



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

**Research Report 1885**

**Training Effectiveness Assessment of Red Cape:  
Crisis Action Planning and Execution**

**Peter S. Schaefer, Scott B. Shadrick**  
U.S. Army Research Institute

**Jeff Beaubien**  
Aptima, Inc.

**Brian T. Crabb**  
U.S. Army Research Institute

**February 2008**

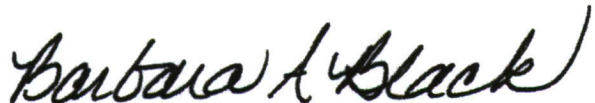
Approved for public release; distribution is unlimited.

**20080424046**

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

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

**Authorized and approved for distribution:**



**BARBARA A. BLACK, Ph.D.  
Research Program Manager  
Training and Leader Development**



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

---

Technical reviews by:

Brooke Schaab, U.S. Army Research Institute  
William R. Sanders, U.S. Army Research Institute

**NOTICES**

**DISTRIBUTION:** Primary distribution of this Research Report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, Attn: DAPE-ARI-ZXM, 2511 Jefferson Davis Highway, Arlington, Virginia 22202-3926.

**FINAL DISPOSITION:** This Research Report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

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

## REPORT DOCUMENTATION PAGE

<b>1. REPORT DATE (dd-mm-yy)</b> February 2008			<b>2. REPORT TYPE</b> Final			<b>3. DATES COVERED (from... to)</b> April 2007 to September 2007					
<b>4. TITLE AND SUBTITLE</b>  Training Effectiveness Assessment of Red Cape: Crisis Action Planning and Execution						<b>5a. CONTRACT OR GRANT NUMBER</b>  _____					
<b>6. AUTHOR(S)</b>  Peter S. Schaefer, Scott B. Shadrick (U.S. Army Research Institute), Jeff Beaubien (Aptima, Inc.), and Brian T. Crabb (U.S. Army Research Institute)						<b>5b. PROGRAM ELEMENT NUMBER</b> 622785					
						<b>5c. PROJECT NUMBER</b> A790					
						<b>5d. TASK NUMBER</b> 272					
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: DAPE-ARI-IK 121 Morande Street Fort Knox, KY 40121-4141						<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>  _____					
									Aptima 12 Gill Street, Suite 1400 Woburn, MA 01801		
									_____		
<b>9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b>  U.S. Army Research Institute for the Behavioral and Social Sciences 2511 Jefferson Davis Highway Arlington, VA 22202-3926						<b>10. MONITOR ACRONYM</b> ARI					
<b>12. DISTRIBUTION/AVAILABILITY STATEMENT</b> Approved for public release; distribution is unlimited.						<b>11. MONITOR REPORT NUMBER</b> Research Report 1885					
						_____					
<b>13. SUPPLEMENTARY NOTES</b>  Subject Matter POC: Peter S. Schaefer											
<b>14. ABSTRACT (Maximum 200 words):</b> The crisis response training program Red Cape: Crisis Action Planning and Execution uses theme-based training and multimedia scenarios to instill expert thinking patterns in crisis response personnel. The training program was assessed in workshops conducted with the Army National Guard. Quantitative data indicate that cognitive performance—as reflected in both independent and self assessments—improves as a function of training. Furthermore, as training progressed, inflation in self-assessments decreased, indicating more accurate self-assessment of cognitive performance. Participant feedback from the participants indicates that the training program was seen as useful, efficient, and effective in improving crisis response cognitions.											
<b>15. SUBJECT TERMS</b> Crisis Response Training, Interagency Training, Training Assessment, Theme-Based Training, Overconfidence											
<b>SECURITY CLASSIFICATION OF</b>			<b>19. LIMITATION OF ABSTRACT</b>  Unlimited		<b>20. NUMBER OF PAGES</b>  _____		<b>21. RESPONSIBLE PERSON</b> Ellen Kinzer Technical Publication Specialist 703-602-8047				
<b>16. REPORT</b>  Unclassified	<b>17. ABSTRACT</b>  Unclassified	<b>18. THIS PAGE</b> Unclassified									

**Research Report 1885**

**Training Effectiveness Assessment of Red Cape:  
Crisis Action Planning and Execution**

**Peter S. Schaefer, Scott B. Shadrick**  
U.S. Army Research Institute

**Jeff Beaubien**  
Aptima, Inc.

**Brian T. Crabb**  
U.S. Army Research Institute

**ARI-Fort Knox Research Unit**  
**James W. Lussier, Acting Chief**

**U.S. Army Research Institute for the Behavioral and Social Sciences**  
**2511 Jefferson Davis Highway, Arlington, Virginia 22202-3926**

**February 2008**

---

**622785A790**

**Personnel Performance  
and Training Technology**

Approved for public release; distribution is unlimited.



# TRAINING EFFECTIVENESS ASSESSMENT OF RED CAPE: CRISIS ACTION PLANNING AND EXECUTION

## EXECUTIVE SUMMARY

---

### Research Requirement:

Successful crisis response efforts require the effective collaboration of various agencies. However, effective interagency collaboration in turn requires frequent interagency training. Such events are rare, and tend to be large scale readiness exercises focusing more upon macro inter-agency behaviors than individual and small team learning. Thus, the events tend to be both too infrequent and lacking the attributes required to foster the development of expertise. To meet this need, the Red Cape: Crisis Action Planning and Execution training program was developed in a collaborative research effort by the U.S. Army Research Institute for the Behavioral and Social Sciences and the Indiana Army National Guard. Red Cape utilizes a theme-based training methodology and brief, engaging multimedia scenarios to instill the appropriate expert thinking patterns in those military, civil and government agencies involved in crisis response efforts. There is a need, however, to evaluate the effectiveness of the training materials.

### Procedure:

The development and content validation of Red Cape has been detailed in a previous report (Shadrick, Schaefer, & Beaubien, 2007). The current report describes the utilization of Red Cape training by military, civil, and government crisis response personnel to instill expert thinking patterns. The workshops were hosted at the request of the Indiana Army National Guard. Cognitive performance—the ability to generate expert considerations—was assessed by independent raters and training participants. Participant feedback regarding various aspects of the training was elicited at the end of each workshop session.

### Findings:

Independent assessment of cognitive performance indicated that the number of vignettes trained was significantly correlated with individual mean cognitive performance ( $p < .01$ ). Analysis of the self-assessment of cognitive performance demonstrated that training also reduced score inflation, suggesting that the theme-based method can help training participants become more accurate judges of their own progress. Participant feedback indicated that Red Cape was seen as effective, realistic, and relevant.

### Utilization and Dissemination of Findings:

The research supports the use of theme-based training to instill higher-order cognitive skills. It demonstrates a low-cost, effective way to maximize the ability of diverse stakeholder groups to cooperate and coordinate in a domain in which failure is measured in lives and property. The Red Cape training is intended to maximize benefits from participating in large-scale readiness exercises. The training has been implemented with the Army National Guard

units and their civil-military, interagency partners. The training program has been briefed to the Commanding General and Deputy Commanding General of the U.S. Army Armor Center, and the results have been briefed to representatives of the U.S. Army Training and Doctrine Command, the U.S. Army Management Staff College, and to multiple Army National Guard units.

TRAINING EFFECTIVENESS ASSESSMENT OF RED CAPE: CRISIS ACTION  
 PLANNING AND EXECUTION

CONTENTS

---

	Page
Introduction.....	1
Think Like a Commander Training Program.....	2
Theme-based Training for the Domain of Crisis Response.....	3
Content and Use of the Red Cape Training Package.....	4
Army National Guard Workshops .....	5
Method .....	5
Measures .....	6
Analyses.....	6
Results.....	8
Discussion .....	12
Conclusions.....	14
References.....	15
Appendix A Themes of Crisis Response .....	A-1
Appendix B Army National Guard Workshop Scenarios.....	B-1
Appendix C Air National Guard Workshop Scenarios.....	C-1

List of Figures

Figure 1. Ratings of cognitive performance as a function of workshop.....	8
Figure 2. Ratings for confidence in trained skills question .....	9
Figure 3. Ratings for challenge level of training exercise question.....	10
Figure 4. Ratings for training expectation question.....	11
Figure 5. Ratings for position relevance question .....	11



# TRAINING EFFECTIVENESS ASSESSMENT OF RED CAPE: CRISIS ACTION PLANNING AND EXECUTION

## Introduction

Crisis response efforts require the joint capabilities of military, civilian, and governmental agencies. However, several factors prevent agencies from collaborating as effectively and efficiently as they might. Some of these factors include specialized terminologies which make cross-agency communication difficult, as well as standing operating procedures (SOPs) which may differ from agency to agency, or may even be contradictory (Dixon, 2006). One obvious way to overcome such obstacles would be to engage in interagency training exercises which allow for increased familiarity with the terminology and SOPs of other agencies.

However, Glenn et al., (2006, pp. 200-201) noted: “Events that substantively bring military decision makers in contact with other-agency representatives, indigenous authorities, or commercial enterprises of relevance to operations are extremely rare.” The problem is further compounded by the fact that those infrequent events which do bring the relevant organizations together are usually large scale readiness exercises which focus more on macro inter-agency behaviors than individual and small team learning. Thus, the events are both too infrequent and lacking the attributes required to foster the development of expertise.

Training events which aim to develop expertise often follow the proven principles of ‘deliberate practice’ (Ericsson, Krampe, & Tesch-Römer, 1993). Following these principles allows training participants to:

1. See what expert behavior “looks like.”
2. Perform the task and note discrepancies from the expert model.
3. Repeat the task, successively approximating the expert model through conscious attention.
4. Repeat the task until behavior is habitual.
5. Reach criterion performance, and then repeat without conscious attention.

Much of the research on deliberate practice has focused upon the development of procedural skills, such as physical agility in canines (Helton, 2007), lap completion time in swimmers (Hodges, Kerr, Starkes, Weir, & Nanaidou, 2004), and skill maintenance in elite pianists (Krampe & Ericsson, 1996). However, there is some precedent for using deliberate practice principles to train cognitive skills as well. For example, the former Soviet Union used these kinds of methods to instill expert thinking patterns in chess masters (Alburt, 2003). Unlike most other nations, prominent researchers in the Soviet Union regarded thinking as a behavior to be trained. More recent research has focused on how deliberate practice may be used in various domains, including medical diagnosis (Guest, Regehr, & Tiberius, 2001) and case analysis skills in lawyers (Farmer & Williams, 2005).

Researchers at the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) hypothesized how deliberate practice might be adapted to the training of tactical thinking skills. However, two initial problems had to be surmounted. First, what was to be trained? And secondly, how could the thoughts of the training participants be measured so that training

efficacy could be estimated? The answer to both questions lay in the development of theme-based training.

Themes are elements or characteristics of expert thinking which are applicable across a range of situations. In ARI's training for tactical thinking, themes of expert tactical thinking were developed via observation, interviews, and experiments (Lussier, 1998; Ross & Lussier, 2000). The themes formed the basis of a cognitive model of tactical expertise which was used to analyze tutorial sessions and mental models of tactical thinking (Phillips, Shafer, Ross, Cox, & Shadrick, 2005). That cognitive model is the cornerstone of the Think Like a Commander (TLAC) training program, an instantiation of the theme-based training method. An overview of TLAC and the theme-based method is provided below.

### *Think Like a Commander Training Program*

The TLAC program instills tactical thinking skills by presenting participants with a series of complex and unpredictable tactical scenarios. While each scenario may have multiple successful solutions, each scenario has a set of considerations which are situational factors that play a role in expert tactical thinking (Lussier, Shadrick, & Prevou, 2003). Each one of the considerations corresponds to one of the overarching themes of expertise. Thus, although the themes themselves are applicable across all scenarios, the manner in which a theme is instantiated in a scenario is determined by the appropriate considerations. This approach was used because of the tradeoffs involved in making the themes too narrow or too broad. If the themes were too specific, then they would be applicable only to a very small subset of scenarios and would not have general benefits. If they were too general, then they would be difficult to train to the point of automaticity, as they would thus require training in a wide variety of situations.

Once a participant has been presented with a TLAC scenario, he/she is asked to think about the scenario and to list items which they believe are important to consider. These items are the participant considerations. After this list has been generated, an instructor leads a discussion in which the second- and third-order effects arising from the participant considerations are discussed. Finally, the list of expert considerations is presented along with the list of participant considerations. Each participant is then given feedback related to the themes of expert cognitive performance. Participants can compare their considerations to those of the experts, and note which themes they are missing in their responses. Participants can thus focus their thinking in subsequent scenarios upon those themes in which they are performing poorly (i.e., those themes whose considerations they failed to identify).

Empirical evaluations support the TLAC training approach. Army captains who undergo TLAC training improve their ability to extract the appropriate considerations from scenarios (Shadrick & Lussier, 2004). A subsequent investigation (Shadrick, Lussier, & Fultz, 2007) found that higher TLAC scores are positively correlated with rank and deployment experience. Most tellingly, however, TLAC-trained captains outscored a sample of lieutenant colonels with deployment experience. This suggests that TLAC training accelerates the development of tactical thinking skills which otherwise might have to be gained through deployment experience and military service. Finally, Shadrick, Crabb, Lussier, and Burke (2007) found that TLAC

training resulted in cognitive performance gains on a far transfer task (writing a Company-level operational order). On the whole, the evidence suggests that TLAC training results in both the acceleration and transfer of tactical thinking skills which are important in the military context.

### *Theme-based Training for the Domain of Crisis Response*

The TLAC program demonstrates that the theme-based training approach is a successful means of ingraining expert thinking patterns. The approach is therefore relevant to crisis response, which also requires the successful assessment of high-stakes situations. However, the application of theme-based training to inter-agency crisis response is more complicated.

The goal of TLAC was not to provide Soldiers with the background *military* knowledge, but rather to train Soldiers so that tactical thinking—in this case, operationally defined as the ability to extract salient aspects of tactical scenarios which must inform appropriate decision making—becomes automatic (Lussier, Shadrick, & Prevou, 2003). However, successful crisis response training must also impart background knowledge. Plans developed by an agency in isolation may involve incorrect assumptions about the procedures, capabilities, and responsibilities of other agencies. Therefore, training which attempts to instill correct thinking patterns must be capable of correcting erroneous mental models. Second, crisis response training must also be able to instill multiple correct mental models. This is inevitable, as mental models will differ from agency to agency—the reasons for which an agency exists are fundamentally distinct from those of another agency. Finally, it is also necessary that the training provide opportunities for interaction with personnel from the various stakeholder agencies so that a broader understanding can be obtained.

To understand the expert behaviors of crisis response, a cognitive task analysis (CTA) was conducted. The CTA included experts from various agencies, including the Department of Homeland Security, the Army National Guard, Weapons of Mass Destruction Civil Support teams, and various local, county, and state police. (For a more complete description of the CTA, see Shadrick, Schaefer, & Beaubien, 2007.)

The results of the CTA were used to develop a list of behavioral themes and a set of training scenarios. The scenarios involved events which were judged likely to occur in the future, but for which the respective agencies were ill-prepared. Each scenario was dynamic and ambiguous in the initial stages. Success required extensive coordination and cooperation among multiple civil and federal agencies. As with TLAC, for each individual scenario a set of expert considerations particular to that scenario were developed by a group of subject matter experts as benchmarks of expert cognitive performance. In addition, because the responsibilities in a given scenario were different from agency to agency, there were four different versions of each scenario as follows: Army National Guard, first responder, coordinating agencies, and supporting agencies.

Titles of three of the nine scenarios used in this evaluation are listed below. Beneath each scenario title, we provide two of the approximately 20-24 expert considerations identified for the Army National Guard version of each scenario.

- Scenario: 'Power Grid Shutdown.'
  - What missions must we delay until full units are available?
  - What is the smallest unit we should allow to operate independently: company, platoon, or other?
- Scenario: 'Industrial Plant Explosion.'
  - How long can the Soldiers operate before they need their first rest period?
  - Should I split the unit to address all of the Incident Commander's needs?
- Scenario: 'Severe Earthquake along New Madrid Fault.'
  - Do I break unit integrity and organize tasks based on strength and mission requirements?
  - What impact will a second quake or major aftershock have on units in convoy?

The CTA also assessed the content validity of the scenarios (Shadrick, Schaefer, & Beaubien, 2007). Content validation procedures are used to evaluate how well important aspects of a domain are represented. Research in test content validation has found a correlation between content validity indices and criterion validities (Carrier, Dalessio, & Brown, 1990; Dubois & Dubois, 2000). The leveraging of such procedures is arguably even more important in tasks wherein the costs of inferior performance are high: that is, when improper training can lead to damaged property or lost lives. The results of the content validity assessment strongly supported the use of Red Cape.

#### Content and Use of the Red Cape Training Package

The Red Cape training package used in the current evaluation consists of 15 multimedia scenarios presented via personal computers. The scenarios are 5-10 minute presentations of audio, visual, and text-based information. Use of the Red Cape training package begins with an overview of the training methodology conducted by a trained facilitator. The facilitator gives background information on how the training content was developed, and outlines the Themes of Crisis Response, such as Maintain Focus on Mission Priorities, Keep the Chain of Command Flexible, and See the Big Picture (see Appendix A for remaining themes and theme definitions). The Themes of Crisis Response are broadly applicable across a range of scenarios, and are characteristic of the cognitions of expert crisis response personnel. They are used as a mental framework to guide the thinking of training participants.

Next, all participants in a training session don headphones and are presented with a scenario. As the participants are wearing headphones, it is possible to train collocated participants from various agencies simultaneously. At various points throughout the scenario, key facts are highlighted. At the end of each scenario, all participants receive a specific challenge appropriate to their agency: that is, they will be asked to perform a specific action or set of actions. After these challenge questions are presented, all participants are instructed to take 10-15 minutes to type in specific pieces of information, hereafter referred to as *expert considerations*, which should form the basis for any decision made or action taken. Cognitive performance is thus operationally defined as the ability to extract those salient aspects of crisis response scenarios which experts use to inform decision-making and action-taking. So that cognitive performance can be independently assessed, the system captures all participant

considerations. Because the scenarios are tailored to specific agencies, the proper considerations will also differ from agency to agency. When participants are generating the considerations, they may also click on a button to retrieve the Themes of Crisis Response.

Once time has expired, the system provides a screen displaying all of the considerations provided by the participant along with definitions of the Themes of Crisis Response. The facilitator then leads a group discussion during which members of each agency discuss key issues amongst their own members and with members of other agencies. This is useful for highlighting misconceptions about the role of other agencies in the scenario. Once the group discussion has been completed, each participant is instructed to bring up the Red Cape self-assessment screen.

The self-assessment screen displays the participant's list of considerations as well as the list of expert considerations previously developed by a group of subject matter experts. Each expert consideration is also nested within a specific Theme of Crisis Response. Training participants assess their cognitive performance by simply checking off those expert considerations which overlap with their self-generated list. By identifying which theme a particular expert consideration belongs to, participants can shore up their mental models relating expert considerations to crisis response themes. Because the expert solution for a given agency may contain a different number of considerations than the expert solution for another agency, a common performance metric must be established. This is accomplished by converting the raw score into a percentage. The percentage is the number of expert considerations identified/checked by the participant divided by the total number of expert considerations available.

## Army National Guard Workshops

### *Method*

#### *Participants*

A total of 40 individuals participated in the series of workshops. The audience was drawn from individuals of varying organizations (Army National Guard  $n = 22$ , coordinating agencies  $n = 14$ , first responder agencies  $n = 4$ ). All of the data reported herein were collected during pre-scheduled, formal training events.

#### *Procedure*

There were three rounds of workshops in all, conducted at two-week intervals. Within each round of workshops, there were a total of three sessions. In each session, participants viewed and responded to three scenarios (see Appendix B). Training participants were allowed to participate in only one session per round, as the same scenarios were used for all sessions within a given round of workshops. Presentation order of the scenarios was randomly varied from session to session. The sessions were conducted in the manner outlined in the previous section titled "Content and Use of the Red Cape Training Package." At the end of each session, participant feedback was gathered.

## Measures

### *Quantitative Data*

*Rater-assigned Scores.* To ensure that the rating assessments were reliable, the participant considerations of a randomly chosen subset of participants ( $n = 9$ ) were evaluated by two independent raters. After assessing the participant considerations, the two raters met to discuss and calibrate their rating procedures. The agreement between their pre-calibration ratings was assessed via the inter-class correlation coefficient (ICC). In general, ICC values ranging from .45 to .75 are indicative of good agreement, and values greater than .75 indicate excellent agreement (Fleiss, 1981). The observed ICC value of .80 ( $p < .05$ ) obtained on the pre-calibrated ratings thus denoted excellent agreement among the raters. The .80 value reported here is consistent with values obtained in previous theme-based TLAC research (e.g., .87, Shadrick, Lussier, & Fultz, 2007; .90, Shadrick & Lussier, 2004). After calibration, the ICC value approached 1.00. Having thus assessed the rating procedure used the remaining participant considerations were evaluated by one of the independent raters.

*Self-assessment Scores.* Self-assessment scores were examined for improvement across workshop sessions. However, another important interest in the self-assessment scores was their correspondence with the rater-assigned scores.

### *Participant Feedback*

Participant feedback consisted of Likert-scale items (1 = strongly disagree, 5 = strongly agree) and open-ended questions. The Likert-scale items asked participants to indicate their agreement with the statements:

- “I am confident that I can perform the trained skills when I return to the field.”
- “The training exercises were challenging.”
- “The training met my expectations.”
- “The training was relevant to my position within my organization.”

The open-ended questions asked participants:

- “What did you like best about Red Cape training?”
- “What did you like least about Red Cape training?”
- “How can we improve Red Cape training?”

## Analyses

### *Quantitative Data*

*Rater-assigned Scores.* Ideally, all participants would have participated in all three workshops and all nine scenarios. However, this was not possible due to the participants' primary mission requirements. For example, one workshop session had to be rescheduled because Army National Guard participants were activated to respond to emergency weather

conditions. While missing data complicates the data analysis, data coding adjustments were made to aid interpretation. We coded the data according to whether or not a given workshop session was the first, second, or third Red Cape training event for an individual. To assess the cognitive performance gains, we graph the overall mean rater-assigned score across participants per workshop. We also correlate the mean rater-assigned scores per participant with the number of scenarios completed by that participant. As the number of scenarios was a non-normally distributed variable, we assess all its relationships via the nonparametric Spearman-rho  $r$ .

*Self-assessment Scores.* The procedure for examining the self-assessment scores was similar to that used for the rater-assigned scores. First, we graph the mean self-assessment score across participants per workshop. As with the rater-assigned scores, we also correlate the mean self-assessment score for an individual with the number of scenarios completed by that individual.

To assess the extent to which the self-assessment scores corresponded to the rater-assigned scores, within each workshop session we correlated each individual's mean self-assessment score with their rater assigned score. The difference between the two sets of scores can be considered a bias indicator. A positive difference score (where self-assessment scores are higher than rater-assigned scores) would arguably indicate overconfidence; conversely, a negative difference score would arguably indicate underconfidence. We focus on this potential bias for two reasons. First, earlier theme-based training research has indicated that participants have a tendency to inflate their self-assessment scores (Shadrick & Lussier, 2004). Second, we wanted to see if the accuracy of self-assessment would increase over the course of Red Cape training.

It can be argued that rater-assigned scores are as subject to bias as self-assessment scores. However, we believe this to be rather implausible. First, the rater-assigned scores were examined for reliability via the calculation of an inter-class correlation coefficient. While reliability does not indicate validity, it is a necessary component of it, and judgment outcomes substantially the same across two independent raters is consistent with valid ratings. Second, the raters made their assessments of cognitive performance without time constraints. Therefore, there is good reason to believe that the rater assessments of cognitive performance were more likely to be accurate than the self-assessments of cognitive performance. Third, overconfidence in self-assessment is a robust phenomenon that appears in many domains (Schaefer, Williams, & Goodie, 2004; Nowell & Alston, 2007; Dunning, Heath, & Suls, 2004).

### *Participant Feedback*

Mean Likert-scale responses are graphed as a function of workshop. Representative comments to each of the open-ended questions are also provided.

## Results

### Quantitative Data

*Rater-assigned Scores.* As Figure 1 indicates, rater-assigned scores increased across workshops. This indicates that, as training progressed, the participants' cognitive performance improved. This finding is supported by the significant correlation (Spearman-rho  $r(40) = +.74$ ,  $p < .01$ ) found between the overall mean rater-assigned score for an individual and the number of scenarios completed by that individual.

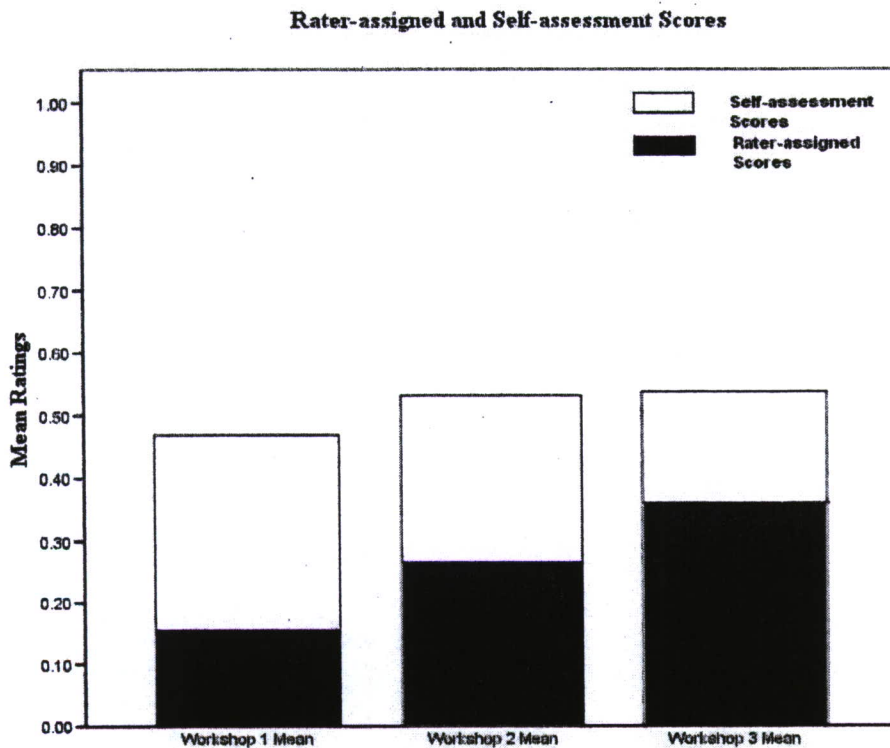


Figure 1. Ratings of cognitive performance as a function of workshop.

*Self-assessment Scores.* The self-assessment scores also improved as a function of training, as indicated in Figure 1. As with the rater-assigned scores, self-assessment scores were positively and significantly correlated with number of completed scenarios (Spearman-rho  $r(40) = +.80$ ,  $p < .01$ ). The overall correlation between the average self-assessment scores and the average rater-assigned scores was also significant (Pearson  $r(40) = +.77$ ,  $p < .01$ ).

It is also evident in Figure 1 that there is a notable disparity between rater-assigned and self-assessment scores, which we have interpreted as overconfidence. The figure also indicates that this presumed overconfidence/inflation decreased as training progressed. In addition, as training proceeded, the correlations between rater-assigned scores and self-assessment scores increased (+.39 in Workshop 1, +.52 in Workshop 2, and +.72 in Workshop 3). The data suggest that self-assessment of cognitive performance becomes more accurate—relative to rater-assigned scores—as Red Cape training proceeds.

### Participant Feedback

Participants were asked to rate their agreement with the statement “I am confident that I can perform the trained skills when I return to the field.” The responses to this item are graphed in Figure 2. Overall, many participants appeared confident about skill transfer. However, significantly lower confidence ratings were seen in Workshop 2 ( $F(2, 12) = 4.5, p < .05$ ). A closer examination of the data indicated that the ‘Severe Earthquake along New Madrid Fault’ scenario was primarily responsible for the drop in expressed confidence. Both rater-assigned and self-assessment scores were lowest for that scenario. It is plausible, therefore, that the confidence ratings were lower as a result of their performance in that particular scenario.

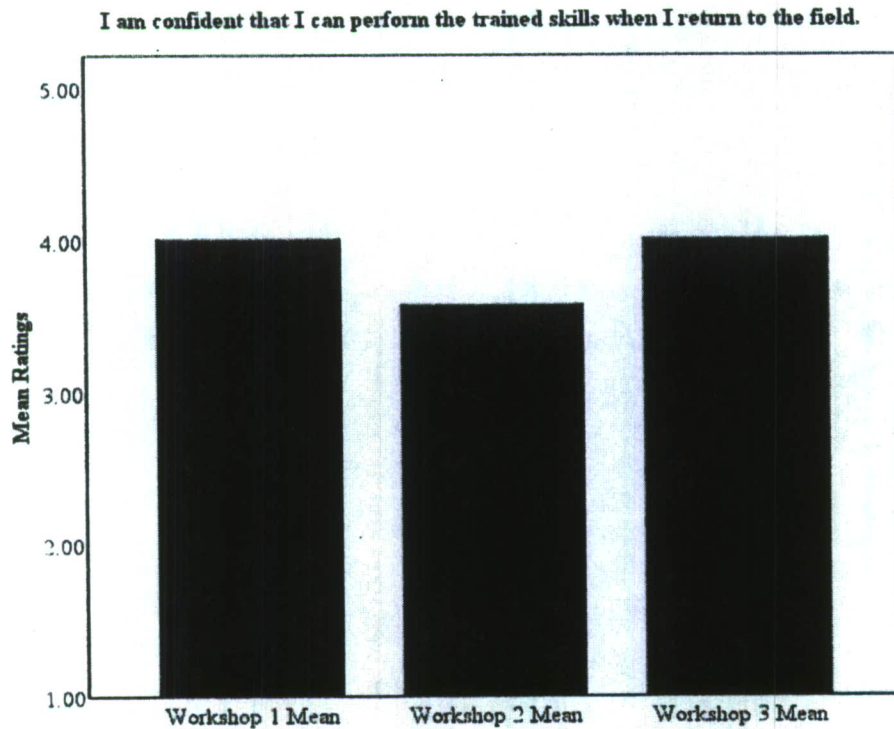


Figure 2. Ratings for confidence in trained skills question.

Anecdotally, there is also the observed behavior of participants during the workshops. When initially presented with the ‘Severe Earthquake along New Madrid Fault’ scenario, a common reaction was stunned silence. One of the participants stated “This is the situation which we have feared, and for which we need to train.” The group discussion after this scenario in particular was both lively and interesting to observe. As the stakes of failure in such a scenario became clearer, interagency conversation became more animated and in-depth. There were several instances in which participants from other agencies saw the need to radically overhaul their own plans because their assumptions involving other agencies turned out to be incorrect.

Participants were also asked to indicate their agreement with the statement “The training exercises were challenging.” (See Figure 3.) As training progressed, participants began to see

the material as more challenging ( $F(2, 12) = 4.96, p < .05$ ). Similarly, responses to the “What did you like least about the Red Cape training?” question dealt with the difficulty of the material. One participant wrote “The difficulty of organizing a qualified response to a situation with little or the barest information available in a short time...It was real brain work!” Another simply wrote “Time constraints.”

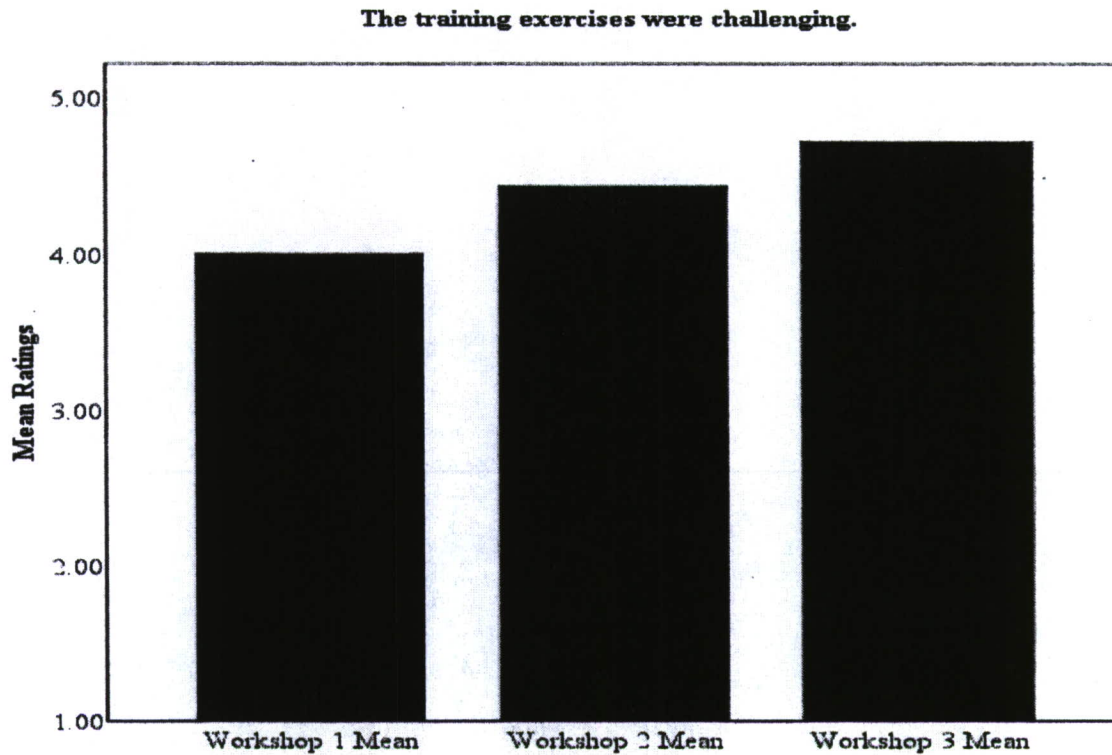


Figure 3. Ratings for challenge level of training exercise question.

Responses to the Likert-scale item “The training met my expectations” (Figure 4) indicate that participant perceptions of the training became more positive as the training progressed ( $F(2, 12) = 4.8, p < .05$ ). A similar but non-significant trend ( $F(2, 12) = .75, p > .05$ ) was seen in the responses to the Likert-scale item “The training was relevant to my position within the organization” (Figure 5).

**The training met my expectations.**

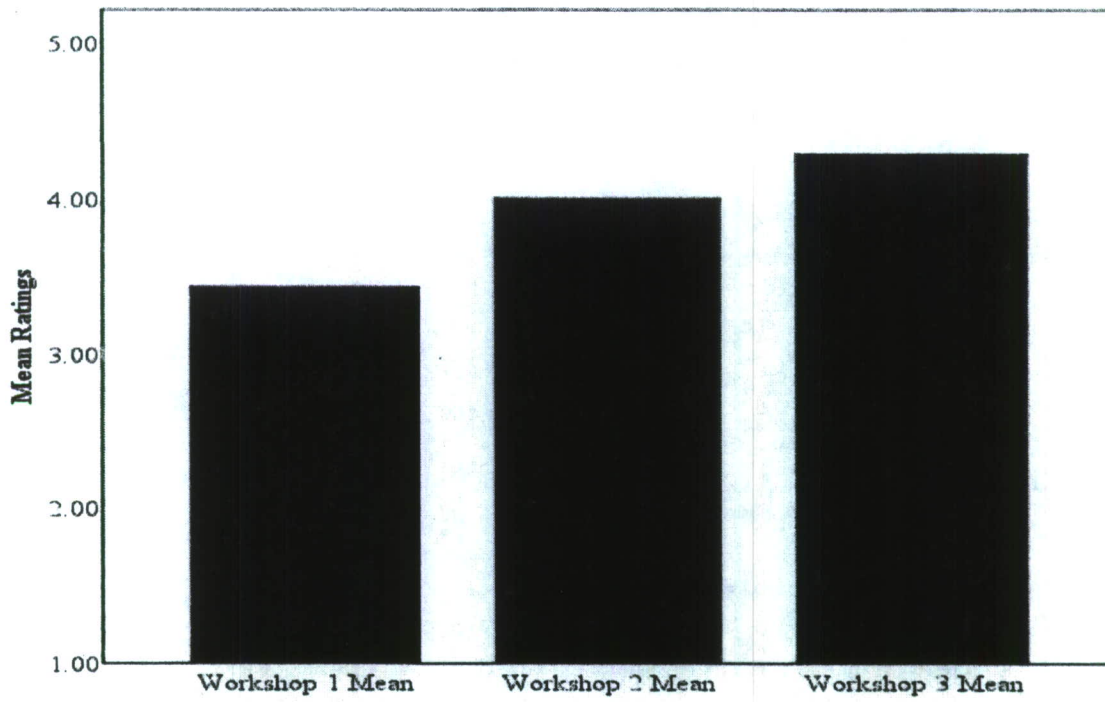


Figure 4. Ratings for training expectation question.

**The training was relevant to my position within my organization.**

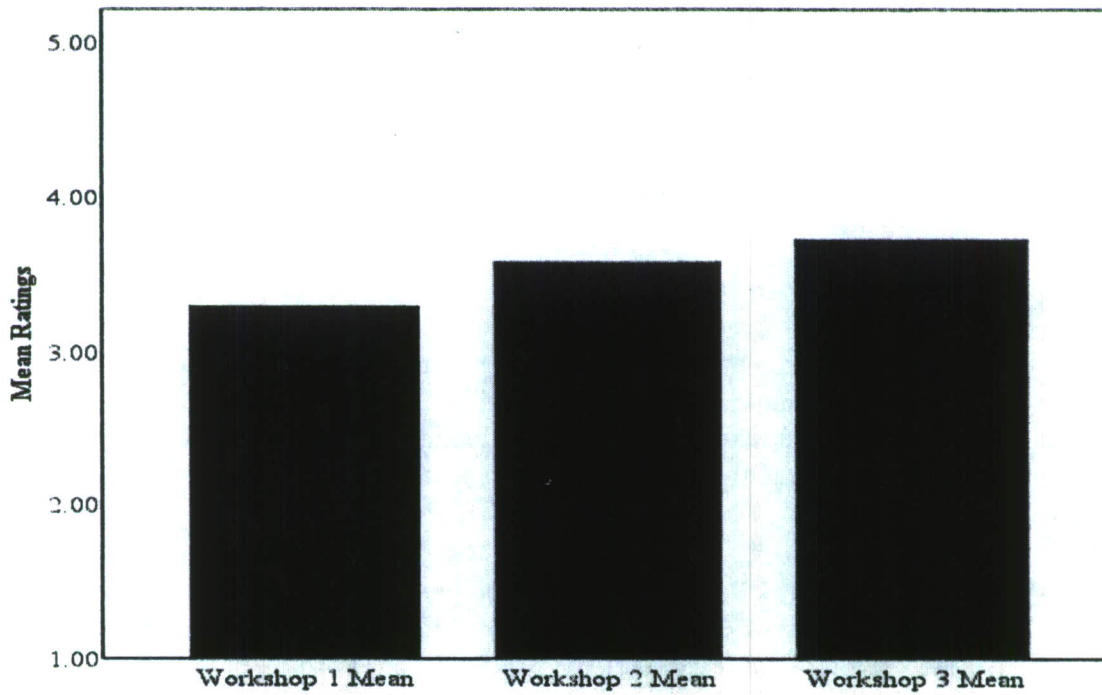


Figure 5. Ratings for position relevance question.

Participants were also asked “How can we improve Red Cape training?” One telling response was “I think more information could be given in the scenarios... however, I understand that we often only get fragmented information prior to a mission...so I guess it is somewhat realistic.” Another participant, while talking about the difficulty imposed by the time constraints, said “...but this really happens.” Participant feedback is thus consistent with the position that Red Cape training is difficult because crisis response is difficult.

When participants were asked “What did you like best about Red Cape training?” the responses related to two aspects of the training: the multimedia nature of the presentations and the group discussions. For example, comments related to the delivery method included:

- “It wasn’t death by Power Point...liked the exercises, very interactive, makes you think.”
- “Interactive media...driving home priorities.”
- “I like the fact that it is computer based and requires 'hands on' with a lot of individual input...kept my attention.”

Comments related to the group discussions included:

- “Discussion from different agencies to facilitate understanding of the big picture.”
- “The thought of interfacing with different disciplines regarding plausible scenarios...Format provokes looking at issues from other perspectives.”
- “The discussion...encouraged all to think outside the box.”

#### Discussion

Results support the efficacy of the Red Cape training program. Learning—as reflected in rater-assigned scores—increased as a function of training. This was demonstrated by the significant, positive correlation between average rater-assigned scores and number of completed scenarios. These results are encouraging when one considers that (a) the time between training sessions was approximately two weeks and (b) the total training period for three workshop sessions was roughly equivalent to 1.5 days of training.

Participant feedback was also favorable. Participants perceived the training as having met their expectations and as relevant to their individual positions within their respective agencies. Even the responses to the “What did you like least about Red Cape training?” question were largely complimentary. Consider, for example, some of the participant responses quoted above. Although several noted the difficulty of incomplete information and time constraints, they also tended to note that these traits were characteristic of real life crisis response situations.

The pattern of positive participant feedback was replicated in a separate effort when ARI personnel conducted a one-day Red Cape workshop at the request of an Air National Guard unit and relevant stakeholder agencies located in a different state. When asked “What did you like least about Red Cape training?” respondents answered “Just making the scenarios Air Force specific” and “Make it region/service specific, although I understand why it wasn’t.” In other

words, the Air National Guard participants saw the relevance of the training, and wanted to tailor it to their own needs. The U.S. Army Research Institute had already anticipated such a response, and has embarked upon a research effort focusing upon the development of a user-friendly modification tool to meet such requests.

However, while training improved cognitive performance, the Army National Guard data reported here also indicate the necessity for further training in the domain of crisis response. Even in Workshop 3, the average cognitive performance of participants was around 40 percent. In other words, participants were able to extract only 40 percent of the expert considerations from the scenarios. Similar theme-based TLAC training enabled U.S. Army captains to provide 70 percent of the appropriate expert tactical considerations (Shadrick & Lussier, 2004). In part, this discrepancy might be due to the interagency nature of crisis response. When the success of a given agency's plan is so contingent upon the performance of other agencies, it is difficult to quickly and accurately formulate an appropriate solution.

Perhaps the most intriguing finding, however, relates to self-assessment. As training progressed, participants began to more accurately self-assess cognitive performance. This is a potentially important finding for two reasons. First, crisis response personnel including Soldiers must be able to accurately self-evaluate and—if necessary—change plans as a situation develops. Secondly, this finding suggests a means of further validating the theme-based training approach used in Red Cape.

Consider for a moment the approach used in Shadrick, Lussier, and Fultz (2007). In that research, the cognitive performance of individuals who varied along several dimensions (i.e., military rank, deployment experience, and TLAC training) was compared. In general, persons of higher rank outperformed those of lower rank, and the deployed outperformed the non-deployed. This pattern suggests that TLAC trains an externally valid skill—namely, tactical thinking. More importantly, however, it was also found that TLAC-trained captains outperformed captains with deployment experience for periods of up to one year. This further suggests that tactical thinking skills gained via combat experience and military service may be, at least partially, teachable in an environment where poor performance does not incur such a heavy cost.

We are suggesting that a similar approach be used in the domain of crisis response. This approach would involve a two-stage process. First, we would choose individuals ranging in amount of crisis response experience. Those individuals would then undergo Red Cape training and then have their cognitive performance and accuracy of self-assessment measured. Second, we would further examine the findings here: namely, that inflation in self-assessment is reduced by undergoing Red Cape training. If more experienced individuals in crisis response efforts are more accurate in self-assessment—and training a new group of participants increases the accuracy of self-assessment—we can plausibly assert that Red Cape training instills important skills which would otherwise have to be gained through a time- and resource-intensive process. In addition, of course, we would be able to train individuals prior to participation in high-stakes crisis response situations and thereby mitigate the sometimes disastrous consequences which can arise from poor decision-making.

## Conclusions

The Red Cape training package appears to deliver on its promise of low-cost, efficient, and effective training. Participant feedback was favorable, and indicated that although the training was challenging, it was also needed. The rater-assigned scores bear out the subjective impressions of difficulty and indicate the need for further crisis response training. Current research is being conducted to determine how Red Cape training materials can be easily and rapidly modified by non-technical personnel so this training method can be adapted to other domains. Further research investigating the reliability and utility of the latter finding is suggested.

In sum, it appears that Red Cape theme-based training captures the difficulty of crisis response situations, and improves responders' cognitive performance and their self-assessments of cognitive performance. While the importance of improving crisis response performance is self-evident, the implications of improved self-assessment or self-awareness are also important. Training events are likely to remain rare, and experienced trainers are a precious commodity. As the U.S. Army has noted (Department of the Army, 2006), persons who accurately reflect upon and improve their own performance are an invaluable resource, especially given the dynamic nature of the current operating environment. Red Cape's ability to improve self-awareness may prove equally important for crisis response.

## References

- Alburt, L. (2003). *Secrets of the Russian chess masters: Fundamentals of the game, Volume 1*. W. W. Norton & Company.
- Carrier, M. R., Dalessio, A. T., & Brown, S. H. (1990). Correspondence between estimates of content and criterion-related validity values. *Personnel Psychology*, 43 (1), 85-100.
- Department of the Army. (2006). *Army Leadership: Competent, Confident, and Agile (FM 6-22)*. Washington, DC: Author.
- Dixon, R. G. (2006). *Systems thinking for integrated operations: Introducing a systematic approach to operational art for disaster relief*. Fort Leavenworth, KS: School of Advanced Military Studies (DTIC No. ADA448434).
- Dubois, D. A., & Dubois, C. L. Z. (2000). An alternate method for content-oriented test construction: An empirical evaluation. *Journal of Business and Psychology*, 15 (2), 197-213.
- Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment. *Psychological Science in the Public Interest*, 5 (3), 69-106.
- Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100 (3), 363-406.
- Farmer, L. C. & Williams, G. R. (2005, October). *The rigorous application of deliberate practice methods in skills courses*. Paper presented at the UCLA/IALS Sixth International Clinical Conference on Enriching Clinical Education, Lake Arrowhead, CA.
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions*. New York: Wiley.
- Glenn, R. W., Jacobs, J., Nichiporuk, B., Paul, C., Raymond, B., Steeb, R., & Thie, H. J. (2006). *Preparing for the proven inevitable: An urban operations training strategy for America's joint force*. Santa Monica, CA: RAND Corporation.
- Guest, C. B., Regehr, G., & Tiberius, R. G. (2001). The life-long challenge of expertise. *Medical Education*, 35 (1), 78-81.
- Helton, W. S. (2007). Deliberate practice in dogs: A canine model of expertise. *Journal of General Psychology*, 134 (2), 247-257.
- Hodges, N. J., Kerr, T., Starks, J. L., Weir, P. L., & Nananidou, A. (2004). Predicting performance times from deliberate practice hours for triathletes and swimmers: What, when, and where is practice important? *Journal of Experimental Psychology: Applied*, 10 (4), 219-237.

- Krampe, R. T., & Ericsson, K. A. (1996). Maintaining excellence: Deliberate practice and elite performance in young and older pianists. *Journal of Experimental Psychology: General*, 125 (4), 331-359.
- Lussier, J. W. (1998, May). *Developing expertise in battlefield thinking: Models of battlefield thinking*. (Unpublished briefing). Fort Leavenworth, KS: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Lussier, J. W., Shadrick, S. B., & Prevou, M. I. (2003). *Think Like a Commander prototype: Instructor's guide to adaptive thinking*. (Research Product 2003-02). Arlington, VA: U. S. Army Research Institute for the Behavioral and Social Sciences.
- Nowell, C., & Alston, R. M. (2007). I thought I got an A! Overconfidence across the economics curriculum. *Journal of Economic Education*, 38 (2), 131-142.
- Phillips, J. K., Shafer, J., Ross, K. G., Cox, D. A., & Shadrick, S. B. (2005). *Behaviorally anchored rating scales for the assessment of tactical thinking mental models*. Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Ross, K. G., & Lussier, J. W. (2000). *Adaptive thinking seminar*. (Available from the ARI-Ft. Knox Research Unit, U.S. Army Research Institute, 121 Morande St., Fort Knox, KY 40121-4141).
- Schaefer, P. S., Williams, C. C., & Goodie, A. S. (2004). Overconfidence and the Big Five. *Journal of Research in Personality*, 38 (5), 473-480.
- Shadrick, S. B., & Lussier, J. W. (2004). *Assessment of the Think Like A Commander training program*. (Research Report 1824). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Shadrick, S. B., Crabb, B. T., Lussier, J. W., & Burke, T. J. (2007). *Positive transfer of adaptive battlefield thinking skills*. (Research Report 1873). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Shadrick, S. B., Lussier, J. W., & Fultz, C. (2007). *Accelerating the development of adaptive performance: Validating the Think Like a Commander training*. (Research Report 1868). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Shadrick, S. B., Schaefer, P. S., & Beaubien, J. (2007). *Development and content validation of crisis response training package Red Cape: Crisis action planning and execution*. (Research Report 1875). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

## Appendix A

### Themes of Crisis Response

#### *Maintain Focus on Mission Priorities*

Effective managers recognize that their primary mission is to protect human life and property, with safety being paramount. Plans and actions should focus on preserving life and property while also addressing the incident and its underlying causes. Effective managers may choose to delay, deflect, or delegate activities that address secondary or tertiary objectives.

#### *Keep the Chain of Command Flexible*

Effective managers recognize that the chain of command does not remain fixed throughout the crisis. Management roles often evolve as the crisis unfolds and additional agencies become involved. Effective managers recognize and transition between manager and subordinate (or participant and observer) as the crisis unfolds.

#### *See the Big Picture*

Effective managers remain aware of what is happening around them, with particular attention of the impact on higher, adjacent, lower, and supported organizations. Effective managers attempt to analyze all factors inside and outside the organization's direct control that could affect its activities during the time line covered by the plan, to include geographical scope, population involvement, incident duration, and lead time required for particular actions.

#### *Plan for and Recognize Decision (Trigger) Points*

Effective managers identify key and measurable indicators that require immediate action. They develop plans of action for these trigger points to avoid premature commitment of key assets and failing to act in a timely manner.

#### *Reprioritize as Necessary*

Effective managers continually reassess the scope and priority of mission requirements as the crisis unfolds. They continually validate new information as it becomes available, and adjust their mission plans as the crisis unfolds and approaches resolution. Effective managers are able to determine "time-critical" versus "resource-critical" requirements and adjust the plan accordingly.

#### *Use All Available Assets*

Effective managers remain aware of what assets are available to them, what their limitations are, and what lead times are required. They listen to the input of subject matter experts and are prepared to defer judgment to them, as necessary. Effective managers use persuasion and cooperation in situations where directive authority is not appropriate.

*Think in Shades of Gray, not Black and White*

Effective managers recognize that they will be working with imperfect and incomplete information from a wide range of organizations and individuals. They consider the source of the data and allow for variations when applying that data.

*Model a Dynamic Situation*

Effective managers recognize that the cause of the incident may be an accident or act of nature, or it may be either criminal or terrorist activity. They avoid the temptation to simplify the situation by treating the cause of the crisis as static or simply reactive.

*Understand the Public Need*

Effective managers recognize that the public will experience a wide range of emotions, from proactive attempts at involvement to a passive acceptance of the situation. Certain reactions could lead to irrational behavior from the public. The effective manager takes the time to assess these reactions, explain (when appropriate) the reasoning behind certain actions and attempts to allay any public fears.

## Appendix B

### Army National Guard Workshop Scenarios

#### Workshop 1 Scenarios

##### *Nuclear Bomb in Shipping Container*

Intelligence reports indicate that Al Qaeda has obtained a five kiloton nuclear weapon and the associated activation codes. That size bomb would decimate a major city the size of New York or Chicago. During the past three weeks, an increase in the volume of cell phone and e-mail intercepts has led the Department of Homeland Security to believe that plans are being made to deploy the bomb. Six days ago, all communications traffic suddenly ceased. The Department of Homeland Security believes that Al Qaeda has gone into "radio silence" and that the bomb may be in transit. The Burns Harbor is identified as a likely point of entry for the weapon.

##### *Dirty Bomb*

An automobile packed with explosives exploded in downtown Indianapolis. The explosion killed a police officer and wounded 12 citizens. As the initial explosion aftershock subsided, a "talc" like powder was seen covering the blast zone. Area citizens rushed to the scene and attempted to render assistance. Local merchants opened their establishments for any relief effort. Emergency medical crews were transporting the injured to the local hospitals. A terrorist group later claimed responsibility for the attack.

##### *Storm of the Century*

On Thanksgiving morning, the state of Indiana was hit with a major storm event. The lake area received 24 inches of snow, in an 18 hour period. The middle section of the state was hit by the same storm system dumping 3.5 inches of ice on the area. Due to the previous weeks' heavy rain, the Ohio River area was showing signs of potential flooding. The flooding conditions are only exacerbated by the current storm. The Army Corps of Engineers predicts for the river to crest in 18 hours around the counties of Clark, Harrison and Spencer. The storm results in numerous problems that hamper health care, first response, and public works organizations.

#### Workshop 2 Scenarios

##### *Power Grid Shutdown*

Wabash Valley Power Company supplies electricity to 27 distribution cooperatives throughout northern and central Indiana. The company recently experienced a shutdown at one of those cooperatives. A power grid failure resulted, thereby shutting down electrical power to two-thirds of the state. The shutdown occurs during a severe major winter storm, leaving many residents, at risk of hypothermia.

### *Industrial Plant Explosion*

A major explosion occurred this morning at a chemical plant in Gary, Indiana. The plant is located adjacent to Interstate 80/90, adjacent to the Grand Calumet River, one quarter mile south of Lake Michigan, one mile north of the Methodist Hospital Northlake Campus, and 13 miles east of the Illinois border. Electrical power, telephone and cable television interruptions are being reported throughout the Gary metropolitan area.

### *Severe Earthquake along New Madrid Fault*

A magnitude 7.5 earthquake has occurred along the New Madrid fault, impacting a multi-state region that includes Arkansas, Illinois, Indiana, Kentucky, Mississippi, Missouri, and Tennessee. Reports from Evansville, Clarksville, and Bloomington indicate catastrophic damage in all three cities. Buildings have collapsed and debris is blocking city streets. Numerous fires are reported burning in each city, and water main breaks have significantly hindered firefighting operations. Casualty and death figures are sketchy but growing rapidly. Hospitals are functioning on generator power and are overwhelmed with victims. Electricity and water are out in major portions of each city.

### Workshop 3 Scenarios

#### *Airplane Crash in Restricted Area*

Forty-five minutes ago, a 12-passenger corporate jet crashed just short of the U.S. Army Newport Chemical Weapons Depot. The jet's flight plan indicated that it was scheduled to fly from Kansas City, Missouri to Cincinnati, Ohio. Twelve minutes prior to the crash, the jet veered 30 degrees off course and began descending from 30,000 feet. No emergency was declared and the FAA was unable to contact the pilot over any radio frequency. Because the crash crossed the Depot fence line, multiple jurisdictions are involved.

#### *Vehicle with Hazardous Material (HAZMAT)*

The Indiana State Police receive a 911 call from a motorist with a cell phone. A commercial truck with a tandem semi-trailer has overturned in the eastbound lane of I-64 between mileposts 32 and 33. The rear tanker trailer is on its side, blocking both lanes at the bridge over Pigeon Creek and leaking heavily from the top hatch. The tractor and front tanker trailer left the highway and went down the embankment, coming to a rest on its side in Pigeon Creek. The caller can see a hazardous cargo placard on the overturned trailer, but does not want to venture closer to the vehicle to determine what it was carrying.

*Sports Riot in a University Town*

In late August, Indiana University is playing host to Notre Dame in an early season football game. Two weeks prior, ESPN contacted the university and announced that they were going to make this a nationally-televised football game. An unrelated three day outdoor rock concert had been previously scheduled outside the city limits, with performances scheduled for 15 hours per day on Thursday, Friday, and Saturday. Small conflicts between the football and music fans, eventually reach a boiling point.



## Appendix C

### Air National Guard Workshop Scenarios

#### *Power Grid Shutdown*

Wabash Valley Power Company supplies electricity to 27 distribution cooperatives throughout northern and central Indiana. The company recently experienced a shutdown at one of those cooperatives. A power grid failure resulted, thereby shutting down electrical power to two-thirds of the state. The shutdown occurs during a severe major winter storm, leaving many residents at risk of hypothermia.

#### *Airplane Crash in Restricted Area*

Forty-five minutes ago, a 12-passenger corporate jet crashed just short of the U.S. Army Newport Chemical Weapons Depot. The jet's flight plan indicated that it was scheduled to fly from Kansas City, Missouri to Cincinnati, Ohio. Twelve minutes prior to the crash, the jet veered 30 degrees off course and began descending from 30,000 feet. No emergency was declared and the FAA was unable to contact the pilot over any radio frequency. Because the crash crossed the Depot fence line, multiple jurisdictions are involved.

#### *Rail Yard Explosion*

The Indiana Army National Guard is simultaneously deploying two battalions to the Middle East. As the units progress through their qualification and pre-deployment procedures, both are required to load their vehicles onto rail cars for transport to the port in Charleston, South Carolina. An explosion occurs during the loading procedure, killing several Guard Soldiers and damaging multiple vehicles.