expeditionary Force, and Commanding General, 1st Marine Expeditionary Brigade. Prior to this assignment, he commanded the Marine Corps' Training and Education Command, and had command of Marine Air Ground Task Force Training Command, Marine Corps Air Ground Combat Center in Twenty Nine Palms, California.

In 2005, Major General Spiese served as the Deputy Director for Strategy, Policy and Assessments at the U.S. European Command (inclusive of Africa). He commanded and served at multiple Marine Corps key training commands, including Commanding General Training Command; Deputy Commanding General Training and Education Command; command of the School of Infantry (East); the Director, Expeditionary Warfare School; and the Director, Tactical Training Exercise Control Group at the Marine Corps Air Ground Combat Center.

Major General Spiese has earned a Master's of Science degree from the University of Southern California, and a Masters of Military Arts and Science degree from the U.S. Army School of Advanced Military Studies. He is a graduate of the Marine Corps Command and Staff College.

Upon leaving active duty, Major General Spiese worked for Cubic Defense Systems as the Vice President for Training Strategy and Effectiveness. He now serves as a Special Advisor to Ravenswood Solutions.

Camilla Knott, Ph.D.

Camilla Knott, Ph.D. is Principal Consultant and Associate Director of Performance Research at TIER1 with expertise in the areas of human cognition, and training and assessment of advanced cognitive skills and leader attributes.

Her recent work has focused on defining training requirements and identifying and implementing novel training methods for developing 21st Century Soldier Competencies (e.g., character and accountability, teamwork and collaboration, problem solving and critical thinking). Dr. Knott's current work also includes the implementation of competency-based training systems and supporting the advancement of training capabilities across various agencies within the Department of Defense including the Air Force, Navy, and the Army. Other efforts led by Dr. Knott have focused on the development of decision support tools for health care applications and other dynamic high-performance environments. She has over 15 years of research experience that spans the development and implementation of research protocols that range from cross-sectional and longitudinal experiments to field studies and survey methods.

Dr. Knott completed a National Research Council post-doctoral fellowship at the Air Force Research Laboratory (AFRL), Wright-Patterson AFB. At AFRL, her work focused on change detection in tactical command and control teams and team decision-making using synthetic task environments.

Dr. Knott holds a Ph.D. in Applied Experimental Psychology from The Catholic University of America.