

Technical Report 1364

Implementing Measures of Individual and Collective Hypothesis Generation: A User's Guide

Drew L. Leins

Jim Leonard

Applied Research Associates

Christopher L. Vowels

U.S. Army Research Institute



May 2018

**United States Army Research Institute
for the Behavioral and Social Sciences**

Approved for public release; distribution is unlimited.

**U.S. Army Research Institute
for the Behavioral and Social Sciences**

**Department of the Army
Deputy Chief of Staff, G1**

Authorized and approved:

**MICHELLE SAMS, Ph.D.
Director**

Research accomplished in-house for the
Department of the Army

Technical Review by

Jayne Allen, U.S. Army Research Institute
Cary Stothart, U.S. Army Research Institute

NOTICES

DISTRIBUTION: This Technical Report has been submitted to the Defense Information Technical Center (DTIC). Address correspondence concerning ARI reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, Attn: DAPE-ARI-ZXM, 6000 6th Street, Building 1464/Mail Stop 5610, Fort Belvoir, VA 22060-5610.

FINAL DISPOSITION: Destroy this Technical Report when it is no longer needed. Do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this Technical Report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy) May 2018			2. REPORT TYPE Final			3. DATES COVERED (from... to) June 2016 to June 2017		
4. TITLE AND SUBTITLE Implementing Measures of Individual and Collective Hypothesis Generation: A User's Guide						5a. CONTRACT OR GRANT NUMBER W5J9CQ-11-D-0003		
						5b. PROGRAM ELEMENT NUMBER 633007		
6. AUTHOR(S) Drew L. Leins, Jim Leonard; Christopher L. Vowels						5c. PROJECT NUMBER A792		
						5d. TASK NUMBER 331		
						5e. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Applied Research Associates 1235 S. Clark Street Suite 1212 Arlington, VA 22202						8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U. S. Army Research Institute for the Behavioral and Social Sciences 6000 6th Street Fort Belvoir, VA 22060-5610						10. MONITOR ACRONYM ARI		
						11. MONITOR REPORT NUMBER Technical Report 1364		
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.								
13. SUPPLEMENTARY NOTES ARI Research and Subject Matter POC: Christopher L. Vowels								
14. ABSTRACT (Maximum 200 words): The goal of this research was to understand how Soldiers use heuristics when generating hypotheses to explain threat risk in operational environments and to develop a measure of cognitive processes leaders/trainers could use to evaluate Soldiers' performance during training for those environments. To translate the findings for training and evaluation purposes, we conducted focus groups with squad level Soldiers and leaders. Our purpose was to elicit feedback that would identify gaps in training and possible applications for measures of decision-making in training to detect/assess threats. We conducted 12 focus groups, each with 3-5 Infantry squad members and leaders and/or platoon sergeants. Several themes emerged: current approaches, gaps and needs, implementation strategies, hurdles and pitfalls, and potential benefits. A frequent challenge was training evaluations can be too subjective and differ, sometimes significantly, across trainers with different experiences and biases. Measures must be developed that reduce subjectivity and also increase reliability in scoring relevant behaviors. Based on Soldier guidance, we developed an implementation strategy for small teams, including an option for how leaders could integrate an assessment and adapt it to different tasks, an assessment leaders could use to score Soldier performance, and an exercise trainers could use to incorporate the assessment.								
15. SUBJECT TERMS hypothesis generation assessment performance measurement collective task squad training Infantry								
SECURITY CLASSIFICATION OF						19. LIMITATION OF ABSTRACT Unlimited	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON Dr. Brian T. Crabb Chief Fort Hood Research Unit 254-288-9403
16. REPORT Unclassified		17. ABSTRACT Unclassified		18. THIS PAGE Unclassified				

Standard Form 298

Technical Report 1364

Implementing Measures of Individual and Collective Hypothesis Generation: A User's Guide

Drew L. Leins

Jim Leonard

Applied Research Associates

Christopher L. Vowels

U.S. Army Research Institute

Fort Hood Research Unit

Brian Crabb, Chief

May 2018

Approved for public release; distribution is unlimited.

ACKNOWLEDGMENTS

We would like to thank the leaders and Soldiers across multiple installations for taking considerable time to provide input and guidance for this research effort. Their experience and wisdom and service are immeasurable.

IMPLEMENTING MEASURES OF INDIVIDUAL AND COLLECTIVE HYPOTHESIS GENERATION: A USER'S GUIDE

EXECUTIVE SUMMARY

Research Requirement:

The goal of this research was to better understand the heuristics Soldiers use when generating hypotheses to explain threat risk in operational environments, and to develop a measure of cognitive processes leaders/trainers could use to evaluate Soldiers' threat assessment performance in training. Across multiple studies, we examined how individual Soldiers, and groups of Soldiers working collectively, use mental shortcuts (heuristics) to assess threats in familiar tasks (e.g., detecting threats in operational environments) and unfamiliar tasks (e.g., diagnosing medical conditions) under different amounts of time pressure.

When working to detect threats individually, Soldiers appeared to use mental shortcuts guided by recognition. When they recognized the information in a scenario to be relevant to potential threats in the scenario, they assessed the related threat risk accordingly; however, when they were likely unable to recognize information as relevant, their assessments were suboptimal. Stated simply, they quickly recognized and assessed familiar threats, but failed to quickly recognize and assess unfamiliar threats.

Time pressure also influenced Soldiers' assessments. When under no time pressure, Soldiers maximized the amount of information they evaluated before making an assessment, leading to better informed assessments. However, when under time pressure, Soldiers evaluated less information, leading to quicker, but less informed assessments. This occurred whether Soldiers were assessing threats in familiar (i.e., operational) decision environments or unfamiliar (i.e., medical) decision environments.

When Soldiers worked to assess threats collectively, we observed different effects. First, groups of Soldiers tended to delay their assessments longer than did Soldiers working individually. But, groups of Soldiers continued to make quicker, better informed assessments when completing familiar scenarios than when completing unfamiliar scenarios. Furthermore, groups of Soldiers with relevant deployment experience were more likely than groups of less experienced Soldiers to make efficient, well informed threat assessments in scenarios depicting operational environments; however, both groups performed similarly in unfamiliar scenarios. Notably, time pressure did not force groups, regardless of experience, to make assessments sooner than necessary.

To determine how best to transition the findings of this research to training and evaluation, we conducted focus groups in which we presented to trainers/leaders our findings and example applications. Our goal was to elicit information that would help identify gaps in training and possible applications for measures of decision-making in training to detect/assess threats.

Procedure:

We conducted 12 focus groups, each with 3-5 squad leaders and/or platoon sergeants, recruited from an infantry brigade. We interviewed these leaders because they are responsible for training at the small unit level. All focus group sessions were documented via hand-written notes. Each session lasted approximately 90 minutes.

Leaders shared their thoughts and opinions about performance elements relevant to threat detection/assessment training. For example, they discussed the importance of Soldiers being able to differentiate relevant from irrelevant cues in their environment and to communicate to members of their team when they perceive something relevant. They discussed whether behaviors critical to threat assessment were or could be observed and measured, and whether they could be measured on a continuous scale different from, or supplemental to, a binary Go vs. No-Go measure. To conclude the focus groups, leaders discussed methods for implementing performance measures and considered an example training task in which performance measures could be implemented.

Findings:

Leader feedback was coded and analyzed for themes. Several themes emerged and clustered under the following categories: current approaches, gaps and needs, implementation strategies, hurdles and pitfalls, and potential benefits. On current approaches to training and evaluating decision-making behaviors such as threat assessment, the most common sentiment shared by leaders was that they already evaluate these behaviors, but their methods are informal, unstandardized and not recorded. On current training gaps and needs, the most common themes were that leaders lack opportunities for developing training, they lack dedicated opportunities for observing and evaluating their Soldiers' decision-making, and that they need a model for training and explicitly evaluating decision-making specific to threat assessment.

When discussing implementation strategies, many leaders indicated that training and evaluation of decision-making in threat assessment should be implemented at the squad or team level. Leaders also suggested implementing this type of training as opportunity training, to be conducted as a supplement to more formal training when unexpected training time becomes available. Alternatively, many leaders suggested that evaluations of critical cognitive processes could be incorporated into existing training, for example, in situation training exercises field training exercises, and live fire exercises. Finally, several leaders suggested that these evaluations would be best incorporated informally in after action reviews (AARs) or as part of monthly counseling.

On preparing new leaders to incorporate measures of decision-making performance, many leaders recommended introducing implementation guidance at the Basic Leader Course. Some leaders also suggested that a separate train-the-trainer would be useful for introducing the critical concepts, behaviors, and measures to new leaders, particularly if they were to implement the new measures at AARs or monthly counseling. Reflecting a reported need for threat assessment training, multiple leaders advised that we develop stand-alone threat assessment training incorporating these measures.

Leaders identified several challenges to implementation. A frequently cited challenge was that these evaluations can be too subjective and differ across leaders with different experiences and biases. Thus, implementation must address how to reduce the use of measures that allow too much subjectivity, but also increase reliability in scoring relevant behaviors. Another challenge to implementing the tool is that to evaluate threat assessment performance, leaders must overcome constraints on training opportunities, time, and resources. Thus, an implementation strategy that allows leaders to evaluate critical behaviors during existing training exercises without requiring additional substantive time or resources will help leaders seamlessly integrate such measures. Other implementation challenges include the need for training to consider the cultural context of the relevant operational environment, the different learning styles across Soldiers, and that training the critical cognitive components of threat assessment may be somewhat foreign to many leaders/trainers.

Despite the challenges to implementing a tool for evaluating the behaviors associated with threat assessment performance, leaders recognized several benefits of implementation. First, an evaluation tool with clear rules for scoring relevant behaviors (and linking them with underlying cognitive processes) can help clarify standards for new leaders. Second, a 5-point scale for measuring performance would provide more thorough information than the binary Go vs. No-Go metric currently used. Leaders suggested that a more detailed understanding of performance will help promote improvement because it offers Soldiers more measurable and more specific feedback on performance, and it could promote healthy competition among Soldiers and squads.

Utilization and Dissemination of Findings:

Based on focus-group feedback, we developed a sample implementation strategy for leaders to consider, including an option for where leaders could implement the evaluation tool, some guidance for how leaders can adapt the tool to different tasks, an example of the form leaders could use to score Soldier performance, and a sample practical exercise that leaders can use to pilot test the evaluation tool.

Presently, the most practical option for implementing this training and evaluation tool may be as opportunity (e.g., hip-pocket or sergeant's time) training. This implementation will not require development of a new formal training program or the modification of an existing training program. Rather, the leader can adapt the example materials to run a brief training exercise to evaluate a team's threat assessment performance. For example, leaders can consider whether existing training lanes are available, and how to access them, or if they have to develop their own training lanes. They can gather materials or identify environments relevant to the types of threats their Soldiers must be prepared to identify and then consider the cognitive processes involved in identifying the threats, assessing the related risk, and communicating assessments to teammates or leaders.

Leaders may find it useful to run their Soldiers through threat assessment training exercises multiple times and under different conditions to observe how different circumstances affect their teams' threat assessment performance. In addition, leaders might ask different squad members to lead the exercise to observe Soldiers' performance in different positions within the squad. In addition to scoring individual Soldiers, leaders may also find it useful to score their

teams' performance collectively to gain better insight regarding how their teams perform over time. This can also provide leaders insight into how individual Soldiers' performance may influence the overall team's performance.

IMPLEMENTING MEASURES OF INDIVIDUAL AND COLLECTIVE HYPOTHESIS GENERATION: A USER’S GUIDE

CONTENTS

	Page
Method	3
Results	4
Current Approaches.....	4
Current Gaps and Needs.....	4
Implementation Strategies	5
Challenges to Implementation	6
Potential Benefits	7
Sample Implementation Strategy.....	8
Implementation	8
Adaptation Guidance.....	8
Evaluation Form (Squad Threat Assessment)	9
Practical Exercise	10
References	13

APPENDICES

Appendix A. Focus Group Slide Deck.....	A-1
Appendix B. Description of Steps for the Squad Threat Assessment (STA) and STA versions	B-1
Appendix C. Trainer’s Slide Deck.....	C-1

IMPLEMENTING MEASURES OF INDIVIDUAL AND COLLECTIVE HYPOTHESIS GENERATION: A USER'S GUIDE

Across multiple projects, a set of studies has been conducted exploring how Soldiers identify and explain threat risks in different environments (Leins, Leonard, Masters, & Vowels, submitted; Leins, Leonard, Zimmerman, Minchew, & Vowels, In Preparation; Leins, Zimmerman, Mueller, Singer, & Vowels, 2016; Zimmerman, Leins, Marcon, Mueller, Singer, & Vowels, 2016; Zimmerman, Leins, Marcon, Perelman, Singer, Mueller, & Vowels, 2013; Zimmerman, Mueller, Daniels, Marcon, & Vowels, 2012; Zimmerman, Mueller, Grover, & Vowels, 2014). Most recently, across three studies, we examined how individual Soldiers, and groups of Soldiers working collectively use mental shortcuts (heuristics) to assess threats in familiar tasks (e.g., detecting threats in operational environments) and unfamiliar tasks (e.g., diagnosing medical conditions) under different amounts of time pressure. These studies used the same general design: Soldiers viewed an image depicting a scenario, assessed the threat risk inherent in the scenario, watched as the scenario changed over time, and then signaled when their thinking changed. When Soldiers saw something in the image that changed the way they thought about the scenario's threat risk, they stopped the scenario and reported their new assessment. We determined their assessment strategies by examining when they stopped a given scenario and whether their new assessments were associated with incoming information that was more informative or less informative. The goal of these studies was to better understand the heuristics Soldiers use when generating hypotheses to explain threat risk in operational environments. Specifically, we wanted to examine whether Soldiers use simple rules for assessing environmental cues and determining threat risk, and whether their use of these rules changes across decision environments.

When working to detect threats individually, Soldiers appeared to use mental shortcuts guided by recognition (see recent reports by Leins et al, for details on methods and results). When they recognized the information in a scenario to be relevant to potential threats in the scenario, they assessed the related threat risk accordingly; however, their assessments also reflected instances when they were unable to recognize information as relevant. Specifically, their assessments tended to coincide with the appearance of the most valuable information when detecting threats in scenarios representing operational environments, but not when detecting threats in scenarios representing medical diagnostic environments. Their familiarity with relevant threat cues (e.g., disturbed earth indicative of improvised explosive device (IED) emplacement) in operational environments allowed them to

HYPOTHESIS GENERATION

Soldiers use environmental cues to generate hypotheses to make sense of their surroundings.

We all use heuristics – mental shortcuts – to help generate hypotheses. Heuristics reduce the amount of information we need to evaluate before making a decision.

When Soldiers use heuristics to generate hypotheses about threat risk, they must ensure that the few pieces of information they rely on are good.

Variables such as time pressure and uncertainty can lead to poor decisions if they cause Soldiers to adopt the wrong decision strategy or rely on the wrong information.

Training Soldiers to adopt the right strategies and attend to the right information will help ensure safety and mission success.

differentiate relevant from irrelevant information and make efficient assessments, sometimes based only on limited but highly relevant information. By contrast, when trying to assess the risk associated with symptoms of medical conditions, their lack of familiarity with those symptoms precluded them from differentiating relevant from irrelevant information, and caused them to evaluate as much information as possible before assessing the inherent threat risk. Thus, even when highly relevant symptoms appeared early in a scenario, Soldiers delayed their assessments until they could evaluate more information. Stated simply, they quickly recognized and assessed familiar threats, but failed to quickly recognize and assess unfamiliar threats.

In addition to familiarity, time pressure influenced the timing of assessments. When trying to assess threats in a series of scenarios, time pressure seemed to force Soldiers to change their assessment strategies. When under no time pressure, Soldiers maximized the amount of information they evaluated before making an assessment, leading to better informed assessments. By contrast, when under time pressure, Soldiers evaluated less information, leading to quicker, but less informed assessments. This happened whether Soldiers were assessing threats in familiar (i.e., operational) decision environments or unfamiliar (i.e., medical) decision environments.

When Soldiers worked to assess threats collectively, we observed slightly different effects. First, groups of Soldiers (up to four personnel) tended to delay their assessments longer than did Soldiers working individually. But, groups of Soldiers continued to make quicker, better informed assessments when completing familiar scenarios than when completing unfamiliar scenarios. Thus, recognition still supported more efficient decision-making. Notably, time pressure did not force groups into making less optimal decisions. Rather, groups delayed their assessments regardless of how much time they had to complete a set of assessments.

Given that familiarity with the decision environment promotes efficient, well informed assessments, it was not surprising that groups of Soldiers with relevant deployment experience were more likely than groups of less experienced Soldiers to make efficient, well informed threat assessments in scenarios depicting operational environments. However, both groups performed similarly on unfamiliar scenarios.

These effects suggest that when under time pressure, Soldiers are susceptible to making assessments based on suboptimal information. Encouraging, however, is that working in groups may mitigate the effects of time pressure. Soldiers working in groups appeared less likely than individuals to base their assessments on suboptimal information. To be sure, they were generally less efficient than individual Soldiers, but the influence of group dynamics and group decision processes may have protected Soldiers against lowering their standards of informativeness required before making a decision. Moreover, these effects suggest that Army leaders/trainers can observe, measure, and train the methods Soldiers use for assessing threat risk, to help accelerate skill acquisition and Soldier readiness. To determine how best to transition the findings of this research to training and evaluation, we conducted focus groups in which we presented to trainers/leaders our findings and example applications. Our goal was to elicit information that would allow us to identify gaps in training and possible applications for measures of decision-making in training that involves threat detection and assessment. Soldiers provided feedback on current training and evaluation practices, gaps in training, current training needs, how best to implement the research findings, hurdles, and pitfalls inherent in

implementing these findings, and the potential benefits of implementing these findings as training and evaluation measures.

Method

We conducted 12 focus groups, each with 3-5 squad leaders and/or platoon sergeants, recruited from an infantry brigade. We interviewed these leaders because they are primarily responsible for training at the small group level and, being Combat Arms, training to identify possible operational threats is paramount. All focus group sessions were documented via handwritten notes. Each session lasted approximately 90 minutes. To provide context, we described to leaders the studies we conducted on threat assessment leading up to the development of the Squad Threat Assessment (see Appendix A).

Leaders shared their thoughts and opinions about performance elements relevant to threat detection/assessment training. For the purposes of the focus groups, we construed threat assessment as the activity involving identifying features of an operational environment that raise or lower the risk of operating in that environment. This includes core cognitive processes such as perceiving cues, interpreting the value or threat-relevance of cues, prioritizing cues, and generating hypotheses to make sense of the environment and guide decision-making. In this context, leaders discussed the importance of Soldiers being able to differentiate relevant from irrelevant cues in their environment and communicate to members of their team when they perceive something relevant. They discussed whether behaviors critical to threat assessment were or could be observed and measured. For example, they discussed measures of time to report a threat assessment and the quality of a threat assessment.

They shared thoughts on whether performance could be measured using a continuous scale, such as the scale depicted in Figure 1. To close the focus groups, leaders then discussed alternative methods for implementing performance measures and considered an example training task in which performance measures could be implemented.

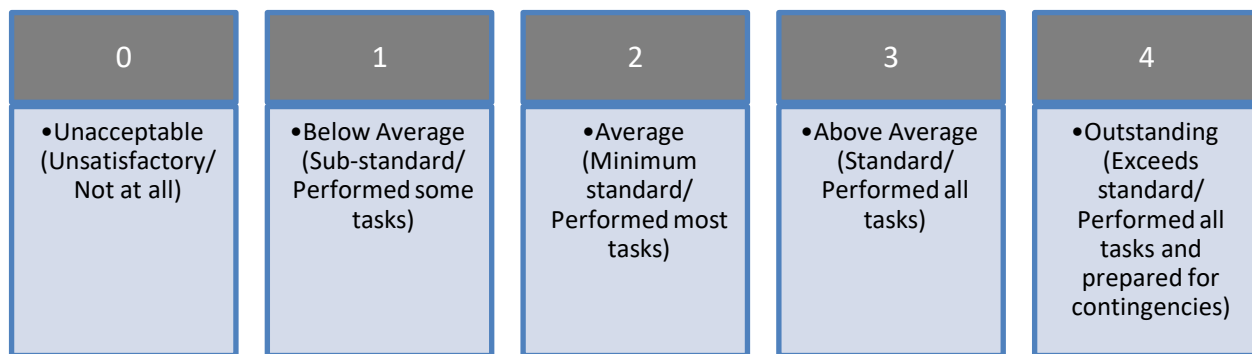


Figure 1. Example of a 5-point scale for evaluating decision-making behaviors.

Following the focus groups, we coded and analyzed the leader feedback for themes. Several themes emerged and clustered under the following categories: current approaches, gaps and needs, implementation strategies, hurdles and pitfalls, and potential benefits. We identify and discuss these themes in the next section.

Results

Current Approaches

When asked to discuss their current approaches to training and evaluating threat assessment, many leaders reported that they already evaluate some of the critical processes underlying threat detection and assessment (e.g., differentiating relevant from irrelevant threats and communicating perceptions). For instance, to evaluate their more junior Soldiers' decision-making when the team structure changes, and existing leadership is not there to guide decision-making, leaders may ask their Soldiers to assume leadership roles in training. Leaders also reported asking their Soldiers to demonstrate their knowledge and understanding of tasks by asking them to teach back lessons those Soldiers had just learned. Thus, leaders reported evaluating decision-making performance, but they also reported doing so informally. For example, in an after action review (AAR), they may ask their Soldiers to justify decisions made during a training exercise. Those leaders who reported conducting such informal evaluations indicated that they did not necessarily measure behaviors or record the evaluations. Rather, they kept only mental notes of Soldiers who seemed to make good versus poor decisions. When they documented relevant performance evaluations, they reported doing so in notebooks maintained for their own records, not as formal evaluations maintained in Soldiers' official records. Leaders reported sometimes using these unofficial evaluations to help guide feedback and remediation plans discussed during monthly counseling.

When discussing one of the critical findings from our research – the influence of time pressure on decision-making – leaders indicated that they incorporate “time hacks” into some of their training. That is, they challenge their Soldiers to complete tasks under time constraints and they measure the time it takes to complete a task or steps within a task. However, they applied these time hacks to existing Training and Evaluation Outlines (T&EO). These T&EOs typically do not contain tasks or subtasks that represent cognitive activities, such as identifying and prioritizing threat cues or developing hypotheses to determine threat level. Thus, although some leaders measured the time taken to complete tasks and subtasks identified in an approved T&EO format, they did not necessarily measure the time taken to complete supporting cognitive tasks.

Regarding current approaches to training and evaluating decision-making behaviors such as threat assessment, the most common sentiment shared by leaders was that they already evaluate these behaviors, but their methods are informal, unstandardized, and not recorded.

Current Gaps and Needs¹

Some leaders suggested that the informal and unstandardized nature of evaluating decision-making, particularly in threat assessment, resulted because they do not have a model for specifically training threat assessment. Critically, however, they claim to need such a model. Moreover, many leaders reported having no control over formal training design. Thus, they do not yet have a model for training threat assessment and they see limited opportunity to develop a

¹ Although many leaders reported limited access and time to train threat assessment, it is important to note that training programs exist that incorporate some of the concepts identified here. For example, the Joint Improvised-Threat Defeat Organization (JIDO) offers resources to train warfighters to counter improvised threats (www.jieddo.mil/index.htm); and, the U.S. Army Maneuver Center of Excellence offers courses in Advanced Situational Awareness (<http://www.benning.army.mil/Armor/316thCav/ASA/index.html>).

model for training threat assessment. Consequently, they either do not train and evaluate this important component that supports several operational tasks (e.g., dismounted patrol, clearing operations), or they train and evaluate it informally. If they train and evaluate it informally, they encounter other challenges. Many leaders reported having too few opportunities to actually observe and evaluate Soldier performance. They have limited opportunities because they must take part in the very same collective tasks that would otherwise offer opportunities to observe and evaluate their Soldiers. Moreover, generating opportunities to observe and evaluate is difficult because leaders often have only limited time to complete their training. They do not often have the luxury of running an exercise enough times to allow them to participate (lead by example) and then observe and evaluate. Because they cannot often just observe, they typically evaluate and provide feedback only after completing the exercise (e.g., during an AAR). This becomes problematic because their evaluations are limited to what they can glean from their point of view operating in the training exercise, gross measures of task performance (e.g., time to complete task, success of meeting final objective), and from what their Soldiers can recall about their own and their teammates' performance during the exercise.

Leaders reported that they want their Soldiers to be good decision makers, but that they don't train decision-making per se. They need training that emphasizes decision-making processes that are critical to completing identified tasks. They reported needing training that offers leaders guidance on what to look for when evaluating decision-making, and how to convey performance requirements to their Soldiers. In addition, leaders reported needing a method for evaluating the quality of decision-making within a training exercise, and a more specific metric that supplements the current Go vs. No-Go metric they currently use.

Regarding current training gaps and needs, the most common sentiments shared by leaders were that they lack opportunities for developing training, they lack opportunities for observing and evaluating their Soldiers' decision-making, and that they need a model specifically for training and evaluating decision-making in threat assessment.

Implementation Strategies

Many leaders suggested that training and evaluation of decision-making in threat assessment should be implemented at the squad or team level. They suggested that rather than developing formal training exercises at the battalion or brigade level, this type of training and evaluation should remain flexible and adaptable to allow platoon sergeants to determine how best to conceptualize and convey the requirements for effective decision-making in a given task. Multiple leaders suggested that allowing platoon sergeants to develop the evaluation criteria would lead to more standardized and reliable evaluations made by squad and team leaders.

Leaders also suggested implementing this type of training as "opportunity" (hip-pocket/sergeant's time) training, to be conducted as a supplement to more formal training when unexpected training time becomes available. Such opportunity training can mitigate some the limitations imposed by the time and resource constraints of formal training exercises. During "opportunity" training, leaders can observe and evaluate their Soldiers' decision-making and offer immediate feedback.

Alternatively, many leaders suggested that evaluations of critical cognitive processes could be incorporated into existing training exercises. Specifically, these measures could be

adapted for evaluations in situational training, field training, and live fire exercises. However, they cautioned that such incorporation should not require a burdensome number of items to score, as it would be impractical to increase the number of subtasks to be scored for many training exercises. In fact, several leaders suggested that these evaluations would be better incorporated informally in AARs. Similarly, multiple leaders suggested these measures would be appropriate to include in monthly counseling sessions. Hence, leaders could use these measures to offer individual feedback as needed.

Leaders in our focus groups also discussed how best to prepare new leaders to incorporate measures of decision-making performance. Several leaders suggested introducing implementation guidance at the Basic Leader Course. Doing so might allow seamless integration of these measures with existing methods of performance evaluation. Some leaders also suggested that a separate train-the-trainer support package would be useful for introducing the critical concepts, behaviors, and measures to new leaders, particularly if they were to implement the new measures at AARs or monthly counseling. They indicated that new leaders might be more flexible and open to adopting these measures than existing leaders who may already have established methods for evaluating (decision-making) performance.

Finally, in line with the reported need for threat assessment training, multiple leaders suggested that we develop stand-alone threat assessment training incorporating these measures. They indicated it would be useful to develop training lanes that focused primarily on threat detection and assessment, lanes that would allow leaders to focus on evaluating the critical behaviors that correspond with threat detection and assessment skills.² Such skills support many of the primary tasks Combat Arms Soldiers, particularly infantry, must accomplish to be successful in their missions.

Regardless of the where and when this evaluation tool is implemented, leaders reported that it must include clear and consistent rules for scoring and guidance on how to apply the rules. A recurring theme in focus groups was that the evaluation tool should contain a limited number of items to score, to reduce the burden on leaders when training and evaluating, as well as to increase the likelihood that leaders will indeed use the tool. Leaders indicated that the flexibility to evaluate Soldiers individually and collectively as teams or squads would enhance the tool's utility. They also identified some form factors that would make the tool more appealing and enhance the likelihood they would adopt it. For example, many leaders suggested that the evaluation tool could include a scoring matrix that they could maintain in their notebook or that would fit in their hip pocket. The scoring matrix could use the 5-point scale presented in Figure 1, again, with concise but clearly defined guidelines for assigning scores.

Challenges to Implementation

When discussing how best to implement a tool for evaluating the cognitive components of threat assessment, leaders also identified some inherent practical challenges. One of the most frequently cited challenges was that these evaluations can be subjective and differ across leaders. Thus, implementation must address how to reduce subjectivity and increase reliability in scoring

² Note that many installations offer resources for training 'threat detection,' including space for setting up training lanes, or have existing lanes, and groups that will advise on how to set them up (e.g., Counter IED Integration Cells [CI2C]). However, these resources do not necessarily include guidance or methods on assessing hypothesis generation or other cognitive components underlying threat detection.

relevant behaviors. As noted above, one route to increasing reliability is to ensure that the scoring rules are clear and that the behavioral criteria for assigning a score are observable and measurable.

Another challenge to implementing the tool is that these evaluations can be difficult or impractical to make: To evaluate threat assessment performance, leaders must overcome constraints on training opportunities, time, and resources. Thus, an implementation strategy that allows leaders to evaluate critical behaviors during existing training exercises without requiring additional substantive time or resources will help leaders to integrate such measures.

Other implementation challenges that were mentioned include the need for training to consider the cultural context of the relevant operational environment, indicating any tool that evaluated threat assessment performance should be adaptable to training to identify the relevant threats. In addition, Soldiers can exhibit different learning styles. Whereas one Soldier may learn best in formal training, another may learn best in less traditional modes, again reinforcing the notion that the evaluation tool should be adaptable to different learning environments. Finally, some leaders echoed the importance of evaluating decision-making in threat assessment, but expressed concern with their proficiency in training the critical cognitive components of threat assessment. This concern highlights the challenge of developing a training and evaluation tool that addresses a critical need. However, it requires a foundation of knowledge that is not yet formally trained and would consume valuable time and resources that are already limited.

Potential Benefits

Despite the challenges to implementing a tool for evaluating the behaviors associated with threat assessment performance, leaders recognized several benefits of implementation. First, an evaluation tool with clear rules for scoring relevant behaviors (and linking them with underlying cognitive processes) can help clarify standards for new leaders. Currently, new leaders may rely on lessons learned from their previous leaders to convey to their Soldiers how to assess threats in operational environments. Although piecing together important features of various lessons can generate a good, comprehensive training plan, it can also generate an incomplete and substandard plan if the parent lessons are also incomplete and substandard. A standardized, empirically supported training and evaluation tool can offer new leaders a more stable starting point from which to train threat assessment.

Regarding the potential form of the evaluation tool, leaders indicated that a 5-point scale for measuring performance would be useful. First, it provides more detailed information than the binary Go vs. No-Go metric currently used for evaluating performance on most tasks. Leaders suggested that, compared to the Go/No-Go metric, a 5-point scale would give them a better understanding of their Soldiers' performance, help them give more specific feedback to Soldiers, and give Soldiers a better understanding of their own level of performance, identifying specific behaviors they need to work on. Leaders suggested that this more nuanced understanding of performance will help promote improvement not only because it offers Soldiers more measurable, detailed feedback on performance, but also because it could promote healthy competition among Soldiers. Several leaders reported that when Soldiers' task performance is quantified and ranked, Soldiers often strive to outperform their teammates, especially in Combat Arms units. Thus, whereas a determination of Go vs. No-Go (or trained vs. untrained) will not necessarily indicate to Soldiers how they rank among their peers, a continuous, quantifiable

measure of performance will (such as the Army Physical Fitness Test (APFT) and weapons qualification scores).

Sample Implementation Strategy

Based on the feedback offered by leaders in our focus groups, we developed a sample implementation strategy for leaders to consider. This sample implementation strategy includes an option for where leaders could implement the evaluation tool, some guidance for how leaders can adapt the tool to satisfy evaluation needs in different tasks, an example of the form leaders would use to score Soldier performance, and a sample practical exercise that leaders can use to pilot test the evaluation tool.

Implementation

Presently, the most practical option for implementing this training and evaluation tool may be as opportunity (e.g., hip-pocket or sergeant's time) training. This implementation will not require development of a new formal training program or the modification of an existing training program. Rather, the leader can adapt the materials included here, as well as commonly available materials, to run a brief training exercise to evaluate a team's threat assessment performance. This training exercise can be simple and brief, requiring few materials and little time, or it can be complex and lengthy, requiring a rich environment and enough time for Soldiers to process all of its elements. Moreover, it can be adapted to the unique features of a known operational environment (e.g., specific threats associated with terrain or culture).

Adaptation Guidance

To ensure that leaders and their Soldiers maximize the utility of this training and evaluation tool, leaders should adapt the tool to best fit the constraints of their training environment as well as the needs for ensuring operational readiness. For example, leaders should consider whether existing training lanes are available, and how to access them, or if they have to develop their own training lanes. Leaders should gather materials or identify environments relevant to the types of threats their Soldiers must be prepared to identify. For example, if training to identify threats in rural environments, leaders should consider the specific type of threat (perhaps buried IEDs) most likely to occur and the cues that indicate the presence of the threat. After identifying the environment, relevant threats, and cues to those threats, leaders should consider the cognitive processes involved in identifying the threats, assessing the related risk, and communicating assessments to teammates or leaders. Some of these processes will transfer across environments; for example, differentiating relevant from irrelevant cues requires the same attention and perception processes regardless of the type of threat indicated by those cues. Therefore, leaders can manipulate the training environment to include threat-relevant and threat-irrelevant features, to help determine how well their Soldiers identify and prioritize those features. As training environments differ with respect to complexity and the task to be completed, leaders can determine cut-off points for scoring the proportion (or number) of relevant threat cues or the time taken to identify those cues. Similarly, leaders can establish and convey appropriate standards for communicating threat risk across different training environments. Thus, leaders can convey to their Soldiers the different elements that contribute to a high quality threat assessment and evaluate them accordingly.

Leaders may find it useful to run their Soldiers through these training exercises multiple times, under different conditions. For example, leaders may want to run their teams through a lane and evaluate their threat assessment performance, then change the configuration of the lane (or introduce different threat-relevant cues) and constrain the time available to their teams to complete the exercise. This will allow leaders to observe how different conditions affect their teams' threat assessment performance. In addition, leaders might ask different squad members to lead the exercise, to observe Soldiers' (change in) performance in different positions within the squad.

In addition to scoring individual Soldiers, leaders may also find it useful to score their teams' performance collectively, to gain better insight regarding how their teams perform over time. This can also provide leaders insight into how individual Soldiers' performance may influence the overall team's performance and vice versa.

Evaluation Form (Squad Threat Assessment)

The evaluation form, the Squad Threat Assessment, in Appendix B offers leaders the ability to record several layers of information.³ First, leaders can evaluate individual Soldier performance by recording raw data and translating these data to ratings using the scale suggested on the form. For example, if they set up a training lane with three critical threat cues, they can record how many of those cues each Soldier identified and how long it took Soldiers to identify the cues. Leaders can then translate the number of threat cues, that are identified within a certain time, to a performance rating. Leaders can also record how long each Soldier took to generate an overall threat assessment for the training environment, record the core or gist of the assessment, and rate the quality of the assessment. Finally, leaders can identify whether each Soldier communicated the assessment to the team and the level of detail that was communicated. Leaders can also use this form to evaluate team performance. In the lower section of the form, leaders can record and rate alternative assessments offered by Soldiers during the training exercise. To provide context for rating and during feedback, leaders can also record the course of action (COA) chosen by the leader and the result of that COA.

Prior to using this form, leaders should identify the performance criteria that associate with each score. For example, when scoring for the number or proportion of relevant⁴ cues identified, leaders should associate those measures with performance scores (a raw score of 0/3 cues identified = a rating of "0," a raw score of 1/3 cues = a rating of "1," and so on). Similarly, time-related criteria can be used to score how long it takes Soldiers to identify a critical cue or a set number of cues. In some scenarios, distance to the threat may be more pertinent than the time taken to identify the threat. Therefore, for some training lanes, distance could be used instead of time to determine level of individual and collective performance. To rate communication behavior, leaders should consider what levels of communication are appropriate for a given

³ In an attempt to bridge academic notions with practical training considerations, one viewpoint may be to consider the Squad Threat Assessment as a collective (knowledge) updating process made observable and recordable. The literal or actual comprehension of the environment may be irrelevant or impossible/unachievable. Thus, what becomes critical is the process of trying to understand or comprehend the environment, which involves generating, ideally, increasingly more effective hypotheses/assessments of the current (and future) environment.

⁴ Trainers should weight 'relevant' cues appropriately for the context, taking into account the severity and likelihood of various threats across differing situations.

exercise, for example, whether Soldiers simply pointed out a potential threat (perhaps a rating of “2”) or they pointed out a threat and gave a reason why they believed it was a threat (perhaps a rating of “3”). When rating the quality of an assessment, leaders may need to consider whether the assessment was accurate and useful for mitigating threat risk (perhaps a rating of “4”). Did a Soldier identify a subtle or critical cue that no other Soldiers identified, leading to an optimally informed decision mitigating the situational risk (perhaps a rating of “4”)”? In contrast, did Soldiers miss cues and make poorly informed assessments (perhaps a rating of “0” or “1”)”? Did Soldiers misinterpret cues, leading to inaccurate assessments that put the squad at risk (perhaps a rating of “0”)”? These scoring criteria should be conveyed in general terms to Soldiers prior to executing the training exercise. For example, leaders can tell their Soldiers that they will be evaluated on how many threats they detect and how long it takes them to detect those threats.

During the training exercise, leaders can score their Soldiers and teams on identified behaviors and record justifications for each score on a comments line. Following completion of the training exercise, leaders may choose to share these scoring forms with their teams or give feedback verbally, using their recorded comments to support the scores they gave. Leaders can generate new forms for each exercise, or they may choose to laminate one sheet to be used during scoring and then wiped clean after entering the related information into a notebook for tracking performance over time or across conditions.

Practical Exercise

This exercise will allow a leader to pilot test the training and evaluation tool to assess individual and collective performance on threat assessment. The evaluation form will guide the leader in observing and measuring some critical cognitive processes that underlie threat detection and assessment (see Appendix B for different versions of the Squad Threat Assessment). The slide deck in Appendix C will provide additional guidance for trainers and instructors to design training and evaluation, highlighting the critical cognitive components of threat detection and assessment. It is important that leaders have a method for observing and evaluating their Soldiers’ abilities to identify and reason about threat cues in training environments. Soldiers who can make sense of their environments and effectively communicate their assessments to their peers and leaders will help ensure team safety and mission success in the operational environment.

We recommend testing this tool with a simple, brief training task (e.g., conducting a dismounted patrol). For example, leaders can manipulate the identified training environment to include inanimate threat cues such as disturbed earth, exposed wires, suspicious containers, as well as innocuous materials common to the environment. Leaders can brief their teams on the purpose of the exercise (e.g., “Identifying potential IED threats”), and then run their teams through the exercise. As each team engages in the exercise, the leader can observe individual Soldier performance regarding how many relevant cues were identified, how long it took to identify them (or how close, by distance, the Soldier was to a cue before identifying it), how long it took to communicate the threat to the team, and whether the Soldier’s assessment was based on current and relevant information. After each iteration of the exercise, leaders can change various aspects, including the number or location of cues, the task, the team leader, and the time allotted for completing the task. This will help leaders identify Soldier performance under different

conditions. After completing all exercises, leaders can review Soldier performance with each Soldier individually and/or with the team, identifying scores, justification for scores, and suggestions for how to improve performance.

Leaders can use the following template (Table 1) to organize an example training and evaluation. Some example behaviors and criteria are provided in Steps 5 and 6.

Table 1.

Example Template for Organizing Training and Evaluation

Step	Description
1	Identify the relevant operational environment and associated threats that Soldiers must be able to identify: [Notes]
2	Identify a location appropriate for setting up a training lane: [Notes]
3	Gather materials or identify relevant features of the environment that will represent cues to critical threats and cues to non-threats: [Notes]
4	Set up training lane to represent threat-present or threat-absent environment: [Notes]
5	Identify the critical behaviors associated with assessing threat risk in the training environment: • [Behavior 1]: <i>Identify relevant threats.</i> • [Behavior 2]: <i>Generate situational assessment.</i> • [Behavior 3]: <i>Communicate assessment to leader (and team).</i>
6	Assign scoring criteria to each critical behavior: • [Behavior 1]: <i>Identify relevant threats.</i> - o 0 = [Criterion]: <i>No threats identified.</i> - o 1 = [Criterion]: <i>Identified 1/3 threats.</i> - o 2 = [Criterion]: <i>Identified 2/3 threats.</i> - o 3 = [Criterion]: <i>Identified 3/3 threats.</i> - o 4 = [Criterion]: <i>All threats identified ahead of time limit.</i> • [Behavior 2]: <i>Generate situational assessment (SA).</i> - o 0 = [Criterion]: <i>No SA given.</i> - o 1 = [Criterion]: <i>SA was not entirely relevant to situation.</i> - o 2 = [Criterion]: <i>SA was generated, but only partially relevant.</i> - o 3 = [Criterion]: <i>Relevant SA was generated.</i> - o 4 = [Criterion]: <i>SA was used to guide actions of the team.</i> • [Behavior 3]: <i>Communicate assessment to leader (and team).</i> - o 0 = [Criterion]: <i>No SA was communicated.</i> - o 1 = [Criterion]: <i>SA was generated, but not communicated (effectively).</i> - o 2 = [Criterion]: <i>SA was generated, but not communicated to entire team.</i> - o 3 = [Criterion]: <i>Context relevant SA was provided within time limit.</i> - o 4 = [Criterion]: <i>Relevant SA was provided ahead of time limit.</i>

Table Continues

7	Convey general scoring criteria to Soldiers: [Notes]
8	Walk Soldiers through training lane: [Notes]
9	Select a Soldier to act as leader and run team through the lane: [Notes]
10	Observe individual and team behavior and assign scores accordingly: [Notes]
11	After team completes training lane, review scores individually/collectively: [Notes]
12	Repeat exercise with same or new lane configuration and time constraints: [Notes]

References

- Leins, D. A., Leonard, J., Zimmerman, L. A., Minchew, M. L., & Vowels, C. L. (2017). *Decision environment and heuristics in individual and collective hypothesis generation*. (Research Report 2013). Fort Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Leins, D. A., Leonard, J., Masters, G., & Vowels, C. L. (In Preparation). *The influence of expertise and decision environment on collective hypothesis generation*. (ARI Research Product). Fort Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Leins, D. A., Zimmerman, L. A., Mueller, R., Singer, J. T., & Vowels, C. L. (2016). *Improving visual threat detection: Operational needs assessment and implementation strategies*. (Research Product 2016-01). Fort Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Zimmerman, L. A., Leins, D. A., Marcon, J., Mueller, R., Singer, S. T., & Vowels, C. L. (2016). *Assessing threat detection scenarios through hypothesis generation and testing*. (Technical Report 1347). Fort Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Zimmerman, L. A., Leins, D. A., Marcon, J., Perelman, B., Singer, J. T., Mueller, R., & Vowels, C. L. (2013). *Improving visual threat detection: Research to validate the threat detection skills trainer*. (Technical Report 1329). Fort Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Zimmerman, L. A., Mueller, S. T., Daniels, J., Marcon, J. L., & Vowels, C. L. (2012). *Developing training exemplars for the requisite components of visual threat detection*. (Technical Report 1322). Fort Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Zimmerman, L. A., Mueller, S. T., Grover, J., & Vowels, C. L. (2014). *Determining the requisite components of visual threat detection to improve operational performance*. (Technical Report 1340). Fort Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Appendix A

Focus Group Slide Deck*



Hypothesis Generation at the squad level



Survey Approval Authority:
U.S. Army Research Institute for the Behavioral and Social Sciences

Survey Control Number: DAPE-ARI-AO-17-15

RCS: MILPC-3

Expires: 03/01/2018

1

*Any notes for the slides will appear below each slide throughout the Appendix.



Hypothesis Generation at the squad level



Research Program Purpose:

Improving decision making of Soldiers and leaders for home station training & deployment

Program highlights:

- Interviewed Soldiers and leaders (with and w/out deployment experience)
- Developed supplemental home station training (Threat Detection Skills Trainer)
- Study cognitive processes underlying decision making (ongoing)
- Develop measures to incorporate lessons learned (current)

2

Overview of research program: We interviewed Soldiers and leaders with and without deployment experience to identify the fundamental cognitive elements of threat detection. We used this information to generate computer-based training modules to help Soldiers acquire, more quickly, and practice the cognitive skills inherent in visual threat detection (Threat Detection Skills Trainer). In our last few tasks, we studied some of the cognitive processes involved in decision-making in threat detection across different contexts. For example, we explored decision-making under uncertainty; and, we most recently looked at conditions under which Soldiers might use mental shortcuts to make threat assessments. Today, we want to learn more about how we can further transition what we've learned from this program of research into training and evaluation. We want to know where our findings are relevant to your training and, when they are relevant, how we can help you measure and hopefully improve Soldier performance.



Incorporating Measures of Hypothesis Generation in Training and Evaluation



What is hypothesis generation?

Using features of the environment to generate explanations of the environment, for example:

Environmental Feature	Explanation
Large dead space	→ High threat risk
No cover/concealment	
Recent enemy activity	

Essentially, threat assessments of the environment.

3

Over the next few slides, we'll go over what we've learned to date.

For the purpose of our research, we have defined hypothesis generation as...



Hypothesis Generation in Context



Visual threat detection

Identifying relevant features in environment



4

In the context of visual threat detection, hypothesis generation involves attending to and identifying relevant features in the environment. This photo is one we used to develop the Threat Detection Skills Trainer. It was annotated by experienced Soldiers who identified threat-relevant features of the environment in the image. For example, places in the image that offered an enemy cover were identified frequently as threat relevant.



Hypothesis Generation in Context



Visual threat detection

Detecting relevant changes in environment



5

Testing whether Soldiers are sensitive to changes in an environment allow us to assess whether they are attending to relevant features of that environment. In addition, hypothesis generation in threat detection can also include change detection. Having a sense of the baseline status of an environment and then detecting changes from that baseline can help inform assessments of threat risk.



Hypothesis Generation in Context



Visual threat detection

Reasoning about the meaning of relevant features



"The brush and ditches can be used to conceal an IED" vs. "grassy area, dirt, and shrubs"

6

In many professions, experts reason about problems a little differently than do novices. For example, they tend to place the features of problems in context and they have the benefit of being able to match patterns of features with situations they've experienced in the past. So, when experts reason about the meaning of features, they often provide context and examples. In contrast, novices are more likely to simply identify features as relevant.



Hypothesis Generation in Context



Threat Assessment under Uncertainty

- Previous unit winning hearts/minds
 - Added potable water source
 - Organized neighborhood watch
 - Helped establish a school
- Recent upswing in violence
- Locals angry with U.S.
- Recent patrol confiscated IED components
- Detainee reports hidden IEDs
- New cleric in village has ties to Taliban
- You are on patrol w/dogs
- You see nothing suspicious, but dogs continue hitting on this location.



- What are the priority threats?
- How do Soldiers test their hypotheses?
- Do they change hypotheses after evaluating more info?

7

In testing how Soldiers generate and test threat assessments, we asked Soldiers to read through scenarios that offered differing levels of certainty and then identify the priority threats in the scenario and explain why they were priority threats. We then allowed Soldiers to search for additional information about the scenarios, to see how they would test their initial assessments – for example, would they seek information on relevant vs. irrelevant topics? Would they seek information that confirmed vs. potentially disconfirmed their initial assessments? Would they change their assessments? And, critically, would experienced Soldiers' processes look different from inexperienced Soldiers' processes?



Hypothesis Generation in Context

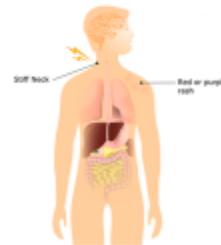
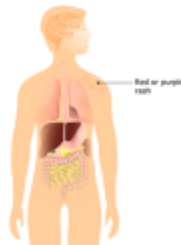


Mental shortcuts in Threat Assessment

- Familiar Task



- Unfamiliar Task



8

Most recently, we tested Soldiers use of heuristics – mental shortcuts – in generating threat assessments. We asked Soldiers to make assessments in familiar tasks (detecting threats in an operational environment) and unfamiliar tasks (diagnosing medical conditions). Sometimes these tasks were completed under time pressure, sometimes under no time pressure. They did this individually and collectively. After Soldiers made initial assessments, we changed the scenario – we added information to it over time – to see if the changes would affect Soldiers' assessments. Critically, we manipulated the order of how the scenario changed – in some scenarios, we made threat-relevant changes early in the evolution of the scenario, whereas in other scenarios, we made threat-relevant changes later in the evolution of the scenario. We wanted to see if the timing of the information would affect if and when Soldiers abandoned their initial assessment and adopted new ones. If Soldiers abandoned their assessments early, it meant that they were likely using a mental shortcut – even though they could evaluate more information about a scenario, they chose to change their assessments based on limited information (although in some cases, the limited set of information was also the most valuable set of information, making the use of the mental shortcut very efficient and effective).



What we found



- **Time pressure = Soldiers changed assessments sooner**
- **Time pressure = Smaller increases in threat ratings**
- **Time pressure = Abandon threat assessments in response to less informative environmental cues**
- **Time pressure led to lower quality decisions**
 - lowered threshold of informativeness to change threat assessment
- **Groups: Time together (especially during deployment) = equal contribution to discussion**

9

Here are some of our more interesting and useful findings. Time pressure appeared to influence how Soldiers employed heuristics in assessing threats. Not only did Soldiers change their hypotheses sooner when the early change was threat-relevant, but they also changed their hypotheses sooner when the early change was less relevant. Also, we tended to see that Soldiers who shared deployment experience, contributed more equally when discussing threat scenarios.



Performance elements to consider



- Do Soldiers appear to respond to the right cues?
- Do Soldiers communicate when they see something relevant?
- How long do they take to assess a situation?
- How do they support their threat assessments? Do they reference:
 - Training
 - Doctrine
 - Experience
 - Logic
- Do they exhibit appropriate confidence levels?
 - E.g., if they are unsure, do they report being unsure?

10

This is a list of questions to be asked when thinking about developing performance steps that include hypothesis generation.

Is there a word that you prefer in place of “assessment?”

Is there a word/term/phrase that you already use that covers the concept we’ve defined as “hypothesis generation” or “assessment?”



How can we implement findings?



Consider observable behaviors when developing performance steps?

- Did unit leader/Soldier respond differentially across iterations/conditions?
 - Time to report
 - Identification of high-value vs. low-value threat cues
- Did unit leader/Soldier communicate with different levels of effectiveness over time? (Evaluated on 0-4 scale at each iteration.)
- The leader/Soldier identified an appropriate course of action.

0	1	2	3	4
• Unacceptable (Unsatisfactory/ Not at all)	• Below Average (Sub-standard/ Performed some tasks)	• Average (Minimum standard/ Performed most tasks)	• Above Average (Standard/ Performed all tasks)	• Outstanding (Exceeds standard/ Performed all tasks and prepared for contingencies)

14

This slide continues with a set of questions about Soldiers, but also incorporates the leader.

Would it be valuable for Soldiers to conduct self-evaluations for these performance measures? This might offer insight to how Soldiers view their performance/competency vs. how trainers view the Soldiers' performance/competency.

How might you measure the suitability (accuracy, relevance, etc.) of an identified course of action stemming from an assessment?



Example measure implementation: Conduct a Security Patrol



Task Steps (typically evaluated as "GO, NO-GO, N/A")

1. Plan – Unit leaders gain/maintain situational understanding...
2. Unit leader receives OPORD (conduct security patrol)
3. Unit leader begins developing a plan
4. Prepare – Unit leader issues OPORD/FRAGORD as necessary...
5. The unit conducts rehearsals
6. The unit leader issues FRAGORDs to address changes to the plan
7. Execute – The unit conducts security patrol

15

Here are the overarching steps comprising the task, Conduct a Security Patrol. These task steps are ordinarily evaluated on the GO vs. NO-GO scale. But these overarching tasks comprise supporting tasks, some of which might be affected by cognitive processes like hypothesis generation (Execution phase of the task may have multiple supporting tasks appropriate for evaluating cognitive processes)....



Evaluate Task 7: Execute Security Patrol



Supporting Collective Task(s)	Potential Evaluation Measures
Conduct Tactical Movement in an Urban Area	GO, NO-GO, NA; Quality of performance on 5-point scale
Supporting Individual Task(s)	Potential Evaluation Measures
Conduct a Mounted/Dismounted Patrol	GO, NO-GO, NA; Quality of performance on 5-point scale
Prepare a Situation Report (SITREP)	
Conduct a Security Patrol	
Conduct a Mounted/Dismounted Patrol	Potential Evaluation Measures
Time taken to report threat assessment	Quantified (x minutes) Quality of performance on 5-point scale
Quality of threat assessment	5-point scale
Identified threat cues	GO, NO-GO, NA; 5-point scale

15

....for example, executing the security patrol involves subtasks like “conduct tactical movement in an urban area” (collective) and “conduct a mounted/dismounted patrol” (individual). Can these be broken down even further? For example, can you observe, measure, and evaluate some of the component parts of “Conduct a Mounted/Dismounted Patrol?”

Can you evaluate the time it takes to report a threat assessment? The quality of the threat assessment? And/or whether/how well did the Soldier identify relevant threat cues?

So, in addition to evaluating tasks broadly, can you scope down to evaluate some of the critical individual tasks/operations that support the broader tasks?



**Informal evaluations of Task 7:
Execute Security Patrol – Dismounted Patrol**



Crawl: Demonstrate processes of effective/efficient hypothesis generation (threat assessment) by:

- Pointing out relevant threat cues in an operational environment,
- Reasoning about why they are relevant, and
- Demonstrating optimal ways of communicating threats.

Walk: Walk students through completing relevant tasks and ask students at regular intervals to:

- Identify priority threats,
- Describe why they are priorities, and
- Seek input from teammates.

Run:

- Students complete relevant tasks at faster pace.
- Instructor evaluates Soldier performance using new objective/subjective metrics.

17

Here is an example of how you can implement and evaluate cognitive processes, like hypothesis generation (threat assessment), using a crawl-walk-run format.

Can you incorporate any of these concepts or performance measures as informal evaluations? For example, can you demonstrate effective hypothesis generation in an example task, then walk Soldiers through the steps sequentially (e.g., what to look for, what not to look for, pitfalls of mental shortcuts to be aware of), and finally evaluate how they do this on their own?

These are the final concepts we wanted to cover. Does anyone have any further feedback we should consider?



Results of Informal evaluations?



Expected Outcomes:

Soldiers

- Able to better assess potentially threatening environments
- As a result of improved assessment, make better decisions
- Improve collective (squad) performance

Trainers

- Better, objective idea of how well your Soldiers are performing
- Improved AARs based on scored data
- Focusing on the thinking gives a more granular understanding of performance
 - As a result, can provide more focused feedback at the individual and collective level.

18

In conclusion, this is what we hoped to accomplish with this research and the related activities.

Any final questions or comments?

Appendix B

Description of Steps for the Squad Threat Assessment (STA) and STA versions

**The rating scale (0 to 4) applies to the Rating under Threats Identified (Step 1), and the Quality of Assessments (Steps 1 and 3). Both the leader/trainer and subordinates should have a clear understanding of how performance will be rated. Some of the examples below are more specific to detecting IEDs. However, the assessment was developed to assist in evaluating several tasks that involve detecting and defeating different types of threats. This assessment might be best employed in the crawl or walk phase of training as it requires the recording of multiple assessments and the communication of those amongst the team.

STEP 1.

Threat Identified

Raw – The number of threats identified out of total possible. Example: 1 / 3.

Time – simply the total time taken to identify the threats. A rough estimate will suffice if no external device is available. *In some situations, distance to threat or target may be more important than time taken to identify the threat (of course, time and distance are often linked).

**Rating – Given the total number identified and the time it took to identify them, what is the Soldier's overall rating? For example, if they only identified 1 / 3 threats and it took longer than should be expected (30 seconds, 60 seconds, etc.), the Soldier should receive a low rating.

Situational Assessment

This is the first assessment given by any team member. It is a brief statement of their understanding of the primary threat(s). Whether they identified any threats, can they tell you what is going on?

Time taken to Commo Assess. With Team (Leader)

(How quickly) did they tell their team and/or leader?

****Quality of Assessment**

This is a rating of their situational assessment. The more detail they give, the better they understand the relevance or primary importance of the threat within the situational context.

If they just provide, "Possible IED," they may only be rated as a 1. If they provide context, such as, "This is a possible buried IED; dirt discoloration; probably pressure plate, no wires seen; recent intel tells us enemy may now be using this tactic," they are likely rated a 4.

STEP 2.

Leader provides his assessment and possible course of action. This could be based on situational assessments already given or a novel/original assessment from the leader.

STEP 3.

Alternative Situational Assessment

This is a brief statement of understanding. This may be given by the same Soldier, but will be alternative or additional to the Situational Assessment from any given in Step 1.

Time to commo to leader

Again, this would be how quickly they informed their team and/or leader of their assessment?

****Quality of Assessment**

Again, this is a rating of the Soldier's situational assessment. The more detail they give, the better they understand the relevance of the threat.

Feedback from Team

This would be any reaction from the team to the given situation assessment.

STEP 4.

Change/update in Leader's Course of Action (indicate # of assessment COA was based on)

This could be based on any of the Situational Assessments indicated. The COA should allow the team to mitigate the threat/risk in an effective manner.

Result/Outcome of COA

This is simply what happened after the action was taken. This feedback is very important for leaders and teams working through/practicing/training on different techniques and maneuvers as accurate feedback will help them sharpen their skills.

STEP 1.

Squad Threat Assessment							Leader:
Exercise / Task:			Location / Lane:			Date:	
Rating Scale: 0 (Unsatisfactory), 1 (Minimum Standard), 2 (Standard), 3 (Above Standard), 4 (Exceeds Standard)							
Iteration:		Threats Identified		Situational Assessment (Number each assessment)	Time taken to communicate assessment	Quality of assessment	Feedback from Team
Soldier	Raw	Time	Rating				
					<u>STEP 2.</u>		
					Leader's Assessment	Leader's Original COA	
<u>STEP 3.</u> <u>STEP 4.</u>							
Soldier	Alternative Situational Assessment	Time to commo to leader	Quality of Assessment	Feedback from Team	(If applicable) Change/update in Leader's COA	Result / Outcome of COA	

*An alternative form for evaluating individual Soldier threat assessment performance:

<i>Squad Threat Assessment</i>							Leader:		
Exercise:							Date:		
Rating Scale:		0 (Unsatisfactory), 1 (Minimum Standard), 2 (Standard), 3 (Above Standard), 4 (Exceeds Standard)							
	Threats Identified		Time taken to Identify Threat		Communicated Assessment with Team		Quality of Assessment		
Soldier	Raw	Rating	Raw	Rating	Raw	Rating	Raw	Rating	Comments:
Group Assessment:									

*Version B focuses largely on the performance of individual Soldier performance, but a group rating can be given at the bottom. One method for calculating the Group Assessment is simply to sum the ratings in each column.

Appendix C

Trainer's Slide Deck*

Trainers/instructors can use these slides to help guide them through developing exercises for training and evaluating threat assessment at the team/squad level.



*Notes accompanying the slides will appear below each slide throughout the Appendix.

This set of slides will help guide trainers/instructors through some of the fundamental cognitive concepts underlying hypothesis generation or situation assessment.



Improving decision making for threat detection and assessment



Training objectives:

- Observe and measure critical cognitive processes underlying effective threat detection/assessment.
- Track Soldier progress for detecting threats in a more meaningful / functional way.
- Provide better feedback to Soldiers to support improvement in decision making.

2

The training objectives are listed on the slide. These objectives could be supplemented by the instructor, given the purpose of the training.

Later in the slide deck, examples of a Squad Threat Assessment are provided.



Some cognitive processes explained



Perception – recognizing, organizing, and giving meaning to sensations received from environmental stimuli.

Example: seeing disturbed earth and becoming aware that it differs from undisturbed earth around it.

Attention – focusing on a small amount of information embedded in a larger context.

Example: concentrating on an area of disturbed earth because it differs from the area around it.

Prioritization – valuing and ranking pieces of information based on a relevant criterion.

Example: ranking an area of disturbed earth as more threatening than an abandoned structure, based on knowledge of recent IED activity.

Heuristics – mental shortcuts or “rules of thumb” that allow for rapid decisions based on a fraction of the information available.

Example: deciding that an environment is highly threatening based solely on the presence of disturbed earth.

3

These are some of the most relevant cognitive processes important for generating hypotheses or situation assessments. Very brief definitions are provided as well as examples. Note: these processes are universal in human beings and can occur very rapidly (sometimes within milliseconds). These processes also run outside of our conscious awareness. In other words, we are unaware of the actual execution of these processes; however, we are aware, and utilize, the output from these processes. Our brains will naturally attempt to make sense of its surroundings based on the information available. Further, even though we can't see or touch these mental processes, they are always taking place in all of us to help us understand/make sense of our surroundings, threatening, or otherwise.

And, as we become more expert or proficient at tasks within our given domain, these processes become functional. That is, these cognitive processes help provide relevant information, often quicker than when we are novices in a domain, while also 'ignoring' irrelevant information.



- Perception, attention, prioritization, & heuristics all support: Hypothesis generation or using features of the environment to generate explanations of the environment, for example:

Environmental Feature	Explanation
Large dead space	→ High threat risk
No cover/concealment	
Recent enemy activity	

- In threat detection, hypotheses are essentially threat assessments of the environment.

Fundamentally, hypothesis generation is another name for attempting to understand or make sense of our surroundings. This overarching process is aided by cues/information/data provided by the processes identified on the previous slide. In an operational environment, especially one with variable types of embedded threats, being able to generate valid hypotheses or assess situations effectively is imperative.



Hypothesis Generation in Context



Visual threat detection

Identifying relevant and irrelevant features in an environment

Some features will be more/less relevant given their specific surroundings.

- In this image, clusters of numbers indicate threat-relevant features identified by multiple Soldiers with combat experience.
- Soldiers were primarily concerned with the circled areas.



5

One of the components of efficient hypothesis generation is identification of the relevant features of an environment apart from irrelevant features. In the image above, Soldiers with combat patrol experience indicated areas in the photo that was most concerning to them (areas that could possibly conceal threats such as snipers, for instance).



Hypothesis Generation in Context

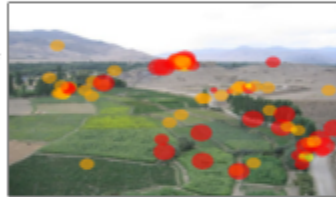


Visual threat detection

Identifying severity and likelihood

Soldiers need a comprehensive understanding that some threats are more severe and some are more likely, given their particular environment.

- In this image, larger clusters of dots indicate greater likelihood of potential threats as provided by Soldiers. The redder the dots, the more severe the threat indicated.



- In this image, Soldiers indicated where and what they thought required their attention, as indicated by red dots.



6

Part of the process of generating good hypotheses about a situation, especially in an environment that might contain threats, is determining which aspects contain information about severity and likelihood. In these images, Soldiers were asked to indicate areas of concern if they had to maneuver through these different types of terrain.



Hypothesis Generation in Context



Visual threat detection

Detecting and prioritizing relevant vs. irrelevant changes in an environment
Not all changes in an environment will lead to changes in threat assessment
In the photo on the right, red circles identify changes from the photo on the left.



Another important aspect that helps guide decision-making in operational settings is the ability to detect changes in locations that you've patrolled before. Of course, not all changes indicate that threats are now absent or present. However, being aware of changes in the environment (whether in buildings, terrain, or people) can help Soldiers make better judgments or assessments across a variety of contexts. One simple set of questions is to have Soldiers ask if this is a change, Yes or No. And, then ask whether the change matters, based on what they currently know about this particular environment.



Hypothesis Generation in Context



Visual threat detection

Reasoning about the meaning of relevant features

- Different levels of experience yield different explanations for environments:
 - Experienced decision makers often include context and examples: "The brush and ditches can be used to conceal an IED."
 - Inexperienced decision makers often omit context: "Grassy area, dirt, and shrubs."



8

Developing good hypotheses can help you with higher order processing or establishing sound reasoning about the situations you are in. And, as Soldiers gain experience, their ability to develop better, comprehensible thoughts to help solve problems across different operational settings will become more efficient. For instance, Soldiers with multiple combat deployments were better at putting different threats into the most relevant, larger context or narrative. That is, if they detected a threat, they also recognized its relevancy or why it mattered for their particular situation.



Hypothesis Generation in Context



Threat Assessment under Uncertainty

Given the following information about this image:

- Previous unit winning hearts/minds
 - Added potable water source
 - Organized neighborhood watch
 - Helped establish a school
- Recent upswing in violence
- Locals angry with U.S.
- Recent patrol confiscated IED components
- Detainee reports hidden IEDs
- New cleric in village has ties to Taliban
- You are on patrol w/dogs
- You see nothing suspicious, but dogs continue hitting on this location



- How do Soldiers determine the most relevant/priority threats?
- How do Soldiers test their hypotheses (or how do they check their assessments)?
- Do Soldiers revise/update their hypotheses after evaluating more info?

9

Most operational environments will involve some level of uncertainty. However, Soldiers can develop the reasoning skills to help them better understand, at least parts of, the environments they will have to operate in. For example, Soldiers and teams can practice their skills by working through brief scenarios with limited, or large, amounts of information. One photo can provide some visual context to better embed the Soldier in the situation. The questions at the bottom (and others) can be used regardless of the context to help Soldiers develop good reasoning about uncertain situations.



Hypothesis Generation in Context



Mental shortcuts in threat assessment

Again, heuristics help us to make quick decisions based on limited information.

- The arrows in the top image indicate threat-relevant features.
- The arrows in the bottom image indicate less relevant features.
- If Soldiers perceive these different features at different times, for instance, they see the roof before the doors, would they generate different hypotheses?
- In other words, how relevant must a feature be before a Soldier evaluates it as a potential threat?
- As Soldiers become more proficient/expert at their jobs, it is important to note that heuristics are often underlying their enhanced ability.



10

Mental shortcuts or heuristics are used by all of us to make sense of situations. Heuristics, which can be thought of as a set of simple decision rules, can be powerful cognitive tools because they allow us to generate hypotheses on a limited set or incomplete amount of information. For instance, the rules might direct the brain what to search for, what is deemed a relevant cue or piece of information and when to stop searching for more cues or information.

As Soldiers become more expert or proficient at their respective jobs, heuristics will help guide visual search of the environment and help develop hypotheses or situation assessments of their operational environment.



Hypothesis Generation in Context



Generating hypotheses across conditions:

- As environments change, we want Soldiers to be able to perceive, attend to, prioritize, and reason about environmental cues reliably.
- The ability to generate good hypotheses across environments and conditions is a hallmark of expertise.
- To help Soldiers become experts in threat assessment:
 - Train them under a variety of conditions
 - Train them in a variety of environments
 - Evaluate them consistently and provide meaningful feedback

11

Again, as Soldiers become more proficient/expert, their critical thinking, problem solving and reasoning about relevant and irrelevant information will improve. However, Soldiers and teams need to be challenged under a variety of conditions, in a variety of environments and need to be provided feedback specific to their individual and small group performance.



Some important variables influencing how Soldiers assess threats



1. **Time pressure**
Individual Soldiers generated hypotheses based on less informative cues when under time pressure
2. **Familiarity**
Soldiers made better decisions when they were more familiar with the decision environment, that is, when they recognized the relevance of environmental cues.
3. **Individual vs. collective decision-making**
Soldiers working together were less susceptible to the influence of time pressure

12

These are a few of the variables that were found to impact performance at the individual and small group (squad) level. These are worth considering when developing training for threat assessments.



Performance elements to consider



- Do Soldiers appear to respond to the right cues?
- How long do they take to assess a situation?
- Do Soldiers communicate when they see something relevant?
- How do they support their threat assessments? (This will help determine the quality of their assessment).
 - Do they reference:
 - Training
 - Experience
 - Logic
 - Doctrine
- Do they exhibit appropriate confidence levels?
 - If they are unsure, do they report being unsure?

13

Above are several performance elements that should be considered for evaluating performance in situations that may involve threats or varying degrees of uncertainty. These questions echo the processes indicated in earlier slides and form the basis for the Squad Threat Assessment, shown in a few slides.



Translating observations to performance measures



Observations:

Identification of relevant threat cues?
Time to report a threat assessment?
Communicated potential threats?
How good is the assessment?

Performance rating:

0	1	2	3	4
• Unacceptable (Unsatisfactory/ Not at all)	• Below Average (Sub-standard/ Performed some tasks)	• Average (Minimum standard/ Performed most tasks)	• Above Average (Standard/ Performed all tasks)	• Outstanding (Exceeds standard/ Performed all tasks and prepared for contingencies)

14

The observations listed are some of the primary behaviors across many operational settings when trying to determine what is going on in the surrounding environment.

An example of a performance ratings scale is provided above. Though tasks are typically evaluated as GO or NO GO, scales that provide more gradation or levels of performance allow for more granular feedback at the individual and collective level. More specific feedback can be helpful for the trainer and the trainee. What determines the range of performance levels is often best left up to the trainers or experts in a given domain. They are most likely to know what separates performance at finer levels.



How to implement measures



- **Consider (operational) environmental variables/features when designing task conditions, for example:**
 - Time constraints – How quickly must the squad complete the task to be effective?
 - Familiarity – How much experience does the squad have conducting this task and how will this influence your expectations for their performance?
 - Uncertainty – How much information will you exclude/include in the pre-task briefing to make the task more/less challenging?
 - Presence of high-value (relevant) vs. low-value (irrelevant) threat cues – How many of each cue will you include and when will the squad encounter them?
- Use these variables to help determine rating criteria to be used in conjunction with the *Squad Threat Assessment*.

15

These are elements that should be considered when developing training lanes/scenarios. Many of these factors are inherent in existing training, but can be modified by the trainer to meet the needs and challenge his/her individual Soldiers and small teams.



How to evaluate performance: Identification & Timing



Consider observable behaviors

- Did Soldiers/leaders prioritize threats appropriately?
 - "The Soldiers/leader identified the priority threat at the objective." (Y/N)
 - Record the number of relevant threats identified & translate to rating (0-4 scale).
- How long did it take to report/verbalize a threat assessment?
 - "The unit leader/Soldier reported a threat assessment within the allotted time." (Y/N)
 - Record an elapsed time & translate to rating (0-4 scale)

16

These brief descriptions of observable behavior will help guide trainers/instructors/leaders when completing the Squad Threat Assessment. Two primary behaviors that should be paid attention are whether or not Soldiers and leaders are prioritizing threats and how long it takes for them to communicate their assessments.

*Experienced Soldiers and leaders may supplement these behaviors with others based on their first-hand deployment (and combat) experiences. The Squad Threat Assessment can be modified accordingly.



How to evaluate performance: Communication & Quality



- Communication of threat assessment? (Y/N)
 - Did unit leader/Soldier support threat assessment with:
 - Context (experiential, logical reasoning)? (Y/N)
 - For example, "I've seen this before. Threat risk is high because we have no cover/concealment and if we continue down this route, we'll have no way out if ambushed."
 - Simple cue identification? (Y/N)
 - For example, "Abandoned car" or "empty marketplace."
 - Translate communication behavior to rating (0-4 scale)
- Did unit Soldiers/leader evaluate alternative threat assessments? (Y/N)
- What was the overall quality of the threat assessment? (0-4 scale)?

17

Again, these are components that can be captured with the Squad Threat Assessment and should be considered when training and evaluating Soldiers to detect threats across a variety of potential operational contexts.



Squad Threat Assessment



• Use this form to guide development of evaluation criteria, and record and track Soldier performance.

• Version B – Incorporates individual and collective performance as well as leader decision making.

Squad Threat Assessment							Leader:	
Exercise / Task: <i>Demounted Patrol</i>				Location / Lane:			Date:	
Rating Scale: 0 (Unsatisfactory), 1 (Minimum Standard), 2 (Standard), 3 (Above Standard), 4 (Exceeds Standard)								
Iteration:	Threats Identified	Situational Assessment (Number each assessment)		Time taken to communicate assessment	Quality of assessment	Feedback from Team		
Soldier	Raw Time Rating							
<i>SPC JONES</i>	<i>1/3</i>	<i>20 sec</i>	<i>2</i>	<i>#1 Possible buried IED; secure perimeter; get Pos ID</i>	<i>60 sec</i>	<i>3</i>		<i>Agreed, probably IED</i>
<i>PFC JONESY</i>	<i>1/3</i>	<i>10 sec</i>	<i>3</i>					
<i>SGT JOHN</i>	<i>0/3</i>	<i>0</i>					STEP 2.	
<i>SGT JOHNNY</i>	<i>0/3</i>	<i>0</i>					Leader's Assessment	Leader's Original COA
<i>PV2 SMITH</i>	<i>0/3</i>	<i>0</i>			<i>#2 Choke point, IEDs possible</i>			Conduct hasty recon
STEP 3.								
Soldier	Alternative Situational Assessment	Time to come to leader	Quality of Assessment	Feedback from Team	(If applicable) Change/update to Leader's COA	Result / Outcome of COA		
<i>PFC JONESY</i>	<i>#1 Possible sniper concealment</i>	<i>30 seconds</i>	<i>2</i>	<i>Not likely</i>	<i>#1 Secure area; Move to safe position; call EOD</i>	<i>IED Destroyed in place</i>		
<i>SGT JOHN</i>	<i>#4 Looks like a fatal funnel</i>	<i>75 seconds</i>	<i>2</i>	<i>None</i>				

19

This version of the Squad Threat Assessment focuses more on the generation of assessments or hypotheses by different members of the team and the integration of that information to make a final determination. Thus, the emphasis is on the critical thinking and communication that should be going on within the team as they move through different training situations.

In Step 1, Soldiers would provide their individual assessments of the situation. Not all Soldiers may provide an assessment. In this example, only SPC Jones provides a situational assessment. The assessments should be ordered in the number they are communicated. In some situations Time may not be as relevant as Distance to a potential threat or target. In that case, for example, the Assessment could be modified to indicate Distance instead of Time, under Threats Identified.

In Step 2, the Leader would provide his/her assessment and what she/he thinks would be a good COA. The assessment could be based on what Soldiers have already communicated or on his/her own novel understanding of the situation.

In Step 3, Soldiers might generate and communicate an alternate assessment based on assessments provided in Steps 1 and 2 or could be a new assessment.

In Step 4, the Leader indicates his/her final assessment and COA. The results of his/her choice should also be recorded. This could be an important piece of feedback as it allows the team to understand how effective their reasoning was as well as the result of the COA they chose.



How to evaluate performance



Consider observable behaviors across conditions

- Did Soldiers/leaders respond differentially across iterations/conditions?
 - Identification of high-value vs. low-value threat cues
 - Time to report
- Did Soldiers/leaders communicate with different levels of effectiveness over time?

Use the following rating scale to compare performance across trials:

0	1	2	3	4
• Unacceptable (Unsatisfactory/ Not at all)	• Below Average (Sub-standard/ Performed some tasks)	• Average (Minimum standard/ Performed most tasks)	• Above Average (Standard/ Performed all tasks)	• Outstanding (Exceeds standard/ Performed all tasks and prepared for contingencies)

20

As training lanes are modified to be more/less difficult, the same Squad Threat Assessment format can be used to evaluate performance.

As training lanes are conducted, hopefully as often as possible to achieve a high level of performance, and the difficulty level changes, it is important for trainers to remember that Soldiers and leaders may react more/less effectively as the level of difficulty increases.



Crawl-walk-run approach



- **Crawl:** Demonstrate processes of effective/efficient hypothesis generation (threat assessment) by:
 - Pointing out relevant threat cues in an operational environment,
 - Reasoning about why they are relevant, and
 - Demonstrating optimal ways of communicating threats.
- **Walk:** Walk students through completing relevant tasks and ask students at regular intervals to:
 - Identify priority threats,
 - Describe why they are priorities, and
 - Seek input from teammates.
- **Run:**
 - Students complete relevant tasks at faster pace.
 - Instructor evaluates Soldier performance using new metrics.

21

The typical training format is to use a crawl-walk-run approach. This is simply laid out above, for the current purposes of training better hypothesis generation/situation assessments.



Expected Outcomes:



Soldiers

- Better able to assess potentially threatening environments
- As a result of improved assessment, make better decisions
- Improve collective (squad) performance

Trainers

- Better, objective idea of how well your Soldiers are performing
- Improved AARs based on scored data
- Focusing on the thinking gives a more granular understanding of performance
 - As a result, can provide more focused feedback at the individual and collective level.

22

These are some of the expected outcomes for Soldiers and Trainers that should be encapsulated in training and evaluating hypothesis generation. Again, experienced trainers can provide more detail, and other expected outcomes, based on their experience and, thus, knowledge of what it takes to prepare for the operational environment.