ITEA Journal 2010; 31: 275–279

Augmenting Test and Evaluation Assessments Using Eye-Tracking and Electroencephalography

Anthony Ries, Ph.D., and Jean Vettel, Ph.D.
U.S. Army Research Laboratory,
Human Research and Engineering Directorate,
Translational Neuroscience Branch, Aberdeen Proving Ground, Maryland

Tools that provide continuous, objective measurements of human-system interactions can augment measures obtained through subjective assessments and/or expert observation by providing near–real time performance metrics. Two tools for the Test and Evaluation (T&E) community will be discussed: eye-tracking applications that are viable for use in T&E today and electroencephalography-based metrics that hold promise for the future.

Key words: Continuous measurements; Electroencephalography (EEG); eye-tracking; mental state; subjective assessments.

System evaluation would be much easier if testers could recruit the ideal operator: someone who never gets fatigued, who maintains a consistent level of concentration on the task, and who can accurately recall their moment-by-moment experience of task difficulty during the testing session. This would ensure that any performance decrements during testing were not a result of the operator being tired, not concentrating, or simply not remembering what happened at a particular point in time. Unfortunately, these operators are hard to find! Instead, evaluators must try and measure the operator's mental state in order to assess how that state affects performance during system interaction, or alternatively, how the system interaction influences mental state, which in turn affects performance. Currently, to evaluate operator mental state, evaluators must rely on self-assessment questionnaires, such as the National Aeronautics and Space Administration Task Load Index (NASA-TLX) (Hart and Staveland 1988), that interrupt the operator at discrete times throughout the testing session to provide an introspective assessment. Not only does the interruption break mental concentration on the task, but self-reports are not sensitive to fluctuations of cognitive state within a task; rather they provide an average subjective estimate over a length of time.

Self-assessment measures are not ideal because they lack objective means to measure particular mental state changes, what influenced the state change, and what happened because of the state change. Consider a scenario where the operator must perform several tasks on a new system, and a self-assessment questionnaire is administered at different intervals throughout the study (i.e., discrete measurements). The operator's fatigue level may fluctuate over the course of the study; however, since individuals must rely on memory to recall past events and are not always accurate in their self-assessment reports, discrete measurements can lead to inaccuracies when trying to capture dynamic mental state fluctuations during the test. The lack of a continuous measure can lead to errors in system evaluation, attributing the decrement in user performance to system design rather than attributing it to changes in the operator's mental state (i.e., level of fatigue). One solution to the problem of tracking performance changes over time is to simply take measurements more frequently; however, this comes with the serious disadvantages of breaking continuity of the task, not to mention the introduction of additional task complexity created by the demands of completing multiple surveys as well as performing the task itself. A more efficient solution to this problem would be the use of tools that permit continuous measurement of task performance, eliminating interruptions for self-assessments.

This article discusses how two measurement tools, eye-tracking and Electroencephalography (EEG), provide continuous measurements related to operator performance without creating task disruption. Eye-tracking provides numerous measures of user eye

Figure 1. Integration of eye-tracking and Electroencephalography (EEG) tools during operator-system interaction. Eye-tracking and EEG provide continuous, objective measurements of operator performance.

movements and visual scanning patterns, while EEG provides a measure of electrical activity in the brain that can be linked to many complex behaviors and operator mental states. These two tools complement one another and may eventually be used effectively together in T&E environments. For example, *Figure 1* shows an operator in a futuristic crew station, wearing a helmet containing EEG electrodes to record brain activity. In addition, a camera mounted within the driving simulation continuously records operator eye measurements. Both of these tools can capture dynamic changes in the operator's mental state throughout the testing session, providing information regarding how the operator interacts with the system interface, without interrupting operator task performance. In this article, potential applications of these two tools to T&E environments will be discussed, with an emphasis on eye-tracking applications that are ready for use today, and EEG applications that provide promise for the future.

Eye-tracking in T&E

Eye-tracking is the process of measuring either the point of gaze (where the operator is looking), or the motion of the eye relative to the head. There are a number of methods for measuring eye movement, including the use of video images from which the eye position is extracted. Advances in computer and video technology have lead to the development of eye-tracking systems that are portable and simple to use (Babcock, Lipps, and Pelz 2002). Eye-tracking offers the evaluator an objective and unobtrusive means to continuously measure human performance in a diverse set of environments and field settings. After a quick calibration, eye-tracking can provide several continuous user measures that can be linked to operator mental state. These measures include blink rate, frequency of eye movement, pupil dilation, the amount of eyelid closure over time (described as Percentage Eye Closure or PERCLOS), and the length of time spent looking at a particular location. Although all of these continuous measures offer advantages, we will highlight just a few of them here to demonstrate how these measures can augment a variety of T&E scenarios today.

Figure 2 shows an example of how an eye-tracker could be used to plot the gaze path of an operator using a crew station computer interface with multiple display screens. In this example, the operator scanned several displays showing urban environments and system status, searching for images of people who could pose a threat to the security of their vehicle. The operator used the touch-screen interface on the center console to complete a threat report anytime a threatening person was seen. The red circles indicate where the operator fixed their gaze on the screen, with the numbers inside the circles indicating the order in which the operator's eyes traveled across the multiple display screens. The size of the red circles describes the relative amount of the time spent looking at each location, with larger circles representing longer gaze fixation times. As can be seen, the operator searched from left to right, starting first with the exit points on the corner building at the left-most display, continuing to the right across the center, and ending at a point between the right-most display and the status display. We can see that the operator spent time looking at the threat report console (circle 3) before scanning the other two buildings (circles 4 and 8). These continuous measures show operator scanning patterns used to perform the task, as well as which system interface features are used most frequently.

Eye-tracking measures have been used successfully to reveal differences between novice and expert operators. Ottati, Hickox, and Richter (1999) compared eye movement patterns of novice and experienced pilots in a flight simulator that required the use of electronic maps for navigation. Traditionally, pilots use printed maps for this task. However, in Ottati's task, the pilots were required to identify critical terrain navigation landmarks from electronic instrumentation. The eye-tracking data revealed search path differences between novices and experts. Unlike the experienced pilots, novice pilots gazed longer and more frequently outside the cockpit instead of at the cockpit instrumentation. In addition, when the novice pilots did look at the mapping instrumentation, they gazed much longer than the experts, suggesting that they struggled to identify the critical landmarks on the electronic map. These differences, revealed by eye-tracking measurements, can be useful for developing training procedures to teach novice operators to use expert-like strategies in reading electronic maps.

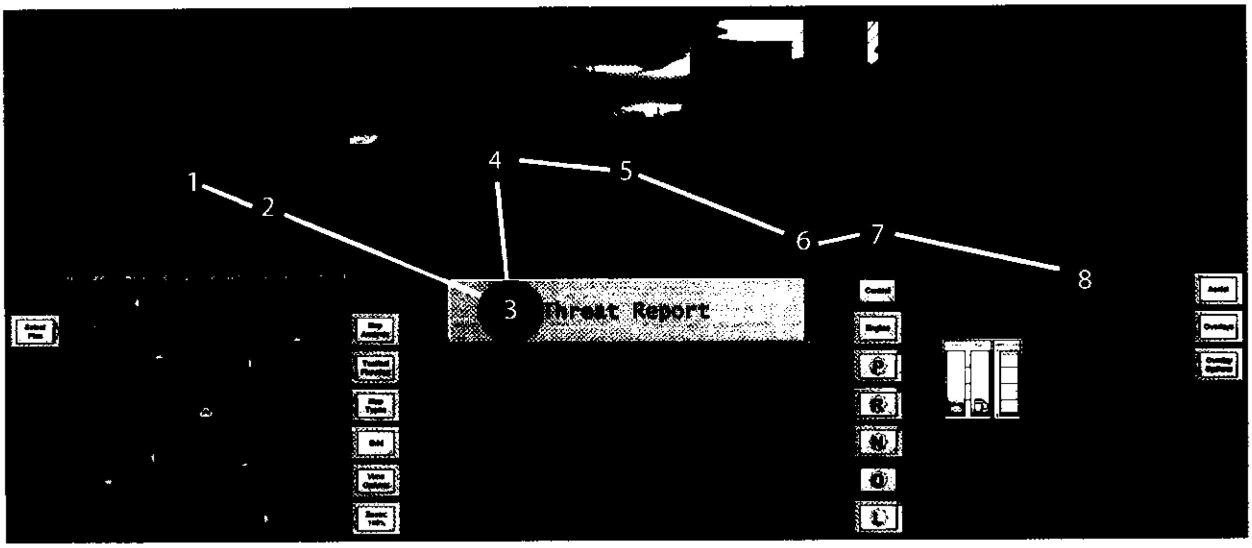

Figure 2. Eye movements superimposed on a crew station computer interface. Numbers indicate the order of eye fixations, and the size of each circle represents the total fixation duration.

Eye-tracking measures such as PERCLOS can help characterize dynamic changes in the operator's mental state owing to factors such as fatigue. In a study by Dinges et al. (1998), operators were intentionally sleep deprived for 42 hours and were then required to monitor a system for target events and to report when the events occurred. The operators completed questionnaires to monitor how sleepy they felt, and an eye-tracking system was used to monitor their visual search. Results of the study indicated that PERCLOS reliably predicted when operators were fatigued more effectively than the questionnaires, and the measure revealed points in time in which operators experienced changes in their level of fatigue. Performance decrements can be linked to the operator's state by capturing these dynamic changes.

This short review only touches the surface of the current-day capabilities of incorporating eye-tracking measures into T&E assessments. The examples presented here highlight potential applications of continuous measures to assess and/or compare how operators interact with system interfaces, as well as monitor dynamic changes in the operator's fatigue level. However, many other applications have been identified in several commercial applications, including the automobile and aviation industries. With their portability and ease-of-use, eye-tracking systems can easily be incorporated into T&E environments to identify additional assessment capabilities.

EEG in T&E

EEG provides a measurement of electrical activity in the brain using recordings from electrodes on the scalp. These measurements have been linked to dynamic changes in behavior and factors related to an operator's mental state. Traditional EEG systems have been very bulky, entailing set-up time to attach electrodes to the scalp with gel, and requiring the operator to minimize any head or body movement while physically tethered to the EEG recording devices. Traditional EEG systems are also highly susceptible to electrical artifacts from nearby equipment and other non-brain sources of electrical activity. System calibration must be performed for each operator before each testing session because the day-to-day variability of operator EEG measurements can be high (East, Bauer, and Lanning 2002). These attributes make traditional EEG systems impractical for current T&E environments; however, in the past few years, EEG systems have been developed that minimize these limitations. These newer systems are light-weight, often incorporating the electrodes into a hat or helmet, and are designed for use in real-world, operational environments. They contain advanced amplification and wireless transmission technology to minimize the impact of electrical interference. One such system is shown in *Figure 3*. As technology develops, these newer EEG systems will likely evolve to be as portable and easy-to-use as the eye-tracking systems in use today.

Results from EEG-based measurements collected in controlled environments show potential for application in operational settings. For example, Pope, Bogart, and Bartolome (1995) utilized a real-time index of operator task engagement, based on the power of EEG spectra, in an adaptive system to mitigate the effects of fatigue. EEG was collected while operators performed several

Figure 3. An example of an electroencephalography system designed to be worn under a helmet. A wireless base station receives and decodes the neural signals from several meters away, allowing the operator to be fully mobile. While this device shows promise in controlled environments, there is still work to be done to increase reliability in operational settings.

tasks including monitoring, resource management, and compensatory tracking. The tracking task was either automated by the system or manually controlled by the operator, and the system dynamically switched between the two modes based on the changes in the EEG-based index in order to minimize the effects of fatigue.

Similarly, Wilson, Lambert, and Russell (2000) also used an EEG-based measure in an automated system that adapted its functioning based upon changes in operator mental state. As in the Pope, Bogart, and Bartolome (1995) study discussed above, operators were required to perform multiple tasks. When the operator could not complete all assigned tasks, the system assisted with the monitoring task, requiring operators to solely focus on the remaining two tasks—resource management and tracking. The system switched into this assist mode only when the EEG-based measure of operator state indicated that they were overloaded and struggling to complete all tasks simultaneously (referred to as "high workload" by the authors). Of importance, by using the EEG measure to dynamically adapt the system, the number of errors decreased significantly in both the resource management task (33 percent) and in the tracking task (44 percent).

Other EEG findings obtained in controlled, laboratory settings show promise for future application in T&E environments. For instance, EEG-based measures have revealed neural signatures that are sensitive to error detection. The error-related negativity (ERN) produces a distinct pattern of brain activity when an individual makes an incorrect response or error. Monitoring operator errors in near–real time would allow testers the ability to identify when certain errors occur during system interaction as well as the ability to adapt system components based on errors committed. The ERN is not only sensitive to when an operator makes a known error, but it is also seen when an operator witnesses someone else making an error (van Schie et al. 2004). Preliminary applications of the ERN are currently being employed during the testing of brain-computer interfaces (Ferrez and Millan 2008). The ability to detect operator errors in real time could substantially augment T&E assessment capabilities.

Another laboratory-based finding demonstrates the utility of EEG for detecting predefined targets or identifying sudden changes in the environment without relying on overt responses from the operator. Using EEG in this way may provide testers with the ability to detect anomalies or unexpected changes that occur during system interaction. One instance of using this approach comes from image classification where an operator must detect a target of interest embedded in a series of rapidly presented images. Systems have shown their ability to accurately identify particular images for further analysis based on EEG measurements alone even though the images were shown for 100 milliseconds (Gerson, Parra, and Sajda 2006). This approach proved to be highly accurate in discriminating between target and nontarget images, and much faster than simply relying on self-reports from the operator performing the target-scanning task. The potential of EEG to rapidly identify what information is critical to the task at hand could greatly enhance how operator performance and system evaluation are analyzed in T&E environments.

These examples highlight the potential of EEG-based measures to index factors related to operator mental state. The examples described here are just a subset of the laboratory-based EEG findings that may have direct relevant applications to T&E. EEG-based measures can greatly augment the types of continuous operator assessments currently available with eye-tracking, and the combined use of both eye-tracking and EEG measurements may improve real-time assessment of operator mental states across many tasks, systems, and environments above and beyond using each method alone.

Conclusions

Eye-tracking and EEG are two tools that provide continuous measurements of operator performance, and both provide powerful analysis tools to the T&E community. Eye-tracking systems are simple to use, portable, and provide measurements related to operator mental state, such as blink rate and PERCLOS, that can be applied in a T&E setting today. Although they are not ready for use in all field settings, EEG systems have the potential to provide additional measures related to operator mental state that may be of use to testers and evaluators. Studies using EEG have successfully adapted systems and improved operator performance, and other more preliminary studies using laboratory-based EEG measures such as the ERN show how real-time assessment of operator performance can be augmented in applied settings. As additional research is conducted, EEG-based measures obtained in controlled settings can be applied to the complex and dynamic environments of T&E. Future research should explore the combined use of eye-tracking and EEG systems to create a broad-based measure of changes in operator performance, expanding assessment capabilities to a diverse set of tasks and environments. ❑

DR. ANTHONY RIES is a research psychologist at Aberdeen Proving Ground in Aberdeen, Maryland. He is currently serving as deputy manager of the "High-Definition Cognition in Operational Environments" U.S. Army Technology Objective (Research), which is focused on enabling assessment of soldier performance in operationally relevant environments using behavioral and physiological measures. Dr. Ries received a bachelor of science degree in psychology from Northwest Missouri State University in 2000 and master of arts and doctor of philosophy degrees in cognitive psychology from the University of North Carolina at Chapel Hill in 2003 and 2007, respectively. E-mail: Anthony.ries@us.army.mil

DR. JEAN VETTEL is employed by the U.S. Army Research Laboratory in the Translational Neuroscience Branch at Aberdeen Proving Ground in Aberdeen, Maryland. Her research focuses on improving soldier–system performance by conducting neuroscience research in complex, operational settings. She received a bachelor of science degree from Carnegie Mellon University in 2000 and master of science and doctor of philosophy degrees from Brown University in 2008 and 2009, respectively. E-mail: Jean.vettel@us.army.mil

References

Babcock, J., M. Lipps, and J. B. Pelz. 2002. How people look at pictures before, during, and after image capture: Buswell revisited. In *Proceedings of SPIE, Human Vision and Electronic Imaging*, 4662:34–47. Bellingham, WA: SPIE.

Dinges, D. F., M. Mallis, G. Maislin, and J. W. Powell. 1998. Evaluation of techniques for ocular measurement as an index of fatigue and the basis for alertness management. Department of Transportation Highway Safety, pub. 808 762.

East, J. A., K. W. Bauer, and J. W. Lanning. 2002. Feature selection for predicting pilot mental workload: A feasibility study. *International Journal of Smart Engineering System Design* 4: 183–193.

Ferrez, P. W., and J. del R. Millan. 2008. Error-related EEG potentials generated during simulated brain-computer interaction. IEEE Trans Biomedical Engineering, March, 923–929.

Gerson, A. D., L. C. Parra, and P. Sajda. 2006. Cortically-coupled computer vision for rapid image search. IEEE Trans Neural Systems and Rehabilitation Engineering, June 1–6.

Hart, S. G., and L. E. Staveland. 1988. Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In *Human mental workload*, ed. P. A. Hancock and N. Meshkati. Amsterdam, The Netherlands: Elsevier.

Ottati, W. L., J. C. Hickox, and J. Richter. 1999. Eye scan patterns of experienced and novice pilots during Visual Flight Rules (VFR) navigation. In *Proceedings of the Human Factors and Ergonomics Society, 43rd Annual Meeting*, Houston, Texas. Santa Monica, CA: Human Factors and Ergonomic Society.

Pope, A. T., E. H. Bogart, and D. Bartolome. 1995. Biocybernetic system evaluates indices of operator engagement. *Biological Psychology* 40: 187–196.

van Schie, H., R. Mars, M. Coles, and H. Bekkering. 2004. Modulation of activity in medial frontal and motor cortices during error observation. *Nature Neuroscience* 7: 549–554.

ITEA Journal 2010; 31: 280–283

Leader Development by Design

Robert A. Cassella
Booz Allen Hamilton, Leavenworth, Kansas

U.S. Army doctrine has recently embraced a "methodology for applying critical and creative thinking" to develop approaches to solve complex, ill-structured problems prior to the initiation of detailed planning. The just-published Army Capstone Concept advocates "a mindset based on flexibility of thought" by leaders who "have a tolerance for ambiguity, and possess the ability and willingness to make rapid adjustments according to the situation." Who are these people, and how will they be developed? While much has been done to produce the doctrine, as much thought must be given to how the prototypical leader of the twenty-first century will be developed and assessed in some set of "vital few" professional competencies.

Key words: Cognition; flexibility of thought; operational adaptability; visualization.

It seems that everyone agrees—the operational environment of the recent past was undeniably complex, today's operational environment is even more complex, and the forecast for the future is no less complex. In the postulated future, the U.S. military will face a daunting array of state and nonstate adversaries armed with everything from antiquated Kalashnikovs to horrific weapons of mass destruction, and the locus of conflict will range from subterranean lairs to outer space and cyberspace.

Military operations will continue to encompass everything from peacetime engagement to humanitarian assistance to irregular and general war. In today's and tomorrow's missions, the salient difference will be the simultaneous nature of operations and the frequency of transitions from operation to operation. This modern phenomenon is what former Marine Commandant, General Charles Krulak,[1] referred to as "three-block war."

> *"In one moment in time, our Service members will be feeding and clothing displaced refugees, providing humanitarian assistance. In the next moment, they will be holding two warring tribes apart—conducting peacekeeping operations—and, finally, they will be fighting a highly lethal mid-intensity battle—all on the same day...all within three city blocks."*

In recognition of this reality, General Martin Dempsey, the current commander of the U.S. Army Training and Doctrine Command, prefaced the most recent Army Capstone Concept with these words: "...[t]he Army must...achieve...operational adaptability...a mindset based on flexibility of thought calling for leaders at all levels who are comfortable with collaborative planning and decentralized execution, have a tolerance for ambiguity, and possess the ability and willingness to make rapid adjustments according to the situation..." General Dempsey went on to emphasize that "[t]he training and education of our entire force must aim to develop the mindset and requisite knowledge, skills, and abilities required to operate effectively under conditions of uncertainty and complexity."[2]

Correspondingly, the 2010 Department of Defense quadrennial defense review states that "...part of our commitment [is] to ensure that tomorrow's leaders are prepared for the difficult missions they will be asked to execute..."[3] Other sources postulate what capabilities will be needed for twenty-first century leaders: innovativeness, creativity, agility, adaptability, critical thinking, versatility—without advancing a method to produce what former Army Chief of Staff General Peter Schoomaker repeatedly described as "soldier-pentathletes."[4]

Recently, Army doctrine has embraced an approach to solve complex, ill-structured problems prior to the initiation of detailed planning. This approach requires practitioners who will endeavor to understand environments populated by diverse actors to define the problem and, ultimately, design an approach to resolve it.[5] Other doctrinal literature prescribes the preferred method of battle command—through decentralized execution—based on orders that emphasize the results to be attained, not how they are to be accomplished.[6]

Table 1. The 2008 Battle Command Battle Lab study of cognition and visualization identified 19 KSAEs grouped into four competency areas.

Domain knowledge	Decision making	Communications	Adaptability
Doctrine	Red teaming	Oral expression	Metacognition
Tactics/operational art	Critical thinking	Written expression	Tolerance for ambiguity/uncertainty
Systems and processes	Fusion	Instructor	Bias recognition
System capabilities/limitations	Assessment/analysis	Interpersonal communication	Attention management
Observer/controller	Second-in-command		Operations and training

This doctrine of mission command demands operators whose appreciation of the situation guides the adaptive employment of forces. Decentralized execution also requires leaders who are willing to act in the absence of orders when existing orders no longer fit the situation or when unforeseen opportunities or threats arise. Although the discussion of decentralized execution usually implies a cogent choice, in the reality of a degraded command and control environment, adaptive leaders will be forced to self-synchronize their actions within the commander's intent and the mission variables to achieve their objective.

All of these requirements for the development of future leaders will certainly necessitate changes to how they are accessed and matured from precommissioning through the gamut of professional military education. Anyone who has been involved in managing the curricula of Service schools recognizes that there is no lack of good ideas, but there is a paucity of available time—first, the amount of time parceled out for leaders to spend away from the operating force and, second, the "zero-sum-game" nature of scheduling leader development opportunities. In this environment, how can leaders receive the education they need to refine their strengths and overcome shortcomings in some number of crucial knowledge, skill, and ability sets?

And it's not just about military professionals—during UNIFIED QUEST 2009, participants generally agreed that, rather than emphasize processes, "...the professional development of civil-military designers and planners should be the critical focus." The UNIFIED QUEST 2009 final analysis report goes on to say that because the development of a single multiagency methodology or process is unlikely and the common ground between government agencies is people, a common education for civilian and military designers and planners would provide a unifying factor between government agencies and the Department of Defense.[7]

One effort to identify requisite knowledge, skills, abilities, and experiences for future full spectrum operators was the 2008 study commissioned by the U.S. Army Battle Command Battle Lab at Fort Leavenworth. The purpose of this study was to identify the cognitive skills required to visualize full spectrum operations; it was conducted in conjunction with the annual OMNI FUSION experiment to inform the evolution of battle command doctrine and the development of future leaders.

The exercise of battle command comprises three interdependent components: (a) understanding, visualizing, and describing a situation; (b) directing and leading forces; and (c) assessing operations—all so as to impose one's will upon an adversary. These component parts of battle command are fundamental for all leaders at all levels in all operations and are generalizable across the depth and breadth of military and civilian leadership.

The results of this research into cognition and visualization coalesced 19 knowledge, skills, abilities, and experiences (KSAE) into four interconnected competencies of domain knowledge, decision making, communications, and adaptability *(Table 1)*.

The study projected that some level of expertise across the identified competencies could produce results that contribute to visualization, shared understanding, collaborative teaming, self-synchronization, self-awareness, sense making, creativity, insight, and agility.

To complete the articulation of these core competencies, the study proposed novice, journeyman, and master level behaviors associated with the competencies that can be used for assessment *(Table 2)*. Because the output, or results, of a competency (adaptability, for example) may be observed and judged, a person could be assessed to be performing well or not so well with respect to a given competency. It is not reasonable, however, to generalize that the person has mastered the depth and breadth of any given competency. Thus, evidence of expertise in a competency is demonstrated by actions (behaviors) that can be observed and assessed to distinguish the outstanding performers from the average.

The study concluded that mastery of these competencies would allow leaders to artfully exploit the capabilities of the human and technological subsystems

Table 2. The behaviors for adaptability illustrate the difference between novices, journeymen, and masters with regard to their tolerance for ambiguity or uncertainty.

Competency = adaptability	Novice behaviors	Journeyman behaviors	Master behaviors
Tolerance for ambiguity/uncertainty	Tolerates risk and uncertainty.	Handles risk and uncertainty comfortably.	Rises to the challenge, thrives on situations involving risk and uncertainty.
	Decides and acts without having a sense of the total picture.	Uses ingenuity to compensate without having a sense of the total picture.	Uses ingenuity in dealing with ambiguous situations, and guides others to cope effectively.
	Copes with change and transitions when necessary.	Effectively copes with change and transitions comfortably.	Anticipates impact of change; plans how to transition and guides others in transitions.

of the command and control system to achieve operational results.[8]

There are obstacles to producing the prototypical "pentathlete"—most notably, the cultural resistance to developing innovative, critically thinking devil's advocates who challenge conventional thinking. Like any significant cultural change in any military service, acceptance of an innovative approach to leader development will require doctrine—the engine of change—and a four-star champion for the leader development programmatics that are necessary to implement the idea of operational adaptation.

The development of tactical, operational, and strategic leaders who can realize and exploit the idea of operational adaptability cannot be deferred until—or contained within—command and staff college level education. The process of growing the generation of leaders who embody the "right stuff" must begin at entry level and continue in and out of professional military education opportunities.

Another challenge would be to overcome the cultural artifact that abhors the idea that producing better warfighters means longer incubation in an educational venue. If future uses of military force will demand more than just warfighting skills, there should be a means to classify requirements for an individual's development within the institutional environment and through their self-development.

There are a number of assessment instruments currently in use throughout the military—primarily used for specialized applications like special operations selection—like the 16 Personality Factor, Wonderlic Personnel Intelligence Test, the Minnesota Multifacet Personality Inventory, and the Individual Adaptability Measure. These tools can be exploited for *classifying*—not *selecting*—entry-level leaders in areas that contribute to the development of the cognitive capacities that enable acquisition of the knowledge and skills, as well as the refinement of innate abilities, required of the objective twenty-first century leader. The results of this classification effort, which would be akin to an Armed Services Vocational Aptitude Battery for leaders, could facilitate personalized development. Emphasis on the individual avoids the tendency to pursue a "sheep dipping" approach to professional growth in which every person receives the same treatment regardless of their strengths and weaknesses.

By systemically determining desired results (such as visualization, collaborative teaming, self-synchronization, self-awareness, sense making, creativity, and agility), we could derive sets of requisite knowledge, skills, abilities, and experiences that contribute to their attainment. Once the required knowledge, skills, and abilities of the objective twenty-first century leader are determined, based on future operational demands, behaviorally based performance assessment tools can be developed and exploited that indicate some level of proficiency (novice, journeyman, master).

The cost of producing a generation of full spectrum warriors will be resource-intensive—the least of which is not time—but these costs can be mitigated by personnel policy changes that require altering the current concept of a "normal" career in content and duration. These changes might include extending time-in-grade gates and redefining the minimum retirement threshold to 25 or 30 years.

Other necessary policy revisions would have to be made to separate developmental assessment and feedback and to the winnowing process that determines which officers will be promoted and selected for positions of greater responsibility. The current Army evaluation reporting system, contained in Army Regulation 623-3, is designed to "select and develop tomorrow's leaders." This bifurcated purpose is at odds with itself—the evaluation narrative that supports selection for schooling, promotion, and command will probably not be developmental in nature and any suggestions for an officers' personal or professional development probably won't get them selected for career advancement.

This conflict of purpose might be mitigated by a solution similar to the Multi-Source Assessment and

Feedback program that provides "...a 360-degree approach...*without ties to Army personnel management processes or systems* [emphasis added]." This leadership assessment tool specifically addresses nine leader competencies such as leads by example, develops leaders, and creates a positive environment. As designed and executed, the Multi-Source Assessment and Feedback program provides feedback to the assessed individual and the chain of command and to the institutional education system, albeit anonymously and in aggregate as formative data to provide command climate insights in units and for the refinement of curricula in service schools.[9]

Another area where improvements must be made is in the delivery of professional military education; leaders must be educated to cope with unstructured problems rather than being given both the problem and all of the attendant information. It is not practical to teach the rudiments of critical thinking in one classroom, for example, and then eschew critical examination of assumptions and the thought underlying these assumptions in an adjacent classroom—and expect students to understand how to solve complex problems or to thrive in an environment that demands decentralized execution.

Depending upon an individual learners' developmental needs, some will need abstract, theoretical explanations while others need concrete, procedural explanations; most will probably need a combination of both approaches. Without a definitive assessment of these needs, we will be forced into applying the same old solution to novel situations and thus defer to the default position rather than leader development by design.

Robert A. Cassella is a consultant with Booz Allen Hamilton in Leavenworth, Kansas, currently supporting the U.S. Army Training and Doctrine Command Analysis Center at Fort Leavenworth. He is the author of Project Management Skills for Kids *(Trafford, 2005). E-mail: cassella_robert@bah.com*

Endnotes

[1]Charles Krulak, 1997. *The Three Block War: Fighting in Urban Areas, an Address to the National Press Club*, October 10, 1997.

[2]U.S. Army Training and Doctrine Command, "The Army Capstone Concept: Operational Adaptability—Operating under Conditions of Uncertainty and Complexity in an Era of Persistent Conflict," December 21, 2009, i.

[3]Department of Defense, "Quadrennial Defense Review Report," February 2010, 54.

[4]Peter Schoomaker, "Farewell Message," April 6, 2007.

[5]Department of the Army, "The Operations Process," February 2010, Chapter 3.

[6]Department of the Army, "Operations," February 2008, 3–6.

[7]U.S. Army Training and Doctrine Command Analysis Center at Fort Leavenworth, "UNIFIED QUEST 2009 Final Analysis Report," July 31, 2009.

[8]U.S. Army Battle Command Battle Lab at Fort Leavenworth, "Cognition and Visualization Study Report," September 26, 2008.

[9]Multi-Source Assessment and Feedback Web site, downloaded from https://msaf.army.mil/Default.aspx (accessed on February 21, 2010).