A Behaviorally Anchored Assessment Tool to Measure Tactical Thinking Proficiency

Jennifer K. Phillips, Jennifer Shafer,
Karol G. Ross, Ph.D , Donald A. Cox.
Klein Associates Inc.
Fairborn, OH
jenni@decisionmaking.com
jshafer@decisionmaking.com
karol@decisionmaking.com
donald@decisionmaking.com

Scott Shadrick
US Army Research Institute
Ft. Knox, KY
scott.shadrick@us.army.mil

ABSTRACT

There is an ongoing need in the Army to enhance soldiers' and leaders' battlefield thinking skills, especially as Future Combat Systems technologies are introduced. To improve cognitive task performance, soldiers and leaders must engage in deliberate practice in context-rich environments, including training scenarios and simulations, field exercises, and actual combat situations. An important aspect of training tactical thinking skills that still needs development, however, is assessment. How do we know that cognitive skills are improving across experiences and over time? Diagnostic assessment will allow us to evaluate training effectiveness and tailor training as expertise develops.

The goal of this project was to develop a technique to assess tactical thinking skills. In previous projects, we created a Behaviorally Anchored Rating Scale for tactical thinking (T-BARS) based on eight themes. The initial T-BARS attempted to capture the progression of tactical skill development, but was developed for limited internal use. The goal of this project was to develop the T-BARS by expanding and validating it to make it more robust and broadly useful. We collected feedback from combat arms Captains, Majors, and Lieutenant Colonels to vignettes provided by the U.S. Army Research Institute. The vignettes were developed as part of their Think Like a Commander training program. We used these responses to improve the T-BARS by adding behavioral anchors, clarifying and proportioning the scale for each of the eight themes, and iteratively testing the T-BARS to reach a goal of 80% inter-rater agreement. The final product was a tool designed to accurately rate the development and quality of tactical thinking. T-BARS in its validated form can be used both as a diagnostic assessment and as a methodology for training effectiveness evaluations.

ABOUT THE AUTHORS

Jennifer K. Phillips is a Senior Research Associate at Klein Associates. Her research interests include skill acquisition and the nature of expertise, and she has applied her research to the development of decision-centered training interventions. Ms. Phillips served as project lead on an Army-sponsored effort to develop decision skills training for small unit leaders in military operations in urban terrain, and worked on a similar effort to provide web-based decision skills training to Air Force personnel for Operations Other Than War. Ms. Phillips has worked on several USMC-sponsored efforts to re-engineer existing command posts and design experimental combat operations centers. She led an Army-sponsored program of research to identify the process by which individuals make sense of situations as they unfold, and to develop training designed to bolster sensemaking skills. Most recently she has been involved in a USMC program of research to improve the use of Tactical Decision-Making Simulations for training decision-making and sensemaking skills. She is experienced in using Cognitive Task Analysis methodologies to capture components of expertise in a range of domains, and has used Cognitive Task Analysis to identify training requirements and for purposes of knowledge management. Ms. Phillips received a B.A. in Psychology from Kenyon College in 1995.

Jennifer Shafer is a Research Analyst at Klein Associates. Her research interests include culture and cognition, information operations, and knowledge representation techniques. She is currently supporting a project to create training and evaluation simulations. Ms. Shafer holds an M.A. in Neuroscience from Johns Hopkins University, Baltimore, MD, and B.S. degrees in Psychology and Biology from the University of Washington, Seattle, WA.

Dr. Karol G. Ross is a Senior Research Associate at Klein Associates. Her areas of expertise are the training and assessment of tactical thinking skills and qualitative research methods. She conducts research and development for the US Army, the USMC, and the Office of Naval Research. She has recently directed and participated in research to produce an online tool to support subject-matter experts in developing training vignettes and Cognitive Task Analysis for the design of anti-terrorism training. Currently, she is directing and conducting research to develop a new assessment method for tactical thinking skills, online training scenarios for coalition warfare, and knowledge management processes and vignettes for the US Army's Battle Command Knowledge System. She has developed and conducted workshops on qualitative research methods for the military and industry. Formerly, as a Research Psychologist for the U.S. Army Research Laboratory, she conducted research and development for training adaptive tactical thinking for digitized Brigade operations at the US Army Command and General Staff College. She also previously served as a Senior Researcher for BDM International studying division-level battle staff performance, and as a Research Psychologist for the Army Research Institute in Germany working in the area of large-scale simulation training for echelons above Corps. In addition, Dr. Ross has eight years of experience working in the healthcare industry and ten years of experience as an adjunct faculty member at the graduate level in education, training, and organizational development. She regularly publishes and presents in the areas of naturalistic research and training. She earned her doctoral degree from the University of Tennessee in Experimental Psychology in 1984.

Donald A. Cox is a Research Associate at Klein Associates. Mr. Cox has been working with computers for over 20 years and has long standing interests in design in all its guises and qualitative methods for informing design. At Klein Associates, Mr. Cox has worked on projects studying the assessment of tactical thinking development, and investigating the cognition of computer hackers and intelligence analysts. Prior to joining Klein Associates, Mr. Cox worked at InContext Enterprises understanding work practice and designing appropriate innovative responses that included new website structure, web applications and e-learning. Before that, he worked for IBM Canada as a User Centered Design specialist. Previous to that, as a developer at IBM Canada, he worked in teams that shipped middleware and development environment products. Mr. Cox holds a M.Sc. in Computer Science from the University of Calgary, Calgary, Alberta, and a B.Sc. in Computing and Information Science from Queen's University, Kingston, Ontario. He has presented his research at the ACM SIGCHI and ACM CSCW conferences.

Scott Shadrick is a Research Psychologist for the US Army Research Institute for the Behavioral and Social Sciences at Fort Knox. He has conducted research on adaptive performance, decision making, distributed training, computer-based learning, learning systems design and evaluation, and leader development.

A Behaviorally Anchored Assessment Tool to Measure Tactical Thinking Proficiency

Jennifer K. Phillips, Jennifer Shafer,
Karol G. Ross, Ph.D., Donald A. Cox
Klein Associates Inc.
Fairborn, OH
jenni@decisionmaking.com
jshafer@decisionmaking.com
karol@decisionmaking.com
donald@decisionmaking.com

Scott Shadrick
US Army Research Institute
Ft. Knox, KY
scott.shadrick@us.army.mil

INTRODUCTION

Assessment is a key component to every training intervention. Without evaluating a training package it is impossible to know whether it is producing meaningful improvement and desirable outcomes in students' knowledge, skill, and performance. Training assessment has largely focused on capturing changes in a student's declarative and procedural knowledge, and testing techniques for that purpose are well-refined. However, the training community continues to address the challenge of assessing training interventions that target thinking skills. By "thinking skills" we refer to the higher-order cognitive functions such as decision making, sensemaking, and the underlying cognitive processes that support those functions. In complex, ill-structured domains[1] such as tactical thinking, medical diagnosis and treatment, and law enforcement, it is insufficient to rely on rote procedures and factual knowledge to perform successfully. Operators require declarative and procedural knowledge as foundations, but different situations in these domains are likely to require application of varying patterns of principles, even in cases of seemingly similar problems or goals. No standard solutions can be employed with regularity. Such domains require professionals to exercise a great deal of judgment to flexibly apply their knowledge. These professions can also require that decisions be made under conditions of time pressure and high stakes. Well-developed "thinking skills" are critical for high levels of performance.

The research program described in this paper addresses the ongoing need in the Army for the enhancement of officers' tactical thinking skills. During their careers, officers amass an impressive command of declarative knowledge and procedural information, but this does not automatically lead to knowing how to make decisions and understanding situations during performance. To improve cognitive task performance, officers must engage in deliberate practice in context-rich environments, including training scenarios, simulations, and field exercises, to prepare for actual combat situations. Deliberate practice is training that is structured to provide an opportunity to develop specifically targeted skills by practicing them and receiving feedback on performance. Previous research has shown that tactical thinking skills can be deliberately practiced and improved (e.g., Lussier, Ross, & Mayes, 2000; Lussier, Shadrick, & Prevou, 2003; Ross & Lussier, 1999; Ross, Lussier, & Klein, 2004).

While training applications have been developed to give officers deliberate practice in tactical decision making, the assessment component required to measure success can be improved. How do we know that thinking skills are improving across experiences and over time? How can we assess an officer's performance in order to target training deliberately to the aspects of expertise he or she most needs to

[1] "An ill-structured knowledge domain is one in which the following two properties hold: (a) Each case or example of knowledge application typically involves the simultaneous interactive involvement of multiple, wide-application conceptual structures (multiple schemas, perspectives, organizational principles, and so on), each of which is individually complex (i.e., the domain involves concept- and case-complexity); and (b) the pattern of conceptual incidence and interaction varies substantially across cases nominally of the same type (i.e., the domain involves across-case irregularity)" (Spiro, Feltovich, Jacobson, & Coulson, 1992, p. 60). In other words, each individual case is complex, and there is considerable variability across cases.

develop? Training professionals and instructors have little to no guidance for assessing or diagnosing aspects of the learner's cognitive proficiency as part of implementation. This paper documents an effort to develop an assessment technique for tactical thinking skills that is not dependent on expert judgments or self-report, and that can be applied across training experiences both in a diagnostic manner and as a method to assess the impact of tactical thinking training.

BEHAVIORALLY ANCHORED RATING SCALES

The challenge in assessing cognitive proficiency is evaluating processes that cannot be seen or easily documented. Cognition is invisible. However, behaviors are an extension of an individual's cognition, and they are visible. A few assessment techniques exist that purport to measure cognition through observable behaviors. Behaviorally Anchored Rating Scales (BARSs) are one such tool. Previous research efforts have examined the use of a BARS assessment technique in the tactical thinking domain and have indicated that it holds promise as a tool for measuring individuals' tactical thinking mental models (Phillips, Shafer, Ross, Baxter, & Harris, 2003; Ross, Battaglia, Hutton, & Crandall, 2003).

Behaviorally anchored rating scales have traditionally been used in organizational settings to measure the effectiveness of individuals performing a wide variety of tasks (Muchinsky, 2003). A typical BARS lists observable behaviors that correspond to a numeric score, with higher numbers indicating more advanced behaviors. BARSs typically utilize five performance points. Thus, a '1' on the scale would correspond to a list of behaviors exhibited by beginners or low-level performers, and a '5' would be tied to a set of behaviors exhibited by the most proficient performers in the domain. To construct each scale, work is observed and/or incidents about work are gathered from subject-matter experts (SMEs). These incidents are placed along a range from poor to excellent. Each rating is given a name, description, and example of typical performance at that level. Once a BARS is developed for a particular task or job position, individuals without domain experience or expertise are able to rate performance by assigning scores to behaviors they observe. The *Code of Best Practices for Experimentation* for the Department of Defense (Alberts & Hayes, 2002) promotes BARSs as a means of conducting performance assessment without continued reliance on SMEs. Raters using BARSs are less prone to common rater biases associated with the use of other rating scales, such as the halo effect and positive leniency (Muchinsky, 2003; Riggio, 2000).

Behaviorally anchored rating scales offer a promising technique for assessing tactical thinking skills because they allow evaluation of invisible cognitive processes by operationalizing them as overt behaviors. They provide a means of judging cognitive proficiency without being an expert, allowing a layman to accurately judge students' performance and improvement. We set out to develop a tactical thinking BARS, or T-BARS, that could be utilized by researchers and training professionals interested in measuring the cognitive proficiency of military tacticians. The T-BARS tool is intended to provide insights as to the tactician's current stage of mental model development, enabling feedback on the effectiveness of training interventions and guidance as to additional training needs of the individual.

DEVELOPMENT OF THE T-BARS

Development of the T-BARS assessment tool built upon previous research conducted with military commanders to define the nature of their command skill, and research defining stages of development for cognitive proficiency. In this section we summarize these findings and models, and then describe the additional research conducted under this effort to generate the T-BARS tool.

Think Like A Commander

The Think Like A Commander (TLAC) research program (Lussier, 1998; Lussier et al., 2003) defined eight "themes" of tactical thinking that expert commanders are thought to use. The themes were derived from interviews with numerous tactical experts (Deckert, Entin, Entin, MacMillan, & Serfaty, 1996). A TLAC program of training was subsequently developed with the goal of training soldiers and leaders to be better adaptive thinkers by becoming proficient at the eight themes during deliberate practice. The TLAC training is currently in use at Fort Knox in the Armor Captain's Career Course, and in the Reserve Component Armor Captain's Career Course as a distance learning application. We used the eight TLAC themes as the basis for the initial version of the T-BARS. They include:

Keep a Focus on the Mission and Higher's Intent (Mission). This theme refers to the need for leaders to always stay aware of the higher purpose and results

they are directed to achieve. Even when unusual and critical events may draw them in a different direction, it is essential to stay focused on the overall mission.

Model a Thinking Enemy (Enemy). The focus of this theme is on the importance of remembering that the adversary is a reasoning human being who is intent on defeating friendly forces. Although it's tempting to simplify the battlefield by treating the enemy as static or simply reactive, this will harm the troops' ability to fight an effective battle.

Consider Effects of Terrain (Terrain). This theme reflects the importance of not losing sight of the operational effects of the terrain on which the troops must fight. Every combination of terrain and weather has a significant effect on what can and should be done to accomplish the mission.

Know and Use All Assets Available (Assets). This theme refers to the necessity for leaders to maintain awareness of the synergistic effects of fighting their command as a combined arms team. Leaders should consider not only all assets under their command, but also those which higher headquarters might bring to bear to assist them.

Consider Timing (Timing). The focus of this theme is on the importance of being cognizant of the time available to get things done. A good sense of how much time it takes to accomplish various battlefield tasks and the proper use of that sense is a vital combat multiplier.

See the Big Picture (Big Picture). This theme refers to the importance of maintaining awareness of what is happening in the environment and how it might affect operations—what courses of action can affect others' operations. A leader with a narrow focus on his or her own fight can easily be blind-sided.

Consider Contingencies and Remain Flexible (Contingencies). Small unit leaders must never lose sight of the old maxim that "no plan survives the first shot." Flexible plans and well thought out contingencies result in rapid, effective responses under fire. Contingencies are characterized by thinking that begins with questions like "What if…?" or "How else can I...?"

Visualize the Battlefield (Visualize). Leaders must be able to visualize a fluid and dynamic battlefield with some accuracy and use this visualization to their advantage. A leader who develops this difficult skill can reason proactively like no other.

Lussier and his colleagues generated general descriptions of the nature of performance along each of the eight themes as skill improves (Lussier, 1998). For example, related to the *Mission*, inexperienced tacticians tend to focus narrowly on their own mission only. Highly experienced individuals, on the other hand, consider the objectives of the larger unit and are able to conduct their mission in a manner that supports the higher intent. Lussier's general descriptors provided the initial basis for the T-BARS.

The Dreyfus Model of Skill Acquisition

Dreyfus and Dreyfus' (1986) stage model of skill acquisition provided the theoretical foundation for the T-BARS assessment tool. The model characterizes five performance levels that individuals go through as they gain skill and proficiency in cognitively complex domains: novice, advanced beginner, competent, proficient, and expert. The model has been applied to training and instruction within domains such as combat aviation, nursing, industrial accounting, psychotherapy, and chess (Benner, 1984; 2004; Dreyfus & Dreyfus, 1986; Houldsworth, O'Brien, Butler, & Edwards, 1997; McElroy, Greiner, & de Chesnay, 1991).

Novices have limited or no experience in situations characteristic of their domain. They exhibit rigid adherence to rules they have been taught, or plans they have been given. They have little situational perception, and they lack the basic domain knowledge needed to perform analysis.

Advanced beginners have enough domain experience that their performance is marginally acceptable. They have a sufficient knowledge base with which to perform analysis in a situation. At this stage they are able to recognize recurring, meaningful "aspects" of situations. Aspects are defined by Dreyfus and Dreyfus (1986) as global characteristics identifiable only through prior experience, where the prior experience provides a comparison case for the current situation. An advanced beginner's knowledge base regarding aspects and attributes of situations enables them to develop their own guidelines for action. However, at this stage all components of the situation tend to be treated as independent pieces without bearing on each other, and as equal in importance rather than differentially weighted based on the circumstances and goals.

At the *Competent* level, performers have mental models that they can apply to new situations. This stage is marked by the ability to envision and predict how a situation is likely to play out, which guides the

formulation, prioritization, and management of longer-term goals. Competent performers are very planful, where advanced beginners are more reactive. However, Competent individuals tend to adhere to the plan as the situation plays out, even when circumstances change. They have difficulty adapting their plan to address new situational demands.

As individuals progress to the *Proficient* stage, their performance shifts from being guided by the plan to being responsive to the situation. They see the situation as an inseparable whole rather than as independent attributes; they have the ability to recognize meaningful patterns of cues without breaking them down into their component parts for analysis. As such, they are able to intuitively assess what is happening and what is most critical for achieving success. They shift their assessment of the situation as it evolves and changes, and they can adjust their course of action accordingly. However, while their situation assessment is recognitional and intuitive, they still perform deliberate analysis when making decisions and devising or adjusting a course of action.

Expert performance is marked by a shift to recognitional decision making. Experts intuitively assess the situation and also intuitively recognize a suitable course of action that will accomplish their goals. They have a substantial base of experience from which to operate. Their mental models are broad, deep, and elaborate. They are able to make fine discriminations between perceptual cues (Klein & Hoffman, 1993), and can diagnose and assess situations that confuse or stump their less-experienced peers. Experts also have a wide range of routines and tactics for getting things done (Klein, 1998).

The five stages of the Dreyfus and Dreyfus model mapped easily onto the general descriptors of tactical thinking performance provided by Lussier for each of the eight TLAC themes. In the following sections we describe the process by which the T-BARS tool was developed, building on the TLAC research and the Dreyfus and Dreyfus model.

Participants

A total of 73 Army and Marine officers were interviewed to capture their thinking processes as they worked through a set of tactical problems. Five data collection efforts were conducted, the first at Fort Campbell, TN; the second at Fort Carson, CO; the third at Fort Sill, OK; the fourth at Fort Hood, TX; and the fifth at Fort Leavenworth with students from the School of Advanced Military Studies (SAMS). To ensure a range of performance levels to inform each of the five performance points on the BARS, participants represented a range of experience levels and specialty areas. We interviewed a Lieutenant, 24 Captains, 27 Majors, 22 Lieutenant Colonels, and 2 Colonels. These officers represented specialty areas including armor, aviation, engineering, field artillery, infantry, military police, and signals.

In addition to the 73 officers who participated in the standardized data collection, we also utilized archival data from 2 Lieutenants and 2 Major Generals to provide additional behavioral descriptors at the novice ('1') and expert ('5') points on the scales, respectively.

Vignettes

The data collection utilized a set of vignettes that placed the participant in the position of a company commander. The vignettes were selected from the set of TLAC scenarios in use at Fort Knox for training Army Captains (US Army Research Institute for the Behavioral and Social Sciences & US Armor Center 16th Cavalry Regiment, 2004). Each vignette consisted of a scenario set in Azerbaijan involving military operations in that area. A Road to War background description was read by participants to set the stage for the set of vignettes. For each vignette, an Operations Order or Fragmentary Order was provided to specify the mission. The vignettes themselves were web-based scenarios containing maps and graphics to indicate movements and locations (e.g., highlightings on the map; see Figure 1). The vignettes were pre-scripted and evolved over time, with narration accompanied by incoming situation reports and other communications from players within the mission (e.g., platoon leaders, local civilians, etc). In all cases, the interviewer controlled the progression of the vignettes.

Five vignettes addressing an array of mission types were employed. In most cases, a single data collection session utilized a subset of three vignettes. In some situation, only two vignettes were conducted due to time constraints. The five vignettes employed were:

1. *Establish a Safe Route*. The participant is required to clear a route through potentially hostile country into an urban area, accompanied by the U.S. ambassador's assistant.
2. *Enable Humanitarian Operations*. While escorting a humanitarian aid convoy to a refugee camp, the participant comes upon a flooded town in need of help.
3. *Man a Border Outpost*. The participant controls five border outposts. In the midst of a holiday

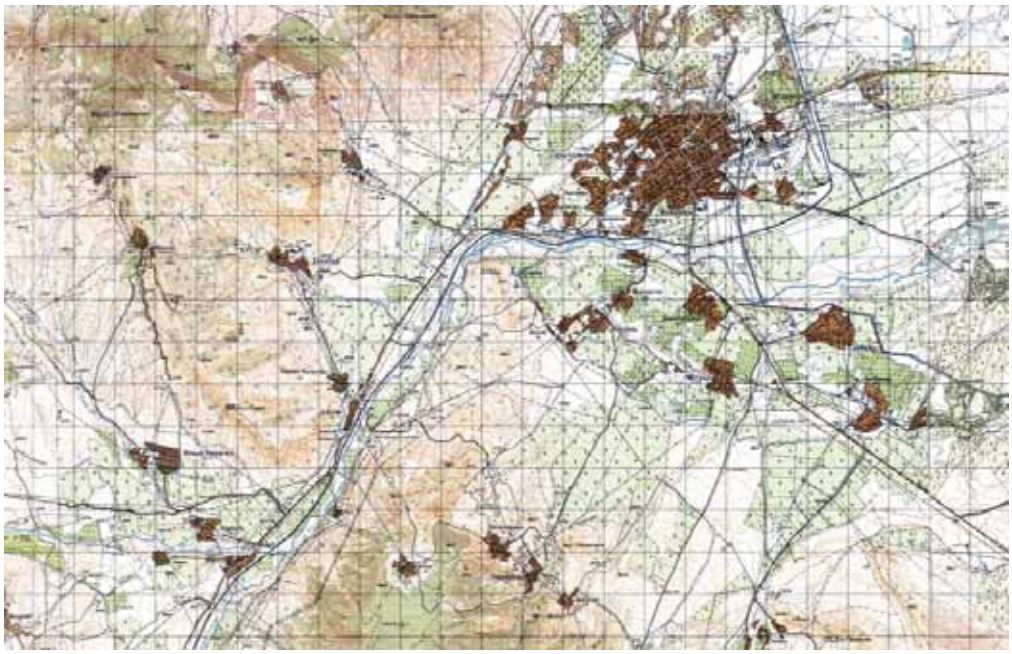

Figure 1. Map for Establish a Safe Route

celebration, an explosion occurs and one outpost no longer responds to communications.

4. *Conduct Presence Patrols.* The participant is forced to determine what to do when a subordinate detains suspected smugglers.
5. *Control a Civil Disturbance.* The participant is required to handle a situation in which two opposing crowds form at a bridge under his control.

For the first three data collection trips we used *Establish a Safe Route, Enable Humanitarian Operations,* and *Man a Border Outpost*. For the fourth data collection trip we used *Establish a Safe Route, Conduct Presence Patrols,* and *Control a Civil Disturbance.*

During the sessions, the interviewer stopped the vignette three times (including at the end) to ask the participant a series of questions from the interview protocol. These questions were designed to elicit the participant's tactical thinking. We chose to use vignettes to elicit participants' tactical thinking because we have found that lived incidents, whether real or simulated, stimulate more and richer information about the cognition concerning that incident (e.g., Hoffman, Crandall, & Shadbolt, 1998; Militello & Hutton, 1998; Pliske, McCloskey, & Klein, 2001).

Interview Methods

Two methods were utilized in the data collection: a think-aloud protocol and a Simulation Interview (SI). After the second data collection effort it was determined that think-aloud data was not as rich as SI data. For the third and fourth data collection efforts we employed only the SI. All data collection sessions were tape-recorded, and transcripts were prepared for use in analysis and scale development.

Think-Aloud

The think-aloud protocol was derived from a technique developed by Klein, Phillips, Battaglia, Wiggins, and Ross (2002). Within each vignette, the participant was asked to "Please think aloud about your responses to the following questions. What's important in this scenario? What information do you need? What will you do now?" If the participant fell silent at any point, the interviewer asked him or her to continue thinking aloud. We found that when given this prompt, participants tended to describe an action plan and a few important items of information, but were unable to expand on their cognitive processes or the reasons behind their plan.

Simulation Interview

The SI protocol was based on Militello and Hutton's (1998) Applied Cognitive Task Analysis technique. The SI developed by Militello and Hutton is intended to give the interviewer a better understanding of

participants' cognitive processes in the context of an incident. In our case, the TLAC vignettes provided the incident. The SI consists of a number of probes about different aspects of the incident. The probes we used were tailored for each of the three stopping points in the vignettes. They focused on what participants perceived as important, why they noticed those things, how they saw the situation developing, their priorities, and what information they sought and why.

Group vs. Individual Interviews

Group interviews were conducted in the first data collection effort. During these sessions, groups of two to six participants were exposed to a vignette in its entirety (i.e., without stopping points). Each participant presented a response to the vignette, which was then discussed by the rest of the group. The interviewers then facilitated a group discussion of the vignette. Think-aloud data was also recorded from each participant individually.

Not surprisingly, we found the group interviews to provide less information and insight into cognitive processes than the individual interviews. Furthermore, the responses tended to be amalgams of the group's thinking rather than genuine, organic responses from a single individual's thought process. After examining the data, we decided to use individual interviews only.

Analysis Methods

The analysis was conducted as an iterative process in two parts: first, developing and refining the scales using the behaviors and thought processes exhibited in the data transcripts, and second, using the newly refined scales to rate other data transcripts. We then used the experience of rating to continue to refine in the next iteration with data collected from a different set of participants.

We began with a previous version of the T-BARS that had been used to compare simulation and paper-and-pencil training techniques for the USMC (Phillips et al., 2003). Interview data from the first data collection effort was rated using this initial version of the T-BARS. The goal was to identify shortcomings in the initial scales in order to guide the remainder of the scale development process.

To rate this first data subset, transcripts of participants' responses to the vignettes were divided into individual segments. A segment was considered to be a portion of the transcript that contained a single, central thought being expressed. These were generally one to three sentences long. Five researchers rated the same data independently and the percent of segments given the same rating was calculated. Inter-rater reliability was low, as expected (mean r = .32, SD = .10). However, the rating exercise enabled us to identify which existing behavioral descriptors were useful and which were confusing. We also identified behaviors from the data set that were not currently represented in the T-BARS, which enabled the addition of behavioral descriptors to the scales.

Three succeeding iterations of scale refinement and ratings took place. Inter-rater reliability was calculated the same way as a check on the reliability of the current scale iteration. Each rating experience continued to guide the refinement of the T-BARSs based on why ratings had not agreed.

The goal of developing the T-BARSs was to assess the maturity of an individual's tactical thinking mental models. In previous research examining the use of the TLAC themes for tactical thinking skills training, it was concluded that a subset of the eight themes represent mental models, while another subset represent cognitive processes (Ross et al., 2003). It was hypothesized that the themes representing mental models – *Mission*, *Enemy*, *Assets*, and *Terrain* – must be built up to some basic level of comprehension before the themes representing cognitive processes – *Timing*, *Big Picture*, *Contingencies*, and *Visualization* – can be implemented (Ross et al., 2003; Ross, Battaglia, Phillips, Domeshek, & Lussier, 2003). Figure 2 illustrates this developmental process. The themes representing cognitive processes are exhibited by experienced, advanced tactical thinkers. They conduct these higher order mental operations in the context of the basic mental models represented by the first four themes. For example, an experienced tactician can estimate how long it will take to move a bridging asset from one point to another (*Timing* in the context of *Assets*) or predict what the enemy will attempt as the situation plays out (*Visualization* in the context of *Enemy*). Accordingly, the T-BARS tool was refined by encapsulating the behaviors associated with the cognitive process themes into the mental models themes, thereby resulting in four T-BARSs rather than eight.

Based on the issues identified with the T-BARSs from the rating iterations, and utilizing a subset of the interview data collected, the scales were continuously refined. Descriptive bullets were added to the T-BARSs within the appropriate theme (e.g., *Mission* or *Assets*) and at the appropriate level (e.g., 1–5). Because the interview data contained both thought processes described by the participants (e.g., "I'm thinking about

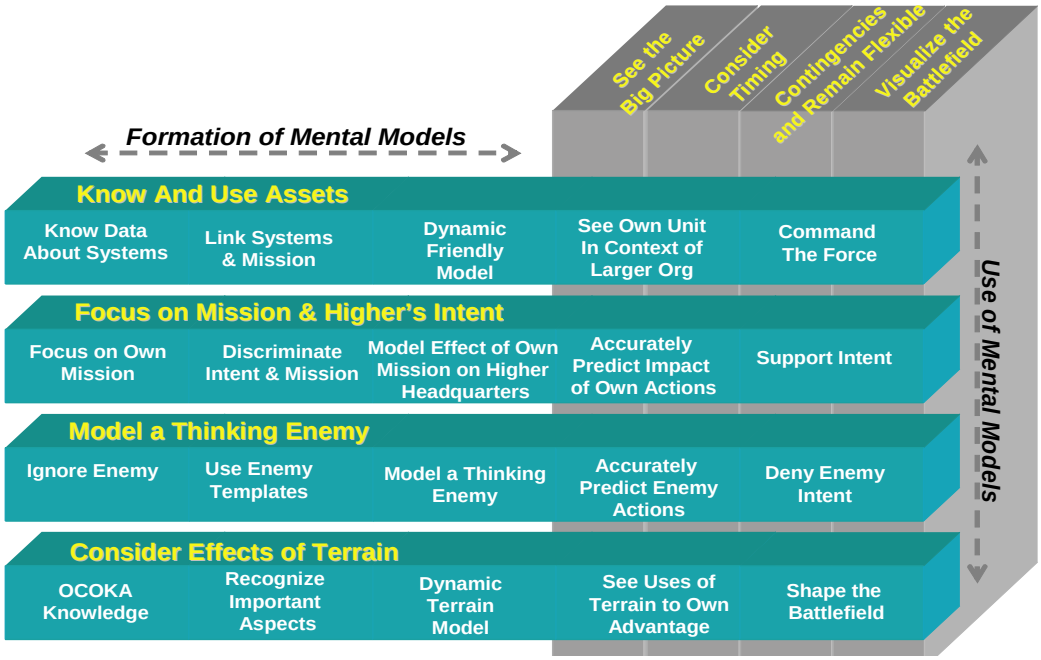

Figure 2. Eight Tactical Thinking Themes and their Hypothesized Developmental Sequence

how I'm going to provide security for my forces") and behaviors (e.g., "I'm going to lead with my tanks to clear the route for the rest of my forces"), we were able to draw linkages between the cognition of the participants and the behaviors that resulted.

While we found it straightforward to match behavioral descriptors to the appropriate TLAC themes, it was more difficult to determine the cognitive proficiency level represented by the behavior. To guide the placement of behavioral descriptors into the levels 1 through 5, we developed tactical thinking profiles based on the Dreyfus and Dreyfus (1986) five-stage model.

Tactical Thinking Profiles

A tactical thinking profile was developed for each of the five levels of performance on each of the four T-BARSs (*Mission*, *Enemy*, *Assets*, and *Terrain*). These profiles utilized the characteristics of the performance stages set forth by Dreyfus and Dreyfus (1986) as a foundation, and incorporated domain-specific instantiations of those characteristics. The profiles were generated not only to guide our refinement of the behavioral descriptors within the scales, but also to assist the intended users of the T-BARS tool with their assessments of individuals' performance. Table 1 presents the tactical thinking profiles for the *Assets* theme.

THE TACTICAL THINKING BARS ASSESSMENT TOOL

At the time of this writing, the T-BARS tool is in the final stages of refinement. Table 2 presents the most recent version of the scale for the *Mission* theme (due to space constraints, we have omitted the behavioral descriptors for levels 2 and 4). The final T-BARS product will include scales for each of the four tactical thinking mental models – *Assets*, *Mission*, *Enemy*, and *Terrain*. It will also include a user's guide with instructions for employing the T-BARS tool to rate individuals' performance and for drawing conclusions about individuals' overall tactical thinking profile and future training requirements.

A final iteration of T-BARS ratings will be conducted to ensure that acceptable inter-rater reliability is established prior to finalizing the scales.

CONCLUSIONS

Tactical thinking is a vital military skill. The essence of being an officer lies in making the right decisions and assessments, on the battlefield and off. Tactical thinking skills therefore should be, and are, key parts of military training. There is a critical need for an assessment technique that objectively measures tactical

Table 1. Tactical Thinking Profile for *Assets* Theme

Know and Use All Assets Available. Commanders must not lose sight of the synergistic effects of fighting their command as a combined arms team—this includes not only all assets under their command, but also those which higher headquarters might bring to bear to assist them.				
Level 1: Novice	**Level 2: Advanced Beginner**	**Level 3: Competent**	**Level 4: Proficient**	**Level 5: Expert**
Performance is abstract and rule-based, and focuses on variables in isolation. Individual know facts about standard capabilities of organic assets such as ranges of fire, number of vehicles per unit, and so forth. The foundational knowledge required to analyze how assets can be applied to the situation has not yet developed.	Performance reflects simple analytical processing using a limited experience base. Organic assets are matched to mission requirements. For example, a unit of tanks would be allocated to the area where heavy armor is needed for protection. Individual has difficulty prioritizing tasks, so asset utilization is driven by capabilities (what the asset can do) over situational demand (what is the most pressing mission task).	Performance reflects a mental model of asset utilization, but remains dependent on analysis and planning rather than recognition and intuition. Individual can prioritize mission tasks and predict how the situation could unfold, and an asset utilization plan is generated against that analysis. However, execution is driven by the plan over the situation, so individual has difficulty adjusting asset utilization to meet changing situational demands.	Performance reflects a recognitional or intuitive assessment of the situation, but analytical decision making where the individual deliberates about a course of action. Individual recognizes the availability of non-organic and non-military assets in addition his own organic assets. For example, civilians are recognized to be valuable sources of intelligence. Situational demands drive asset utilization, rather than the plan or the organic assets at the individual's disposal.	Performance reflects a recognitional ability to assess and decide. Individual can visualize specific outcomes of asset utilization and has the ability to avoid unwanted consequences. For example, he or she knows how to command and maneuver his or her forces to avoid an uprising by the locals. Individual leverages and coordinates organic, non-organic, and non-military assets to achieve mission objectives.

thinking skill, and the T-BARS tool is being developed to meet that requirement.

Assessing tactical thinking has a variety of applications. Evaluating students' current skill level is beneficial for determining their fitness for jobs of high responsibility and pressure. A tool for assessing tactical thinking will also be useful in evaluating the effectiveness of various training interventions. Finally, an assessment tool can be useful for identifying the type of training a student needs by determining his or her strengths and weaknesses.

ACKNOWLEDGMENTS

This work was funded by contract W74-V8H-C-00018 for the Army Research Institute. We wish to acknowledge the valuable guidance, support, and sponsorship provided by Jim Lussier of the Army Research Institute at Fort Knox. We also wish to express our gratitude to the many Army officers who kindly devoted their time and energy so that we could better understand their tactical thinking process.

Table 2. Excerpts from T-BARS for *Mission* Theme

Keep a Focus on the Mission and Higher's Intent. Commanders must never lose sight of the purpose and results they are directed to achieve – even when unusual and critical events may draw them in a different direction.		
Focus on Own Mission	**Models the Interaction between Mission and Intent**	**Actions are Guided by Intent**
1	**3**	**5**
• Adheres rigidly to the mission. • Articulates an understanding of the mission without any consideration of higher intent. • Neglects to keep HQ informed of plans and situation. • Takes passive stance when action is more appropriate. • Asks questions or seeks clarification about events in the scenario or situation. • Asks questions or seeks clarification on mission or situation.	• Articulates a consideration of whether mission will support the intent. • Analyzes intent statement in order to determine what has to be accomplished. • Recognizes a need to prioritize mission tasks or subtasks. • Mentally simulates and articulates what would be a favorable outcome for a particular task. • Projects ahead and articulates own tasks that will happen next. • Identifies something that needs to be accomplished as part of a mission task, but not how to do it. • Recognizes that there are second or third order effects of a particular action, but only understands them in general. • Forges ahead with a particular COA when it is apparent that the COA is no longer relevant. • Articulates general sequencing of mission tasks, but does not articulate *how* to accomplish each task. • Notes the importance of timing and/or sequencing for the particular mission.	• Sizes up the situation and generates a COA that supports intent, rather than being guided by the explicit mission. • Articulates how and/or why his plan or COA supports the commander's intent. • In execution, recognizes that the situation has changed and articulates adjustments to the COA or approach in order to accomplish mission or achieve intent. • In planning or execution, allocates assets based on a prediction about the enemy. • Exhibits ability to respond to a variety of battlefield states as a result of contingency planning, or mentally simulating contingencies. • Acts to create advantages for higher. • Presents opportunities to higher or adjacent units. • In execution, articulates second or third order effects of a particular action. • In execution, articulates actions necessary to ensure mission accomplishment when faced with a threat to mission accomplishment.

REFERENCES

Alberts, D. S., & Hayes, R. E. (2002). *Code of best practices for experimentation*, from http://ww.dodccrp.org/publications/pdf/Alberts_Experimentations.pdf

Benner, P. (1984). *From novice to expert: Excellence and power in clinical nursing practice*. Menlo Park: CA: Addison-Wesley Publishing Company Nursing Division.

Benner, P. (2004). Using the Dreyfus model of skill acquisition to describe and interpret skill acquisition and clinical judgment in nursing practice and education. Bulletin of Science, Technology & Society, 24(3), 189-199.

Deckert, J. C., Entin, E. B., Entin, E. E., MacMillan, J., & Serfaty, D. (1996). *Military decision-making expertise* (ARI Research Note 96-15). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Dreyfus, H. L., & Dreyfus, S. E. (1986). *Mind over machine: The power of human intuitive expertise in the era of the computer*. New York: The Free Press.

Hoffman, R. R., Crandall, B. W., & Shadbolt, N. R. (1998). Use of the critical decision method to elicit expert knowledge: A case study in cognitive task analysis methodology. *Human Factors, 40*(2), 254-276.

Houldsworth, B., O'Brien, J., Butler, J., & Edwards, J. (1997). Learning in the restructured workplace: A

case study. *Education and Training, 39*(6), 211-218.

Klein, G. (1998). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.

Klein, G., Phillips, J. K., Battaglia, D. A., Wiggins, S. L., & Ross, K. G. (2002). *Focus: A model of sensemaking* (Interim Report - Year 1 prepared under Contract 1435-01-01-CT-31161 [Dept. of the Interior] for the U.S. Army Research Institute for the Behavioral and Social Sciences, Alexandria, VA). Fairborn, OH: Klein Associates Inc.

Klein, G. A., & Hoffman, R. (1993). Seeing the invisible: Perceptual/cognitive aspects of expertise. In M. Rabinowitz (Ed.), *Cognitive science foundations of instruction* (pp. 203-226). Mahwah, NJ: Lawrence Erlbaum Associates.

Lussier, J. W. (1998). *Developing expertise in battlefield thinking: Models of battlefield thinking* (Unpublished briefing). Ft. Leavenworth, KS: U. S. Army Research Institute.

Lussier, J. W., Ross, K. G., & Mayes, R. (2000). *Coaching techniques for adaptive thinking.* Paper presented at the 2000 Interservice/Industry Training, Simulation, and Education Conference, Orlando, FL.

Lussier, J. W., Shadrick, S. B., & Prevou, M. I. (2003). *Think like a commander prototype: Instructor's guide to adaptive thinking* (ARI Research Product 2003-01). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

McElroy, E., Greiner, D., & de Chesnay, M. (1991). Application of the skill acquisition model to the teaching of psychotherapy. *Archives of Psychiatric Nursing, 5*(2), 113-117.

Militello, L. G., & Hutton, R. J. B. (1998). Applied Cognitive Task Analysis (ACTA): A practitioner's toolkit for understanding cognitive task demands. *Ergonomics, Special Issue: Task Analysis, 41*(11), 1618-1641.

Muchinsky, P. M. (2003). *Psychology applied to work.* Melmont, CA: Wadsworth/Thomson Learning.

Phillips, J. K., Shafer, J., Ross, K. G., Baxter, H., & Harris, D. S. (2003). *Framework for assessment of tactical decision-making simulations: Results and discussion for final report* (Final Report prepared under Subcontract CHI 00001-07 (Prime IDIQ N61339-99-D0012) for CHI Systems). Fairborn, OH: Klein Associates Inc.

Pliske, R. M., McCloskey, M. J., & Klein, G. (2001). Decision skills training: Facilitating learning from experience. In E. Salas & G. Klein (Eds.), *Linking expertise and naturalistic decision making* (pp. 37-53). Mahwah, NJ: Lawrence Erlbaum Associates.

Riggio, R. E. (2000). *Introduction to industrial/organizational psychology*. Upper Saddle River, NJ: Prentice Hall.

Ross, K. G., Battaglia, D. A., Hutton, R. J. B., & Crandall, B. (2003). *Development of an instructional model for tutoring tactical thinking* (Final Technical Report for Subcontract No. SHAI-COMM-01; Prime Contract No. DASW01-01-C-0039 submitted to Stottler Henke Associates Inc., San Mateo, CA). Fairborn, OH: Klein Associates.

Ross, K. G., & Lussier, J. W. (1999). *A training solution for adaptive battlefield performance.* Paper presented at the 1999 Interservice/Industry Training, Simulation, and Education Conference, Orlando, FL.

Ross, K. G., Lussier, J. W., & Klein, G. (2004). From recognition-primed decision making to decision skills training. In S. Haberstroh & T. Betsch (Eds.), *Routines of decision making*. Mahwah, NJ: Lawrence Erlbaum Associates.

Ross, K. G., Battaglia, D. A., Phillips, J. K., Domeshek, E. A., & Lussier, J. W. (2003). Mental models underlying tactical thinking skills. In *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference 2003*.

Spiro, R. J., Feltovich, P. J., Jacobson, M. J., & Coulson, R. L. (1992). Cognitive flexibility, constructivism, and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domains. In T. Duffy & D. Jonassen (Eds.), *Constructivism and the technology of instruction: A conversation* (pp. 121-128). Mahwah, NJ: Lawrence Erlbaum Associates.

United States Army Research Institute for the Behavioral and Social Sciences, & United States Armor Center 16th Cavalry Regiment. (2004). *Think Like A Commander – Mission to Azerbaijan: Instructor materials* (ARI Research Product 2004-6). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.