REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
21-11-2014	Final	25092013 - 25092014

4. TITLE AND SUBTITLE
Small Unit Decision Making (SUDM)Assessment Battery Final Report: Option III

5a. CONTRACT NUMBER
N00014-12-C-0459

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Ross, Karol, G., PhD
Phillips, Jennifer, K., PhD
Rivera, Iris, D., PhD
Rosopa, Patrick, J., PhD

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Cognitive Performance Group of Florida, LLC.
3662 Avalon Park East Blvd., Suite 205
Orlando, FL 32828

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
875 North Randolph St.
Arlington, VA 22203-1995

10. SPONSOR/MONITOR'S ACRONYM(S)
ONR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
A002

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The objective of this project was to develop a Small Unit Decision Making (SUDM) Assessment Battery to measure decision-making proficiency of infantry small unit leaders. The target audiences are maneuver squad leaders and platoon commanders. The battery contains off-the shelf self-report and custom performance instruments. This report summarizes the finalization of the battery. Item analysis supports the quality of the instruments. Each off-the-shelf instrument was found to measure a discrete construct. The instruments for resilience, metacognition, and attentional control, and the SUDM Situational Judgment Test (consisting of subscales for sensemaking, situational assessment, and decision making), were the strongest predictors of decision making performance as measured by the Decision Requirements Interview, a performance instrument created for this project. The battery significantly discriminates between differently performing groups. The seven instruments addressing resilience, problem solving, attentional control, ambiguity tolerance, self-regulation, and self-awareness, and the SUDM Situational Judgment Test, made the largest contribution towards differentiating between the performance level of the NCO and Lt groups participating in the project. Two additional products from this effort are the Maneuver Squad Leader Mastery Model and the Small Unit Decision Making Behaviorally Anchored Rating Scale. To support implementation by the Marine Corps, the battery should be converted into a computer-administered version, followed by transition to the Marine Corps including system demonstration, testing in an operational environment, and training for battery implementation.

15. SUBJECT TERMS
small unit decision making, decision making assessment, assessment battery, psychometric analysis, mastery model, behaviorally anchored rating scale

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					Karol G. Ross, PhD Chief Scientist
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code)*
U	U	U	UU	53	407.282.4433

Small Unit Decision Making (SUDM) Assessment Battery Final Report: Option III

Prepared by:

Karol G. Ross, Jennifer K. Phillips, and Iris D. Rivera

Cognitive Performance Group
3662 Avalon Park East Blvd., Suite 205
Orlando, FL 32828

Patrick J. Rosopa
Department of Psychology
College of Business and Behavioral Science
Clemson University
418 Brackett Hall
Clemson, SC 29634

Prepared for:

Office of Naval Research
875 North Randolph St.
Arlington VA 22203-1995

Contract Number: N00014-12-C-0459
CDRL A002
21 Nov 2014

Approved for public release; distribution is unlimited

Abstract

The objective of this project was to develop a Small Unit Decision Making (SUDM) Assessment Battery to measure the decision-making proficiency of infantry small unit leaders over time. The purpose of this report is to summarize the finalization of the SUDM Assessment Battery and project. During the finalization phase, a psychometric analysis was conducted. Item analysis and factor analysis supported the removal of a number of items and one instrument that shortened the battery. All instruments were found to measure one discrete construct. The instruments for the constructs of resilience, metacognition, and attentional control, and the SUDM Situational Judgment Test (including subscales for the constructs of sensemaking, situational assessment, and decision making), were found to be the strongest predictors of decision making performance as measured by the Decision Requirements Interview, an instrument created for this project. The battery was found to have the ability to discriminate between differently performing groups. We established that the noncommissioned officer (NCO) and Lieutenant (Lt) groups in the study were statistically different in experience and scores on the Decision Requirements Interview. The scores on the seven instruments addressing resilience, problem solving, attentional control, ambiguity tolerance, self-regulation, and self-awareness, and the SUDM Situational Judgment Test, made the largest contribution towards differentiating between the performance level of NCOs and Lts. Additionally, two other products generated during this effort are described and their application discussed. They are the Maneuver Squad Leader Mastery Model and the Small Unit Decision Making Behaviorally Anchored Rating Scale. Aspects of the battery have merit for continued research and development to make the battery as robust and as useful to the Marine Corps as possible in order to help ensure the best possible success for small unit leaders in the future. Recommendations include investigation of four constructs initially considered as important aspects of small unit decision making (adaptability, cognitive flexibility, change detection, and anomaly detection) for which acceptable instruments were not found or developed. We also recommend that the final battery be implemented with a relatively large sample size from the desired population—prospective and current maneuver squad leaders and platoon commanders—to establish a robust reference group for scoring and interpretation. Finally, to support implementation by the Marine Corps, the battery should be converted into a computer-administered version to mitigate test fatigue and cognitive overload by allowing the respondents to save their work, stop, and return when refreshed. The computer version should also automate scoring and interpretation. This migration to a computer-supported version should be followed by steps to transition that product to the Marine Corps including system demonstration, testing in an operational environment, and training for battery implementation.

Acknowledgements

Cognitive Performance Group would like to thank The Basic School of the US Marine Corps for the access given to our research team to collect data from the Basic Officer Course. Without their support, this project could not have been completed. Specifically, we would like to thank Mr. John DeForest, Academics Director of The Basic School, and Col Todd Desgrosseilliers, the Commanding Officer of The Basic School at the time of this research, for their continued support and interest in this project.

Contents

Tables

Figures

Executive Summary

The Problem

The *Marine Corps Vision and Strategy (MCV&S) 2025* calls for the Marine Corps to be the nation's expeditionary force of choice and to demonstrate the ability to rapidly deploy to a wide range of complex and irregular operating environments as lean, agile, and adaptable individuals and units. This vision can be supported not only by changes to training, education, and experiences for small unit leaders, but also by creating better options to assess decision-making proficiency as a means of assessing the status of and improvements over time in *cognitive readiness*[1] across the Force. The decision dilemmas faced by squad leaders are too numerous to count, let alone test as individual performance items. Furthermore, as is the case in cognitively complex performance environments, seldom if ever can a single best decision be identified for a given tactical problem. Prior assessment efforts have overcome these challenges by scoping the assessment space to a specific, well-defined set of performance parameters, or by relying on subject-matter expert (SME) ratings of decision quality as a means of quantifying decision performance. The multidimensionality of decision making is lost in the assessment process. These approaches also do not lend themselves to the Marine Corps' requirement for a scalable, generalizable assessment capability that predicts decision performance across a range of operational settings. Therefore, the Small Unit Decision Making (SUDM) Assessment Battery research project was undertaken to fill that gap.

Method

The SUDM Assessment Battery project consisted of four phases—*preparation, development, testing, and finalization*—to achieve a reliable and valid battery sufficient for understanding small unit tactical decision making. Preparation consisted of instrument selection and development and the generation of the Maneuver Squad Leader Mastery Model. During the *development phase* instruments were examined and refined with a pilot group to develop a battery suitable for testing with a larger audience. The purpose of the *testing phase* was to make final adjustments to the instruments, administration, and scoring protocols, as needed, using a larger sample than the pilot group.

To support the testing phase, a convenience sample was identified at The Basic School (TBS) in Quantico, VA. Data were collected by administering the battery at TBS to the Basic Officer Course (BOC) companies completing the six-month course, both before and after the course. Participants consisted of a sample of noncommissioned officers (NCOs) in two FY13 BOC companies followed by Lieutenants (Lts) provided by TBS for each course beginning in FY14 and all the NCOs who were participating in each FY14 course to improve their ability to perform

[1] "*Cognitive readiness* is the mental preparation (including skills, knowledge, abilities, motivations, and personal dispositions) an individual needs to establish and sustain competent performance in the complex and unpredictable environment of modern military operations" (Morrison and Fletcher, 2002, p. I-3).

as TBS instructors as part of the Enlisted Instructor-Advisor Initiative (Desgrosseilliers & Hoffman, 2014).

In exchange for the opportunity to collect a large sample of data, our research team will be offering insights into the impact of the Enlisted Instructor-Advisor Initiative under a separate project. The SUDM Assessment Battery project will use the portion of the data collected from FY13 and FY14 BOC companies between July 2013 and July 2014 to support completion of the battery. Data collection will continue at TBS through March 2015 to complete the impact analysis study under a separate contract. The final impact analysis report is to be delivered in June 2015.

In the finalization phase of this research conducted during Option III, the results of the testing phase provided the data for psychometric analysis to examine the reliability and the construct and predictive validity of the battery. During finalization, instruments were examined to derive their most meaningful items, and reduction in the length of the battery was achieved. Constructs were examined to understand their contributions to decision-making performance and to classify participants at different levels of performance. Finally, the battery was extended to a version for platoon commander assessment.

Findings

Results of the analysis indicate battery quality, the predictive ability of the battery, and the ability of the battery to distinguish levels of performance. The sample population was found to be normal for all instruments and conducive to various analyses. The reliabilities of most of the instruments increased after removal of items that did not meet established criteria. The item level analysis for the Decision Requirements Interview resulted in a more evenly distributed scoring system and parallel scoring forms for the two alternate forms of the instrument. Factor analysis results indicate that each instrument measures one construct, providing support for our ability to analyze the multidimensional nature of decision making. The instruments measuring the following constructs were found to best predict decision-making performance on the Decision Requirements Interview: resilience, metacognition, and attentional control, and the SUDM Situational Judgment Test (SJT) consisting of subscales for the constructs of sensemaking, situational assessment, and decision making. As expected, NCOs and Lts are significantly different in experience and performance level, and the battery is able to distinguish between the groups. The scores on the seven scales as follows made the largest contribution towards discriminating between the performance level of NCOs and Lts: resilience, problem solving, attentional control, ambiguity tolerance, self-regulation, self-awareness, and SUDM SJT subscales of situational assessment, sensemaking, and decision making. The battery can correctly, significantly classify participants into different groups of performance levels. Instructor ratings have a poor relationship with the battery, possibly indicating issues with the ability for instructors to rate cognitively complex constructs, but the small number of ratings available for analysis means the findings are not necessarily conclusive in support of the lack of relationship or usability of the rating forms.

Recommendations

- Because a combined sample of Lts and NCOs was used in the factor analyses it is plausible that some instruments may have appeared to have more than one underlying latent construct because the factor structure of small unit decision-making ability is likely to differ across Lts and NCOs. Thus, one recommendation for future work is to ensure that in future development, a large enough sample size is available for both Lts and NCOs to examine the measurement equivalence/invariance of the factor structure across Lts and NCOs.
- All constructs in this effort should be retained for future study as the current analysis is limited by the use of a convenience sample rather than a sample of the target users and the only criterion variable available was the Decision Requirements Interview. Further investigation could result in more meaningful relationships with the remaining constructs with a more relevant sample and an external criterion measure.
- Acceptable instruments were not found for four constructs initially considered as important aspects of small unit decision making—adaptability, change detection, anomaly detection, and cognitive flexibility. For adaptability, we recommend consideration of construction of a scenario-based instrument and consideration of the I-ADAPT instrument which has a 55-item format and evidence of predictive validity. (See Baard, Rench, and Kozlowski, 2014 for a review of the adaptation construct and the I-ADAPT instrument.) For the constructs of change detection and anomaly detection, we recommend adapting or devising a measure that fits our operational definitions as no useful instruments were found. Note that our definition involves military relevant performance and not basic perceptual skills. For the construct of cognitive flexibility we suggest the development of a performance instrument in line with our definition that requires the participant to transfer principles of performance from one scenario to another.
- We recommend that at the conclusion of the project, the final battery be implemented with a relatively large sample size from the desired populations—prospective and current maneuver squad leaders and platoon commanders—to establish a robust reference group for scoring and interpretation.
- Finally, to support use by the Marine Corps, the battery should be converted into a computer-administered version following this project to mitigate test fatigue and cognitive overload by allowing the respondents to save their work, stop, and return when refreshed to a password protected assessment that must be completed within an adequate, designated amount of time from first login. The computer version should also automate scoring and interpretation. This migration to a computer-supported version should be followed by steps to transition that product to the Marine Corps including system demonstration, testing in an operational environment, and training for battery implementation.

Introduction

Improving Small Unit Decision Making

Infantry small unit leaders represent one of the most critical positions on the modern battlefield. They form the tip of the spear against irregular threats in mission environments characterized by extreme levels of complexity. Operations in Iraq and Afghanistan have plainly demonstrated the broad decision-making responsibility of small unit leaders and the strategic failures that result from poor judgment. Future operations are likewise expected to require small unit leaders who can quickly recognize and adapt to evolving situations and make sound decisions that achieve the mission objectives while mitigating against negative second and third order effects.

The Marine Corps recognizes the vital role of the small unit leader. The stated vision of the *Marine Corps Vision and Strategy (MCV&S) 2025* is for the Marine Corps to be the nation's expeditionary force of choice and to demonstrate the ability to rapidly deploy to a wide range of complex and irregular operating environments as lean, agile, and adaptable individuals and units (U.S. Marine Corps, n.d.-a). In recognition of the small unit leader's role in that vision, one directive of the *MCV&S 2025 Implementation Planning Guidance* document (U.S. Marine Corps, n.d.-b) is to develop a plan to improve the small unit leader's ability to assess, decide, and act in a more decentralized manner. Similarly, the Commandant's Planning Guidance (CPG) specifies a task to improve training and experience levels for maneuver unit squad leaders in support of decentralized operations in the 21st century hybrid threat environment (U.S. Marine Corps, 2010). In response to these demands, the Marine Corps Training and Education Command (TECOM) institutionalized a Small Unit Decision Making (SUDM) initiative (U.S. Marine Corps Training and Education Command, 2011). The goals of the initiative were not only to improve the training and overall proficiency level of decision-making skills across the population of noncommissioned officers (NCOs) who may serve as maneuver squad leaders, but also to measure decision-making abilities as a means of assessing improvements over time in decision-making readiness (the potential to perform) across the Force.

The challenges associated with measuring decision-making performance are many. Tactical decision making at the small unit level is a broad and unwieldy concept that cannot be defined as a discrete cognitive activity. While the work of Klein (1989) describes the Recognition Primed Decision process (RPD) as the most widely used process in situations such as squad leader decision making during operations, a single cognitive process fails to account for all the activities associated with RPD, and is not endorsed by the model. Instead, decision making involves a number of cognitive processes and access to a knowledge base. Therefore, assessing and improving decision making requires a multidimensional approach to performance assessment.

The decision dilemmas faced by small unit leaders are too numerous to count, let alone test as individual performance items. Furthermore, as is the case in cognitively complex performance environments, seldom if ever can a single best decision be identified for a given tactical problem. Prior assessment efforts have overcome these challenges by scoping the assessment space to a specific, well-defined set of performance parameters, or by relying on subject matter expert

(SME) ratings of decision quality as a means of quantifying decision performance. The multidimensionality of decision making is lost in the assessment process. These approaches also do not lend themselves to the Marine Corps' requirement for a scalable, generalizable assessment capability that predicts decision performance across a range of operational settings. Therefore, the SUDM Assessment Battery research project was undertaken to fill that gap.

Conceptual Framework of the SUDM Assessment Battery

Our approach to the development of an assessment battery was to consider small unit decision making as a multidimensional construct. The SUDM Assessment Battery was designed to measure skills and attributes that work together to determine the decision-making proficiency of small unit leaders—Marine Corps maneuver squad leaders, initially, and then an extension to platoon commanders. Experienced researchers (including members of our research team) and Marine Corps SMEs identified these enabling skills and attributes through a series of workshops and surveys conducted prior to the start of this project (U.S. Marine Corps Training and Education Command, 2011).

From these findings, TECOM selected five competencies and ten cognitive and relational skills (CARS) for further study as the basis for the generation of a decision-making assessment battery. The five cognitive competencies are sensemaking, problem solving, adaptability, metacognition, and attentional control. The ten CARS are perspective taking, analytical reasoning, anomaly detection, change detection, situational assessment, cognitive flexibility, ambiguity tolerance, resilience, self-regulation, and self-awareness. Our research team added the overarching construct of decision making to the requirement.

A multidimensional conceptualization of decision making immediately recognizes the complex cognitive skills required and processes undertaken. This conceptualization overcomes the narrow measurement approaches often taken that look at overarching measures of effectiveness because they are easy to observe and measure, and replaces ineffective measures of performance that are procedural and do not yield information about how decision making is taking place in terms of the complex processes that yield performance.

Morrison and Fletcher (2002) defined cognitive readiness as "the mental preparation (including skills, knowledge, abilities, motivations, and personal dispositions) an individual needs to establish and sustain competent performance in the complex and unpredictable environment of modern military operations" (p. I-3). Interestingly, they hypothesized a set of 10 "components," with some similarities to the constructs we are using, as relevant to cognitive readiness and suggested that these be measured even though "some aspects of cognitive readiness are not amenable to training…" (p. III-1). Their components of cognitive readiness are (1) situation awareness, (2) memory, (3) transfer of training (ability to apply knowledge and skills in one context to another context), (4) metacognition, (5) automaticity (rapid responses that do not substantially impair other processes), (6) problem solving (situation analysis, understanding goals, and developing a course of action to achieve goals), (7) decision making (reviewing different plans of action, assessing the probable impact of each, selecting one, and committing

resources to it), (8) mental flexibility and creativity, (9) leadership, and (10) emotion (devise and select courses of action under stress).

Subsequent work in the area of cognitive readiness has continued to debate the dimensions of cognitive readiness, often with the rationale that different definitions serve different purposes such as a training focus versus a selection focus. (See O'Neil, Perez, & Baker, 2014 for a current discussion of the definitions of cognitive readiness as well as issues of training and assessment.)

Fletcher and Wind (2014), in particular, updated the components originally proposed by Morrison and Fletcher (2002) which yielded a set of constructs more similar to those that are the focus of this effort. They include (1) situation awareness, (2) problem solving, (3) metacognition, (4) decision making, (5) adaptability, (6) creativity, (7) pattern recognition, (8) teamwork, (9) communication, (10) interpersonal skills, (11) resilience, (12) and critical thinking.

Though Fletcher and Wind propose that ideally the components should be relatively context- and content-free to avoid focusing on the anticipated and expected challenges to cognitive readiness, we have taken somewhat the opposite approach. We have found that decision making, considered one component by both Morrison and Fletcher (2002) and Fletcher and Wind (2014), is in itself multidimensional and assessing it as such allows a richer, clearer picture of this critical performance to emerge, avoids a limiting reductionist approach to assessment, and allows us to better address the original, holistic definition of cognitive readiness from Morrison and Fletcher (2002). *We devised research-based, military relevant operational definitions and examples of the constructs we examined that guided our work.* While we included measures that were context- and content-free, we have assembled them into a battery with performance measures that are definitely context- and content-focused. We do not believe the performance elements of the battery take us outside "the land of the unanticipated and unexpected" given our rich background in devising cognitively authentic scenarios (Ross, Halterman, Pierce, & Ross, 1998; Ross & Pierce, 2000).

The concept for developing an assessment of the multidimensional overarching construct of decision making was to (1) test the soundness of the proposed dimensions of decision making by identifying measures for each construct and testing how the measures work together; (2) develop a deep understanding of the small unit leader by operationalizing the construct definitions in a manner meaningful to the target audience, (3) include and/or adapt readily available, individual measures of states and traits, and devise performance tests of decision making that are content- and context-specific to best assess the constellation of constructs yielding decision-making proficiency, and (4) create a battery that is concise and focused so that it will be used and can be applied to large numbers of participants to inform service-level issues. In this manner, it was our goal to serve the desire of the Marine Corps to understand the decision-making readiness of their small unit leaders and to demonstrate the utility of a multidimensional measurement.

Components of the Battery

The instruments selected or developed for the battery shown in Table 1 include those that measure *traits* (difficult to change; require long periods of time and/or targeted training and

Table 1. Constructs, Assessment Instruments, and Instrument Type

Constructs	Assessment Instrument(s)	Acronym	State, Trait, Performance
Problem Solving	Personal Problem Solving Inventory	PPSI	S
Metacognition	Metacognitive Awareness Inventory	MAWI	S
Attention Control	Neuro-Cognitive Assessment	NCA	T
Adaptability	Adaptive Force Scale Situational Judgment Test	ASJT	P
Sensemaking	SUDM Situational Judgment Test	SUDM SJT	P
Perspective Taking	Differences in Empathy Scale	DES	T
Analytical Reasoning	Metacognitive Activities Inventory	MAI	S
Anomaly Detection			
Resilience	Brief Resilience Scale	BRS	T
	Connor-Davidson Resilience Scale	CDRS	T
Change Detection			
Situational Assessment	SUDM Situational Judgment Test	SUDM SJT	P
Cognitive Flexibility			
Ambiguity Tolerance	Multiple Stimulus Types Ambiguity Tolerance	MSTAT	T
Self-Regulation	Problem Solving Scale	PSS	S
Self-Awareness	Freiberg Mindfulness Inventory	FMI	S
Decision Making	Decision Requirements Interview	DRI	P
	SUDM Situational Judgment Test	SUDM SJT	P

experience to change), *states* (change with knowledge and experience more easily than traits; trainable), and *performance* (domain- and situation-specific decision making that can change with knowledge and practice). Those instruments identified as measuring states and traits are self-report instruments with indirect questions that provide scores allowing insight into the relative degree of the state or trait. The Small Unit Decision Making Situational Judgment Test (SUDM SJT) and the Decision Requirements Interview (DRI) are performance tests that were developed as part of the project. Four of the initial constructs do not have assessment instruments in the final form of the battery (adaptability, change detection, cognitive flexibility, and anomaly detection). The adaptability measure did not perform well, and no measures relevant to the other operational definitions or applicable to measuring various sizes of groups could be identified.

For operationalized definitions and examples of each construct see the *Marine Corps Maneuver Squad Leader Mastery Model* (Ross, Phillips, & Rivera, 2013). For more information, the literature review to identify and select instruments was documented in Vogel-Walcutt, Ross, & Knarr (2013), as well as Vogel-Walcutt, Ross, Smith, & Brown (2012).

Some scores on the assessment can be expected to change as a result of knowledge and experiences more quickly than others. However, generally, changes in the scores from the battery occur over long periods of time as mastery matures, and change varies based on experiences that broaden the knowledge base of individuals; practice and reflection opportunities; support to reflect on learning; and, the strength of the trait in the individual. While traits are difficult to change, the Marine Corps needs to be aware of the distribution of factors contributing to good decision making under stress to understand the cognitive readiness of the Force. As noted above, our approach was to understand the proposed multidimensional nature of decision making and to assess states, traits, and performance to best understand what determines decision-making proficiency. Our research team did not feel bound to limit our assessment to “trainable” skills, in particular because we have seen through many qualitative interviews that constructs considered traits can mature over time with training and experience, especially given our operational definitions.

Uses and Benefits of the SUDM Assessment Battery

To improve policy development and implementation, the SUDM Assessment Battery offers the ability to understand the decision-making proficiency of small unit leaders across the Marine Corps. Given the multidimensional nature of the battery construction, the results allow the Marine Corps, at a high level, to *gain insight into the overall proficiency of the group of current and prospective maneuver squad leaders and platoon commanders* at any given point in time. Scores on performance instruments and underlying cognitive constructs can be aggregated to paint a picture of strengths and needs for improvement that can be addressed at the service level in line with the *MCV&S 2025 Implementation Planning Guidance* task of "improving Small Unit Leader intuitive ability to assess, decide and act...." Therefore, the intended use of the battery is at the policy level to influence the training, education, and experiences of the small unit leader and to assess the impact of such actions in the overall community or sub-communities no more regularly than once a year. *Return on investment for policy interventions that are meant to improve readiness over time can be informed by a clearer, quantitative understanding of small*

unit leader development. Policy impact does not need to be limited to inferences from expert judgment, less nuanced "go/no go" assessments, or measures of effectiveness that do not take cognitive readiness into account.

The SUDM Assessment Battery supports insight into cognitive readiness at the operational and strategic levels as defined by Grier (2012) through use of the aggregate scores of the target audience. The operational level refers to the preparedness to engage in anticipated missions or deployments, while the strategic level has a longer view and refers to preparedness to engage in a wide range of potential situations.

To determine the impact of a discrete training event or to support research programs, researchers potentially can use portions of the battery. The battery was designed to support an understanding of a specific target audience in a manner not currently available in order to provide insight into policy impacts. It was not designed to be used as an intact training effectiveness evaluation tool or to support short-term research given that it was designed to be sensitive at the level of overall development over extended periods of time. It is not appropriate as a pre- and post-test for short training events as a complete battery. Uses of the battery in support of research and development could include (1) administration of the battery before research or training interventions to predict different performance across groups, (2) use of the SUDM SJT portion as a pre- and post-test, or (3) use of the two DRI alternate forms as a pre- and post-test. The researcher should understand which constructs the instruments assess. These uses assume the target audience is appropriate and the research or training intervention is appropriate for the cognitive and performance skills assessed by the battery or its components.

One drawback to using the performance instruments from the battery is that wide distribution of the instruments potentially weakens the validity of the battery for establishing a baseline and maintaining a trend line for the Force for the target groups of interest. Parallel forms of the DRI scenarios could be developed for research purposes. Guidelines can be adapted from Brummel, Rupp, and Spain (2009) for constructing future parallel scenarios by combining that information with information about the structure of the existing DRI scenarios.

Other Products Developed During This Research Effort

In addition to the development of the battery, this project produced two other products, the *Marine Corps Maneuver Squad Leader Mastery Model* (Ross, Phillips, & Rivera, 2013) and the SUDM Behaviorally Anchored Rating Scale (SUDM BARS). Each of these products also has application in the areas of training assessment, research, and policy support.

The Mastery Model was developed based on interviews with participants who were or had been squad leaders or those who supervised or trained squad leaders. It contains operationalized definitions of the constructs and examples, and its centerpiece is the five-stage model of maneuver squad leader development with multiple performance indicators for each Key Performance Area (KPA) at each stage of development.

The SUDM BARS is an observation rubric developed from the Mastery Model. Observable performance indicators are provided for each KPA and are drawn directly from the model. Originally, the BARS was targeted to become part of the battery but this was not practical for battery administration. The SUDM BARS is designed to assist SMEs in conducting objective, standardized observation-based evaluations of individual small unit leader performance. The SUDM BARS was developed by selecting from the Mastery Model the most salient and observable performance indicators from each subcategory in each KPA. The selected performance indicators were then revised for clarity and brevity, with the objective of stating characteristics of performance that would serve as differentiators across the stages of proficiency, and that would likely be apparent in a tactical exercise setting. The format of the rubric was then designed to support quick, easy reference and rating by the evaluator. The SUDM BARS is presented in Appendix A.

To apply the SUDM BARS, users—typically instructors, training officers, or supervisors—observe an individual's performance and then choose the performance description that best matches what was observed by checking the box next to the most accurate performance descriptor. When performance is rated at the Expert level, or Stage 5, the performer receives five points. Proficient, Stage 4 performance, is awarded four points, and so on down to Novice, Stage 1 performance, which is awarded one point. When the subcategory of the KPA is not relevant to the exercise or was not observed, the evaluator can select "not observed." The SUDM BARS can be applied during tactical exercises at any level of fidelity, from tactical decision games and sand table exercises, to course final exercises, live fire exercises, or pre-deployment training missions.

The SUDM BARS instrument has already successfully been used as a research tool. Design Interactive employed the SUDM BARS as part of the Small-unit Training for Adaptability and Resiliency in Decision Making (STAR-DM) research being carried out for ONR. Instructors at the School of Infantry-East used the assessment instrument following a series of scenario-based exercises to rate student performance. Differences were found in groups that the researcher had established based on variables of stress under study.

The *Marine Corps Maneuver Squad Leader Mastery Model* has been distributed by TECOM as part of the new Squad Leader Development Program (SLDP). The Mastery Model can be applied in the program in a number of ways. The application to the SLDP is in line with the definition of strategic level cognitive readiness as defined by Grier (2012). The strategic level refers to preparedness for a wide range of potential situations and encompasses, for the SLDP, overall development and readiness of the squad leader. Four areas of application that can be supported are illustrated graphically in Appendix B. The first is an understanding of performance in terms of the overall mastery orientation of the model which provides a framework for defined, observable performance indicators in key performance areas, thus promoting constructive feedback. Second, is the ability to screen or select based on the framework, but requires that the model be interpreted into a tool for that purpose. Third is the ability to construct a road map of development, and fourth is assessment such as application of the battery and the BARS derived from the model. For example, the SLDP could benefit from the establishment of a baseline and then create and examine a trend line to see if the program is improving decision-making

proficiency, and if either of two squad leader development tracks the SLDP proposed produces the desired improvement in the targeted audience.

Method

Structure of the Project: Prepare, Develop, Test, Finalize

The SUDM Assessment Battery project consisted of four phases—*preparation, development, testing, and finalization*—to achieve a reliable and valid battery sufficient for understanding small unit tactical decision making. This report is the final report for the project which consisted of a base contract and three options. The report provides an overview of the project with an emphasis on the finalization phase.

Preparation consisted of instrument selection and development and the generation of the Mastery Model. During the *development phase* instruments were examined and refined with a pilot group to develop a battery suitable for testing with a larger audience. The purpose of the *testing phase* was to make final adjustments to the instruments, administration, and scoring protocols, as needed, using a larger sample than the pilot group.

To support the testing phase, a convenience sample was identified at The Basic School (TBS) in Quantico, VA. Data were collected by administering the battery at TBS to the Basic Officer Course (BOC) companies completing the six-month course, both before and after the course. Participants consisted of a sample of NCOs in two FY13 BOC companies followed by Lieutenants (Lts) provided by TBS for each course beginning in FY14 and all the NCOs who were participating in each FY14 course to improve their ability to perform as TBS instructors as part of the Enlisted Instructor-Advisor Initiative (Desgrosseilliers & Hoffman, 2014).

In exchange for the opportunity to collect a large sample of data, our research team will be offering insights into the impact of the Enlisted Instructor-Advisor Initiative under a separate project. The SUDM Assessment Battery project will use the portion of the data collected from FY13 and FY14 companies between July 2013 and July 2014 to support completion of the battery. Data collection will continue at TBS through March 2015 to complete the impact analysis study under a separate contract. The final impact analysis report is to be delivered in June 2015.

In the *finalization phase* of this research conducted during Option III, the results of the testing phase provided the data for psychometric analysis to examine the reliability and the construct and predictive validity of the battery. Results determined which constructs underlying squad leader decision making can be meaningfully measured to assess overall decision-making proficiency, and supported insight into the multidimensional nature of that performance. During finalization, instruments were examined to derive their most meaningful items, and reduction in the length of the battery was achieved. Finally, the battery was extended to a version for platoon commander assessment.

Participants

For the psychometric analysis conducted under Option III to finalize the battery, *only the data collected prior to the BOC for each company (i.e., pre-BOC data) were used for the majority of the analysis* in order to create a more consistent data set and not introduce response variations based on learning experiences in BOC. Selected portions of the data were used for some analyses when missing data based on instrument responses made some portions of the data unusable. Some analyses included post-BOC data and are identified as such in the Findings section below. Table 2 provides a description of the participants by company. Data collected between July 2013 and July 2014 were available for analysis, and "x" indicates that future data collections will be conducted for that group and used as part of another project.

Table 2. Participants from the Basic Officer Course (BOC) at The Basic School (TBS)

	Pre-BOC Data Collection			Post-BOC Data Collection		
Company	**NCO**	**LT**	**Total**	**NCO**	**LT**	**Total**
FY 13 E	15	0	15	14	0	14
FY 13 F	12	0	12	9	0	9
FY 14 A	15	45	60	10	43	53
FY 14 B	14	32	46	12	29	41
FY 14 C	7	57	64	x	x	x
FY 14 D	4	59	63	x	x	x
FY 14 E	5	51	56	x	x	x
FY 14 F	x	x	x	x	x	x
Total	72	244	316	45	72	117

Materials

The SUDM Assessment Battery measures the competencies and CARS previously determined by TECOM to be supportive of decision-making proficiency. Each of these constructs and the associated assessment instrument are shown in Table 1 above. Additional materials consisted of a supervisor rating form developed by our team to rate NCO performance in the BOC. Our research team also collected Command Evaluation Forms, a TBS product created to rate Lts who comprise the student population. Rating forms were intended to function as a criterion variable for analysis. Rating forms were not completed consistently throughout the period of data collection and only a small number of ratings for NCO participants were available for analysis.

Procedure

All SUDM Assessment Battery administrations consisted of two parts: (1) a classroom session for the administration of the test booklet in a group setting, and (2) individual interview sessions for the DRI. To reduce cognitive load on the participant, the classroom and interview sessions were typically conducted on separate days. During the classroom sessions the participants were allotted three hours to complete the test booklet. On average, participants completed the booklet in less than two hours. The interview sessions were allocated two hours to complete. Informed consent was obtained at the start of either the classroom session or the interview, whichever occurred first. At all administrations, TBS provided someone to speak to the participants about the importance of diligently completing the assessment. Not all participants, especially the NCOs, attended those informational sessions.

Analysis

The process to prepare the data from each data collection for analysis took approximately three weeks to complete. After all data entry was completed, accuracy checks were conducted to ensure all data entry was correct. All data entry was completed by hand via Microsoft Excel and then exported into SPSS for data cleaning. During the data cleaning process in SPSS, all missing data, reverse coding, subscales, and composite scores were computed. Preliminary analyses (i.e., descriptives, histograms, reliability) were computed to ensure no mistakes were made during the data preparation process.

The psychometric analysis for the finalization of the SUDM Assessment battery was based on three levels of analysis: item, battery, and external relationships. The analysis began with a detailed look at each instrument individually and more specifically the items that make up each instrument. This level included a normality test, reliability and validity analyses, and analyses to ensure the DRI alternate forms were parallel. At the next level, the focus was on the battery and included factor analysis and multiple regression analysis. Finally, we examined the relationships between the battery and external variables.

Findings

Item Level Analysis

The focus at the item level was to better understand whether the off-the-shelf (OTS) and custom-instruments are of sufficient quality. We conducted tests of normality, ensured validity, examined the reliability, and identified any issues with inter-item correlations.

Normality Test

To test whether the distribution of the sample population was normal for each instrument, Q-Q plots were examined. Q-Q plots were used rather than tests of normality like the Shapiro-Wilk test or the Kolmogorov-Smirnov test because those tend to be overly sensitive to sample size, suggesting that distributions are non-normal particularly when sample size is greater than 200. With Q-Q plots, extreme deviations from normality are easier to detect as values (i.e., dots/circles on the graph) that are far away from the line.

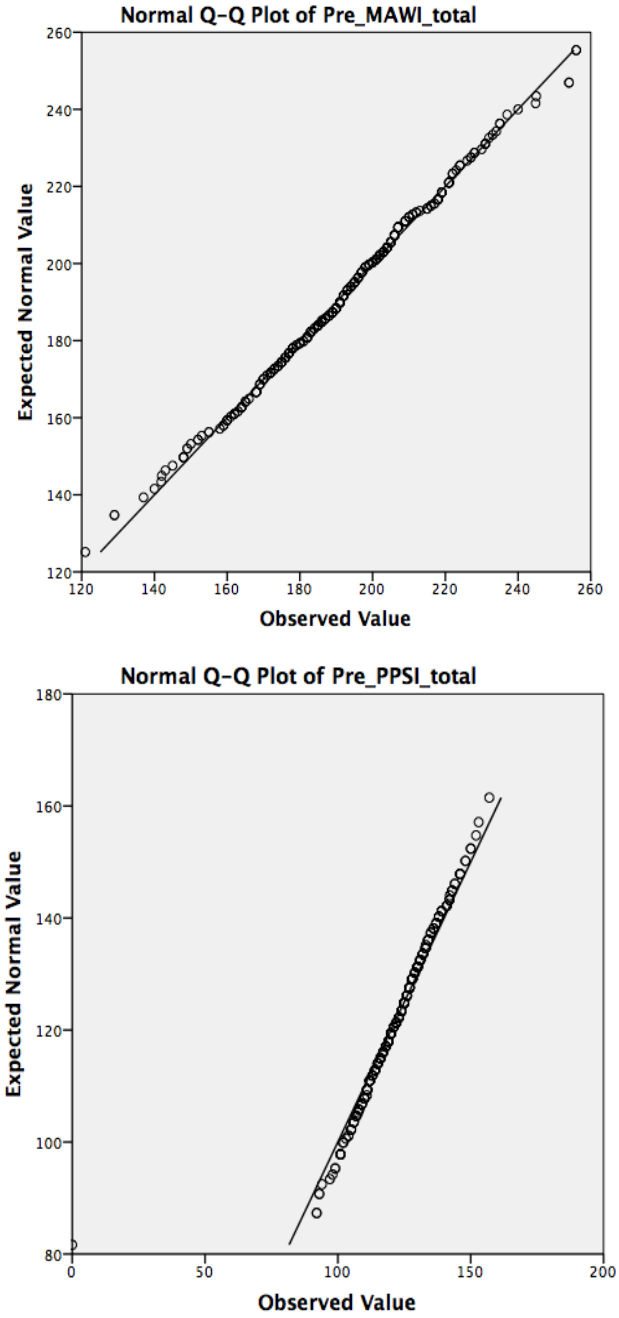


Figure 1. Example Q-Q Plots illustrating the method for determing a normal population for each instrument in the battery

The normality test showed that the sample population was normal for each instrument. Figure 1 shows two examples of normally distributed Q-Q plots, one for the Metacognitive Awareness Inventory (MAWI) and one for the Personal Problem Solving Inventory (PPSI). These illustrations show the manner in which normality can be examined by this method.

Validity and Reliability

Prior validity and internal consistency reliability data were available from the literature for the OTS instruments. The internal consistency reliability cutoff was established at 0.70 (Cronbach, 1951; Nunnally, 1978). Internal consistency reliability of each scale for our sample was previously established as adequate during the testing phase of the project. The reliability and validity data in Table 3 were updated during the current analysis for each instrument using the final version of each instrument after items were removed as described below in this report.

Table 3. Reliability and Validity of the Final Battery Instruments

Instrument	Validity from Literature	Internal Consistency Reliability from Literature	Internal Consistency Reliability for Final Version	
			Pre	Post
Brief Resilience Scale	Convergent; Divergent	0.8-0.91	0.81	0.83*
Personal Problem Solving Inventory	Construct	0.90	0.87	0.88*
Problem Solving Scale	Divergent	0.81	0.77	0.87
Freiberg Mindfulness Inventory	Construct	0.86	0.59*	0.64*
Connor-Davidson Resilience Scale	Convergent; Divergent	0.89	0.90*	0.90*
Metacognitive Awareness Inventory	Construct	0.95	0.94	0.96
Neuro-Cognitive Assessment	Construct	0.98	0.95	0.96
Metacognitive Activities Inventory	Construct	-	0.90*	0.91*
Multiple Stimulus Types Ambiguity Tolerance	Criterion	0.86	0.86	0.89*
Differences in Empathy Scale	N/A	-	0.75*	0.74
SUDM Situational Judgment Test	N/A	-	0.23	0.54*

**Reliability increased for the final form of the instrument.*

In addition, means and standard deviations of the items were also examined to ensure they were within reasonable levels given an instrument's range of possible values. The reliabilities highlighted in red are lower than the acceptable value of 0.70. However, it is no surprise the Situational Judgment Test had a lower reliability because of the multidimensional nature of SJTs. The Freiburg Mindfulness Inventory, which measures self-awareness, was cut down from its original set of items to a total of 7 items to best match the operational definition. This adjustment may be the reason for the lower reliability than previously published. Overall, the majority of the instruments had acceptable to high internal consistency reliability providing evidence for the quality of the instruments.

To further investigate the quality of the items, corrected item-total correlations were calculated for each measure. Positive corrected item-total correlations are desirable because this suggests that the focal item correlates with the total score without including the focal item into the total score. For most instruments, corrected item-total correlations were positive and at acceptable levels. However, some instruments included items that had low, or in some instances, negative corrected item-total correlations. Note that the negative corrected item-total correlations were checked for errors in reverse-coding and that was confirmed not to be an issue. As a result of the inter-item correlations and reliability analysis, items with negative and low correlations were removed to ensure sufficient quality for continued analysis and administration. See Table 4. After the removal of the items, the estimated internal consistency reliability improved while reducing the length of the battery.

Table 4. Items Removed to Increase Instrument Quality

Instruments	Items Removed
Freiburg Mindfulness Inventory (FMI)	6
Connor-Davidson Resilience Scale (CDRS)	3
Metacognitive Activities Inventory (MAI)	11, 12, 20, 21, 22, 23, 24, 25, 26, 27
Multiple Stimulus Types Ambiguity Tolerance (MSTAT)	3, 14, 15
Differences in Empathy Scale (DES)	4, 8
Personal Problem Solving Inventory (PPSI)	Personal Control Subscale

DRI Item Analysis to Ensure Parallel Forms

The purpose of the DRI is to measure the performance of decision making as a separate construct. *We defined decision making as the act of recognizing and interpreting situational indicators and understanding the perspective of others, in the context of a specific goal or goals, in order to commit assets with an understanding of the likely first, second, and third order effects.*

The DRI currently consists of two scenarios that offer alternate forms of assessment for the instrument (Jafarani Patrol and Bangui Offensive). Each scenario consists of several subscales that make up the requirements of decision-making performance: Cue Recognition, Recognition, Perspective Taking, Asset Employment, First Order Effects, and Second and Third Order Effects. These elements of decision making and their performance indicators (Desired Actions) were informed by the stage model of tactical thinking produced by members of our team in an earlier research effort (Ross, Phillips, Klein, & Cohn, 2005) as they were being developed, and once the Mastery Model was completed, KPAs were linked to each Desired Action. The Desired Actions were validated at the USMC School of Infantry-West.

Each subscale is scored on Desired Actions that the participant either does or does not indicate in his or her response. (Participants freely discuss the actions they would take at scenario pause points and do not choose from a list of available Desired Actions. Desired Actions are scored by analyzing the free form answers during the interviews against our list of Desired Actions and checking that analysis after the interview.) Table 5 below provides a summary of the DRI subscales.

Table 5. Subscales of the Decision Requirements Inventory (DRI)

DRI Subscale	Definition
Cue Recognition	Recognition and interpretation of relevant cues to assess situation and factors given the goal or mission
Recognition	Automatic and recognition-primed situational interpretation; recognition versus deliberation
Perspective Taking	Consideration of other perspectives (e.g., enemy, noncombatants, host nation forces); looking at the situation from another person's perspective, often in order to identify the cues others are looking for or actions others might take
Asset Employment	Knowing and using assets effectively
First Order Effects	Understanding of the potential impact of actions on the situation (first order effects); direct impact on leader, unit, or situation in the immediate timeframe
Second and Third Order Effects	Consideration of second/third order effects of actions; through mental simulation, understanding how actions or reactions affect the leader, own unit, other units (higher or adjacent), or situation in the future

The scenario-based alternate forms of the DRI were constructed independently and were not necessarily equivalent. Therefore, two objectives guided the DRI item analysis: (1) ensure parallel scoring between the two DRI alternate forms and (2) reduce the scoring guide to only the Desired Actions that distinguish among performers.

First, a qualitative examination revealed that the Recognition Subscale was distinctly different between the Bangui Offensive and Jafarani Patrol scenarios. In Jafarani Patrol, participants were only penalized for the recognition items (i.e., pre-defined Desired Actions); while in Bangui

Offensive, they were both penalized and awarded points. For example, a positive Bangui Offensive Desired Action is "Deciding on a COA quickly," which receives 5 points while "Changing mind about COA" receives -3 points. Also, the total points available varied across scenarios.

The Recognition Subscale in each scenario was used differently in participant responses and resulted in different scores across the two scenarios. Comparing histograms for the two scenarios, we found that the Jafarani Patrol Recognition items most frequently resulted in a score of 0, which shows that the Recognition Desired Actions are rarely given as responses by the participants. The histogram for Bangui Offensive Recognition items shows the most frequent score of 10 with 15 people from the population attaining this score. Therefore, the most different aspect of the two scenarios was the Recognition Subscale in terms of the value it contributed to the total scores.

To test whether an actual difference existed between the two scenarios, a t-test was run with the Recognition Subscale retained in each scenario. The result was no significant differences between the two scenarios ($t= -1.344$, $p= 0.180$) leading to an initial conclusion that the two scenarios were of equal difficulty. However, in the Bangui Offensive scenario the participants have opportunities to score higher on Recognition items, which we postulated may inflate the overall Bangui Offensive score in comparison to the Jafarani Patrol score leading to false equivalency between two scenarios that are not equal.

To further test the equivalency of the scenarios, we removed the Recognition items in both scenarios and the t-tests were conducted again. Significant differences were found, with Jafarani Patrol ($M=0.22$) having a higher mean than Bangui Offensive ($M=0.19$) ($t = -2.28$, $p= 0.027$). The significantly lower scores on the Bangui Offensive without the added benefit of points from the recognition items suggest Bangui may be a more difficult scenario than the Jafarani Patrol. Because findings suggest that the Recognition scale was inflating the Bangui Offensive scores and it was rarely used to contribute to the Jafarani score, *the recommendation was to remove the Recognition Subscale from both scenarios.*

Second, all the Desired Actions across all scales were examined. The objectives were to determine which items were discriminators and distribute effective Desired Action items equally across the two scenarios. The percentage of times a Desired Action was used in the responses was calculated. The Desired Actions rarely displayed (typically ranging from 0% to 8%) were removed. Similarly, the Desired Actions with very high percentages (typically ranging from 66% to 79%) were also removed, because those actions are not discriminators if most participants demonstrate them. The goal was to have an average level of difficulty with a smaller distribution at the ends of the scoring continuum (i.e., fewer Desired Actions that are really hard or really easy). This approach provided an even score distribution and equivalent final scores.

Appendix C provides the list of the items removed from each scenario along with the percentage of participants that displayed that Desired Action. Before items were removed from the two forms of the DRI, the total possible score for Bangui Offensive was 397, and the total score available for Jafarani Patrol was 406. After all the changes were made the maximum scores

possible for each scenario were equivalent at 374. Table 6 below provides the final scores available for each subscale for each form of the DRI.

Table 6. Maximum Total Scores for Each Decision Requirements Interview Subscale

Scenario	First Order Effects	Second and Third Order Effects	Perspective Taking	Cue Recognition	Asset Employment	Total
Bangui Offensive	107	41	49	74	103	374
Jafarani Patrol	109	45	56	91	73	374

Battery Level Analysis

All problematic items for the OTS instruments and the DRI were removed prior to the battery level of analysis. In this level of analysis, the instruments in Table 1 above were compared and evaluated as a complete unit using the techniques of factor analysis and multiple regression analysis. In this way, it is possible to identify the most efficient structure of the battery, including which set of instruments are the best predictors of decision making performance. Moreover, it provides a better understanding of the interaction among the components of decision making.

Factor Analysis

The goal of the factor analysis was to determine whether each of the instruments truly measured a discrete construct. Because the OTS instruments were previously developed and validated, each instrument should theoretically be one-dimensional. For this reason, a principal axis factor analysis was conducted on correlation matrices of each instrument in the SUDM Assessment Battery for the sample of Lts and NCOs. To demonstrate unidimensionality, the factor analysis would have to show that the first factor (i.e., underlying construct) extracted had an eigenvalue greater than 1.00 and greater than all remaining factors. Examination of the scree plots should also show that after one factor the line drops sharply.

Results of the factor analysis indicate that for all the instruments the first factor had an eigenvalue that was large (greater than 1.00) relative to all the remaining factor eigenvalues, showing support for unidimensionality. In the case of BRS (resilience), FMI (self-awareness), NCA (attentional control), and DES (perspective taking), the magnitude of eigenvalues and the scree plot made clear that one underlying construct existed for each of those instruments. However, for the remaining instruments, more than one factor showed an eigenvalue greater than 1.00 and scree plots indicated the potential for more than one underlying construct. This outcome required further investigation and additional factor analysis to determine whether the instruments were measuring more than one underlying construct.

Additional factor analyses were conducted on the remaining instruments to determine whether more than one factor could be extracted with an eigenvalue greater than 1.00. Because BRS, FMI, NCA, and DES resulted in single factors, the solution was not rotated. To further examine the other instruments, the extracted factors with an eigenvalue greater than 1.00 were rotated using an orthogonal rotation—varimax. It is important to note that all methods of rotation

redistribute the amount of variance explained by each factor such that the first factor will no longer explain the maximal amount of variance, but the variance is instead spread across the extracted factors. The purpose of conducting a rotation is to make the output easier to interpret. It puts the results on an axis that is typically easier to read and understand.

A number of notable findings emerged. The cut off for grouping an item under the factor was 0.39. For all unrotated solutions, the factor loadings (i.e., the correlation between the item and that factor) on the first factor were large in magnitude and positive. Relatively few loadings were less than .40, which is ideal. For the remaining potential additional factors, very few factor loadings were greater than .40, suggesting that the first factor can account for most of the variance in the items. After the varimax rotation, the interpretation of the factor loadings typically should improve. However, in all instances, the interpretation became more difficult. For this reason, only the unrotated solution was used to interpret findings. Table 7 provides a summary of the unrotated factor analysis results for each instrument. (See Table 1 for the full name and related construct for each instrument.) ***Variable*** indicates the instrument name and ***Number of items*** is the number of items in the scale after removal of items as described in the Item Level Analysis section above. ***KMO MSA*** or Keiser-Meyer-Olkin Measure of Sampling tests whether it is appropriate to conduct a factor analysis with the sample population. A KMO MSA lower than 0.6 is inadequate for conducting a factor analysis and anything over a 0.8 is ideal. ***Number of loadings > .39*** indicates the number of items that fell under the first factor and ***Range of loadings > .30 on Factor 1*** indicates the strength of the correlation between those items that fell on factor one and the factor.

Table 7. Summary of Unrotated Factor Analysis for Each Instrument

Variable	Number of Items	KMO MSA	Number of Loadings > .39	Range of Loadings > .39 on Factor 1
BRS	6	0.856	6	.506-.733
PPSI	27	0.888	19	.444-.676
PSS	15	0.794	10	.415-.584
FMI	5	0.714	4	.444-.577
CDRS	24	0.914	21	.391-.708
MAWI	52	0.908	43	.397-.648
NCA	17	0.950	17	.556-.837
MAI	17	0.912	17	.527-.704
MSTAT	19	0.878	16	.409-.744
DES	7	0.829	6	.417-.743
ASJT	15	0.535	N/A	N/A
SUDM SJT	17	0.527	N/A	N/A

The results demonstrate that all the instruments except for the ASJT and SUDM SJT were suitable for conducting factor analysis. This finding is probably due to the multidimensional nature of the SJTs. Additionally, for most of the instruments, the majority of the items loaded on

only one factor with a strong positive magnitude. This finding, along with previous content analysis conducted by our team comparing items to our operational definitions of each construct, demonstrates that each of the instruments does measure only the construct it is intended to measure.

In addition to identifying the number of constructs measured by each instrument, the unrotated solution also provided suggestions for additional items that could be removed to improve the efficiency of the battery. Table 8 summarizes further items that were removed for the final version of the battery in addition to those removed as documented in the Item Level Analysis section above.

Table 8. Items Removed as a Result of Factor Analysis

Instruments	Item Number Removed	Reason for Removal
Personal Problem Solving Inventory (PPSI)	1, 5, 8, 9, 13, 15, 16, 17	No loadings on first factor > .39
Problem Solving Scale (PSS)	1, 5, 6, 8, 12, 15	High factor loadings across potential factors
Connor-Davidson Resilience Scale (CDRS)	2, 6, 9, 20	High factor loadings across potential factors
Metacognitive Awareness Inventory (MAWI)	5, 7, 15, 17, 25, 37, 46, 48, 50	High factor loadings across potential factors
Metacognitive Activities Inventory (MAI)	6, 7, 17	High factor loadings across potential factors
Multiple Stimulus Types Ambiguity Tolerance (MSTAT)	1, 5, 16, 17, 18, 19, 21	High factor loadings across potential factors

Multiple Regression Analysis

Multiple regression analysis provides insight into the constructs that best predict decision-making performance while accounting for the relationships among all the constructs simultaneously. The analysis was performed using the DRI decision-making scores as the criterion variable to understand how the other constructs contributed to understanding decision-making performance.

Prior to conducting a multiple regression analysis, we established that NCOs and Lts as groups that were significantly different from one another. The NCOs averaged 6.1 years of service while the Lts averaged 1.4 years of service. In addition, on the DRI, NCOs obtained a significantly greater mean (M = 87.18, SD = 40.51) than Lts (M = 75.15, SD = 28.03) as expected since the DRI is significantly correlated with years in service (r = 0.137, p = 0.017). Thus, conducting a multiple regression on the pre-BOC data from the SUDM Assessment Battery to predict

performance (DRI scores) would generally be difficult to interpret due to the two heterogeneous populations of participants. Consistent with this difference in groups, Box's *M* test for the equality of covariance matrices was conducted using the 12 instruments shown in Table 9 (*M* = 145.44, $p < .001$), indicating that the covariance matrix among the instruments was not the same for NCOs and Lts.

Because NCOs and Lts differed from one another in terms of years of service and DRI scores, a strictly bivariate approach to examining relationships with the SUDM Assessment Battery was first used. In other words, correlations were calculated between each instrument and the DRI, but separately for NCOs and Lts. Then, tests on the equality of independent correlations were conducted. The purpose of conducting these tests was to examine whether the correlation between an instrument and DRI scores was different for NCOs versus Lts. None of the correlation coefficients differed significantly between NCOs and Lts as shown in Table 9. See Table 1 for the full name and related construct for each instrument.

Table 9. Correlation Coefficient of Each Instrument with DRI Scores

	NCOs		Lts		
Instrument	Pearson's *r*	*n*	Pearson's *r*	*n*	Test on Independent Correlations
BRS	0.070	69	0.117	228	-0.33878
PPSI	0.008	73	-0.006	232	0.10251
PSS	-0.100	73	0.006	232	-0.77859
FMI	-0.066	73	0.066	231	-0.96742
CDRS	0.042	73	0.024	232	0.13194
MAWI	-0.112	73	-0.073	232	-0.28806
NCA	-0.241	73	-0.085	232	-1.17614
MAI	-0.132	73	0.016	232	-1.08934
MSTAT	0.002	73	-0.070	232	0.52803
DES	-0.006	73	0.033	232	-0.28565
ASJT	0.007	69	-0.036	230	0.30759
SUDM SJT	0.131	69	0.080	230	0.36888

Although tests on independent correlations can be useful to determine whether a bivariate relationship exists between two groups, it does not take into account the relationships among all the variables at once. For this reason, a multiple regression using all 12 instruments to predict DRI scores was conducted while controlling for Rank. This analysis resulted in an overall significant model $R = .283$, $F(13, 275) = 1.846$, $p = .036$. Note that the regression coefficient for the dummy variable for Rank ($\beta = .153$, $p = .012$) was statistically significant, further supporting that the multidimensional surface of the 12 instruments when predicting DRI scores was not the same for NCOs and Lts. This finding suggests that the battery as a whole can predict decision-making performance as defined by DRI scores, accounting for 28.3% of the variance.

Using the 12 instruments and a backward selection approach, *four instruments [BRS (resilience), MAWI (metacognition), NCA (attentional control), and SUDM SJT (consisting of the subscales sensemaking, situational assessment, decision making)] significantly predicted DRI scores at alpha = .05 (one-tailed).* See Table 10. Note that the dummy variable for Rank remained statistically significant suggesting that the multidimensional regression surface differed between NCOs and Lts. The resulting model, $R = .279$, $F(6,283) = 3.981$, $p < .001$, indicates that this subset of four instruments accounts for 28% of the variance in the DRI scores.

Table 10. Prediction of DRI Scores as the Criterion Variable

			p value	
Predictor	**β**	***t***	**two-tailed**	**one-tailed**
BRS	0.148	2.223	0.027	0.0135
MAWI	-0.177	-2.739	0.007	0.0035
NCA	-0.106	-1.788	0.075	0.0375
SUDM SJT	0.106	1.794	0.074	0.037
Rank_dummy	0.153	2.613	0.009	0.0045

Relationship Level Analysis

After an examination of the items within the instruments, as well as the behavior of the battery as a whole, we examined how the battery is related to external variables. Investigating how the battery relates to years of experience, rank, and instructor ratings provides a better understanding of the interpretations derived from the battery results.

Rank and Years of Service

Given that NCOs and Lts are different from one another in terms of years of service and DRI scores, another approach was used to understand how the relationships measured in the SUDM Assessment Battery could discriminate between the performance of the NCO and Lt groups. Specifically, a discriminant function analysis was conducted, using the instruments as dependent variables and Rank as the independent grouping variable. This approach can identify a set of weights associated with the dependent variables which result in a linear combination that provides maximum separation (or discrimination) between groups (Rencher, 2002), that is, NCOs (more experience, higher DRI scores) and Lts (less experience, lower DRI scores). In addition to assessing the magnitude of the relationship between the linear combination of dependent variables and group membership, the standardized discriminant function coefficients were inspected and a classification analysis was conducted.

The discriminant function was statistically significant, χ^2 (12) = 25.28, $p = .014$. The canonical correlation, analogous to a multiple correlation, was .288, indicating that the percentage of explained variance between the linear combination of dependent variables and group membership was 28.8%. *An inspection of the standardized discriminant function coefficients suggests that scores on BRS (resilience), PPSI (problem solving), NCA (attentional control), MSTAT (ambiguity tolerance), PSS (self-regulation), FMI (self-awareness), and SUDM SJT (situational assessment, sensemaking, decision making) made the largest relative contribution*

towards the separation of NCOs and Lts. Stated differently, scores on these seven instruments tend to distinguish between NCOs (more experience, higher DRI scores) and Lts (less experience, lower DRI scores). See Table 11. It is also worth noting that the signs of these coefficients are not all positive and that the maximum separation between NCOs and Lts on the battery is provided by a contrast between scores on BRS, PPSI, NCA, and MSTAT (i.e., positive sign) on the one hand and PSS, FMI, and SUDM SJT (i.e., negative sign) on the other.

Table 11. Standardized Canonical Discriminant Function Coefficients

Instrument	Function
BRS	.673
PPSI	.463
PSS	-.340
FMI	-.347
CDRS	.154
MAWI	-.259
NCA	.495
MAI	.257
MSTAT	.430
DES	-.017
ASJT	.223
SUDM SJT	-.437

Related to discriminant function analysis, a classification analysis using Fisher's linear discriminant functions (Rencher, 2002) was conducted to determine which instruments could be used to correctly identify whether a respondent was an NCO (more experience, higher DRI scores) or Lt (less experience, lower DRI scores) and thus, differentiate between individuals functioning at different levels of decision-making proficiency. Generally, there were relatively few false positives and false negatives. *The function correctly classified 65.2% of the sample into either the NCO or Lt group.* However, classification rates can be overly optimistic. To reduce this bias, the holdout method was also used which removes each case prior to computing the classification functions (Rencher, 2002). Then, the classification functions are used to classify the omitted case. Using the holdout method, we correctly classified 60.5% of the sample into either the NCO or Lt group. Overall, even with the holdout method, the classification rate was good, suggesting that the scores on the SUDM Assessment Battery can be useful in classifying individuals.

Instructor Ratings

The NCO performance ratings provided by instructors were compared to the SUDM Assessment Battery to identify whether the battery can measure performance similar to levels of performance as rated by instructors that are with students every day. A limitation of this analysis is that only 38 instructor ratings were available for all NCOs. For the Lts only instructor ratings (i.e., Command Evaluation Forms) were available for FY14 A Company with a sample size less than 45. To prevent bias from using only one company and due to the small sample size and scoring

approach (i.e., Command Evaluation Forms are based on learning objectives/areas of proficiency and not constructs), the Lt Instructor ratings were not analyzed.

The NCO instructor ratings were first correlated with the DRI using a Pearson's Moment Correlation to determine whether they both measured similar performance. No significant correlations were found between total Instructor rating scores and DRI ($r = 0.223$, $p = 0.184$). This was also the case for subscales of instructor ratings and DRI performance.

Each of the instrument pre- and post-scores was correlated with total instructor ratings. For post-test scores there were no significant correlations across the instruments. However, for the pre-test scores there were significant negative correlations between instructor ratings and PSS (self-regulation) ($r = -0.339$, $p = 0.037$), MAWI (metacognition) ($r = -0.437$, $p = 0.006$), and MAI (analytical reasoning) ($r = -0.340$, $p = 0.037$). Although caution should be taken before interpreting relationships based on small sample sizes, this may be an indicator that asking instructors to rate performance on cognitively complex constructs may be difficult for them and not effective.

To get an in-depth view of the relationship between instructor ratings and the instruments, the instructor rating subscales were correlated with their relevant instruments. For example, the instructor rating subscale of perspective-taking was correlated with the DES. Likewise, the instructor rating subscale of resilience was correlated with the BRS and CDRS. In all accounts, no significant correlations were found between the instructor ratings subscales and the individual related instruments.

Summary of the Analysis

The items below summarize the key findings of the analysis that indicate battery quality, the predictive ability of the battery, and the ability of the battery to distinguish levels of performance.

- The sample population was found to be normal for all instruments and conducive to various analyses.
- The reliabilities of most of the instruments increased after removal of items that did not meet established criteria.
- The item level analysis for the DRI resulted in a more evenly distributed scoring system and parallel scoring forms.
- Factor analysis results indicate that each instrument measures one construct, providing support for our ability to analyze the multidimensional nature of decision making.
- The BRS (resilience), MAWI (metacognition), NCA (attentional control), and SUDM SJT (sensemaking, situational assessment, decision making) instruments were found to be the strongest predictors of decision-making performance as measured by the DRI.
- As expected, NCOs and Lts are significantly different in experience and performance level and the battery is able to distinguish between the groups. The scores on seven scales—BRS (resilience), PPSI (problem solving), NCA (attentional control), MSTAT (ambiguity tolerance), PSS (self-regulation), FMI (self-awareness), and SUDM SJT

(situational assessment, sensemaking, and decision making)—made the largest contribution towards the difference between the performance level of NCOs and Lts.
- The battery can correctly classify participants into different groups of performance levels. A classification analysis using Fisher's linear discriminant functions (Rencher, 2002) was conducted to determine which instruments could be used to correctly identify whether a respondent was an NCO (more experience, higher DRI scores) or Lt (less experience, lower DRI scores) and thus, differentiate between individuals functioning at different levels of decision-making proficiency. The analysis correctly classified 60.5% of the sample into either the NCO or Lt group.
- Instructor ratings have a poor relationship with the battery, possibly indicating issues with the ability for instructors to rate cognitively complex constructs, but the small number of ratings available for analysis means the findings are not necessarily conclusive in support of the lack of relationship or usability of the rating forms.

Finalization of the Battery

Summary of Changes from the Test Version to the Final Version

As a result of the psychometric analyses, the following revisions were made to the finalize the battery:

- Removed items recommended by internal consistency reliability, inter-item correlation, and factor analysis.
- Changed rating scales for two instruments to mirror the numbers that are entered in the SUDM Score Coding Form (the Excel file into which Battery data are entered) to make them easier to score.
- Re-formatted all instruments to make them similar and easier for respondents to complete.
- Removed the Adaptive Force Scale SJT (ASJT) from the final battery for the following reasons:
    - Does not correlate well with any other measures in the battery;
    - Internal consistency reliability is well below the cut-off;
    - Not a predictor of performance (as measured by the DRI);
    - Does not distinguish between experienced and non-experienced performers (as measured by rank);
    - Does not match well with the operationalized definition of adaptability or the type of decision making that is intended to be measured by the SUDM Battery;
    - Adds considerable time to the battery administration without results in line with the time investment.
- Four constructs are not addressed in the battery. No suitable instruments were found for Change Detection or Anomaly Detection. These constructs were never included in the Testing Phase. Cognitive Flexibility and Adaptability were addressed by testing two instruments identified in the literature, but those instruments were eliminated.

Interpretation of Scores

To aid in the interpretation of the total scores, we focused on a subset of respondents judged to be the most similar to highly skilled participants in the future battery administration sample to create a provisional reference group for the Administration Manual. The group drawn from the total sample consisted of 21 NCOs and 1 Lt. Appendix D shows the demographics of the sample used. The criterion to be included in the reference group was more than 5 years in service, at least one deployment and experience conducting dismounted patrols. If the information on the type of patrol was not available but the other criteria were met, then the participant was included to achieve the largest sample possible.

We used post-test data from this sample because we reasoned that those with the most experience and training would attain higher scores on the SUDM Assessment Battery at post-test and that this would be more representative of a population of skilled performers compared to using pre-test scores of respondents or using respondents with no deployments. Thus, because data from a reference group of relevant experts was not available, this subset of respondents served as a pseudo-reference group through the application of bootstrapping to create a new sample from a group selected within the convenience sample used in this study.

It is important to note that the sample size was small (n = 22) for this reference group (i.e., experienced respondents with available post-test data). There could be concerns that percentiles associated with scores on the tests in the SUDM Assessment Battery might be unstable. To ameliorate this concern, we applied a bootstrapping procedure.

Bootstrapping is a method that "generates new samples from the original sample and requires no assumptions about the distribution of scores in the population" (King, Rosopa, & Minium, 2010, p. 386). It considers a sample as though it were the population. A sample of observations is taken from the original sample with replacement and relevant statistics (e.g., mean or median) are calculated. This sample is called a bootstrap sample. Note that an observation in the original sample could appear in the bootstrap sample more than once. On the other hand, it is also possible that an observation in the original sample might not appear in the bootstrap sample. After calculating the relevant statistics using the first bootstrap sample, a second bootstrap sample is obtained and the relevant statistics calculated again. Typically, thousands of bootstrap samples are obtained. Therefore, there would be thousands of statistics (e.g., means or medians). The statistics associated with each bootstrap sample are stored/recorded by the analyst, thus creating a sampling distribution of the relevant statistics (Efron & Tibshirani, 1993).

In the current report, the percentiles are actually the mean percentiles based on 5,000 bootstrap samples. Note that these percentiles could be considered to be a much fairer or unbiased representation of the percentiles associated with scores on the tests in the SUDM Assessment Battery from our reference group than those based on the original sample of 22. The resulting percentile scores are used in the Administration Manual to support interpretation of scores. The manual for administering, scoring, and interpreting the battery has been created and submitted separately from this report using these percentiles to support interpretation. The percentiles aid in the interpretation of individual scores relative to scores from an experienced population. For

interpretation of unit level scores, an average of all individual scores should be compared to percentiles in the table and the performance interpretation descriptions provided in the manual.

Discussion

The Challenge of Assessing Small Unit Decision Making

Decision making is an attractive construct to address in the research community, because the essence of what happens in military operations is dependent on decision making. Most of us may feel we would know decision making when we see it, but upon closer examination, the complexity of the process does not lend itself to a consistently agreed upon definition. Previous attempts to understand decision making have approached it as a singular construct. Instruments developed from that theoretical basis can lack the sensitivity required to distinguish all the cognitive processes that are exercised concurrently when making decisions. To measure decision making, we cannot examine only the act of comparing options, study the Marine's analysis of the constraints and benefits to committing resources in a particular way in the context of a set of goals, or measure the outcomes of carrying out a plan. Instead, our approach to understanding and assessing decision making is dependent on the assertion that the decision making that matters in today's hybrid warfighting environment is multidimensional and the different cognitive dimensions that work together during decision making can be assessed and supported to understand and improve military decision making.

Inherently, good assessment of decision making is time consuming. Subject-matter experts generally need extensive time and observation to understand and assess proficiency. To improve the decision-making proficiency of the Marine Corps small unit leader, we must improve the assessment capabilities of the Marine Corps. The goal of this project was a battery that is minimally time-consuming but still able to provide nuanced information, that can be easily administered, scored, and interpreted by non-researchers, and that does not place a heavy burden on the participants causing them to provide data that is not optimally useful. The battery must take into account and measure the multiple dimensions of decision making and avoid reducing complex performance to that which is easiest to measure to produce the most useful information.

During Option III, the final part of the project, we conducted a psychometric analysis, improved and revised the battery into a final form within the bounds of the time and resources of this project, reduced the time needed to administer the battery, and extended the battery to another small unit leader domain (platoon commander). The psychometric analysis concluded that the battery instruments are of good quality, are measuring discrete constructs, and have predictive validity and the ability to discriminate levels of performance. Additional products developed in this effort have also already shown utility for supporting small unit leader development and decision making research.

Moving Forward

Several aspects of the battery merit continued research and development to make the battery as robust and as useful to the Marine Corps as possible in order to help ensure the best possible

success for small unit leaders in the future. The following items summarize areas for continued research and development:

- Because a combined sample of Lts and NCOs was used in the factor analyses it is plausible that some instruments may have appeared to have more than one underlying latent construct because the factor structure of small unit decision-making ability is likely to differ across Lts and NCOs. Thus, one recommendation for future work is to ensure that in future development, a large enough sample size is available for both Lts and NCOs to examine the measurement equivalence/invariance of the factor structure across Lts and NCOs. If the factor structure is the same, then any differences in performance between Lts and NCOs can be more meaningfully interpreted. However, if the factor structure is not the same for Lts and NCOs, then comparisons in performance between Lts and NCOs could be ambiguous.
- It may be beneficial to conduct additional item-level analysis of the Situational Judgment Tests to identify whether items can be removed to improve reliability, and also conduct a factor analysis to identify the different constructs being measured by subsets of the items.
- All constructs in this effort should be retained for future study as the current analysis is limited by the use of a convenience sample rather than a sample of the target users and the only criterion variable available was the DRI. Further investigation could result in more meaningful relationships with the remaining constructs with a more relevant sample and an external criterion measure.
- Acceptable instruments were not found for four constructs initially considered as important aspects of small unit decision making. For adaptability, we recommend consideration of construction of a scenario-based instrument and consideration of the I-ADAPT instrument which has a 55 item format and evidence of predictive validity. (See Baard, Rench, and Kozlowski, 2014 for a review of the adaptation construct and the I-ADAPT instrument.) For the constructs of change detection and anomaly detection, we recommend adapting or devising a measure that fits our operational definitions as no useful instruments were found. Note that our definition involves military relevant performance and not basic perceptual skills. For the construct of cognitive flexibility we suggest the development of a performance instrument in line with our definition that requires the participant to transfer principles of performance from one scenario to another.
- We recommend that at the conclusion of the project, the final battery be implemented with a relatively large sample size from the desired populations—prospective and current maneuver squad leaders and platoon commanders—to establish a robust reference group for scoring and interpretation.
- Finally, to support use by the USMC, the battery should be converted into a computer-administered version following this project to mitigate test fatigue and cognitive overload by allowing the respondents to save their work, stop, and return when refreshed to a password protected assessment that must be completed within an adequate, designated amount of time from first login. The computer version should also automate scoring and interpretation. This migration to a computer-supported version should be followed by steps to transition that product to the Marine Corps including system demonstration, testing in an operational environment, and training for battery implementation.

References

Baard, S. K., Rench, T. A., Kozlowski, S. W. J. (2014). Performance adaptation: A theoretical integration and review. *Journal of Management, 40(1),* 48-99.

Brummel, B. J., Rupp, D. E., & Spain, S. M. (2009). Constructing parallel simulation exercises for assessment centers and other forms of behavioral assessment. *Personnel Psychology, 62*, 137-170.

Cronbach L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika* 16 (3): 297–334.

Desgrosseilliers, T. S., & Hoffman, R. (2014, March). The Basic School. *Marine Corps Gazette*, *98*(3), 8-20. Quantico, VA: Marine Corps Association and Foundation.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman and Hall.

Fletcher, J. D., & Wind, A. P. (2014). The evolving definition of cognitive readiness for military operations. In O'Neil, H. F., Perez, R. S, & Baker, E. L. (Eds.), *Teaching and measuring cognitive readiness*. New York: Springer.

Grier, R. A. (2012). Military cognitive readiness at the operational and strategic levels: A theoretical model for measurement development. *Journal of Cognitive Engineering and Decision Making***,** *6* (4), 358-392

King, B. M., Rosopa, P. J., & Minium, E. W. (2010). *Statistical reasoning in the behavioral sciences (6th ed.).* Hoboken, NJ: Wiley.

Klein, G.A. (1989). Recognition-primed decisions. In W.B. Rouse (Ed.), *Advances in man-machine systems research* (pp. 47–92). Greenwich, CT: JAI Press.

Morrison, J. E., & Fletcher, J. D. (2002). *Cognitive readiness*. (IDA Paper P-3735). Alexandria, VA: Institute for Defense Analysis.

Nunnally, J. C. (1978). *Psychometric theory (2nd ed.).* New York: McGraw-Hill.

O'Neil, H. F., Perez, R. S, & Baker, E. L. (Eds.) (2014). *Teaching and measuring cognitive readiness*. New York: Springer.

Rencher, A. C. (2002). *Methods of multivariate analysis* (2nd ed.). New York: Wiley.

Ross, K. G., Halterman, J. A., Pierce, L. G., & Ross, W. A. (1998) Preparing for the instructional technology gap – A constructivist approach. *Proceedings of the 1998 Interservice/Industry Training, Simulation, and Education Conference*, Orlando, FL

Ross, K. G., Phillips, J. K., & Rivera, I. D. (2013). *Marine Corps maneuver squad leader mastery model.* Cognitive Performance Group: Orlando, FL.

Ross, K. G., Phillips, J. K., Klein, G., & Cohn, J. (2005). *Creating expertise: A framework to guide technology-based training* (Final Technical Report for Contract #M67854-04-C-8035 for the Marine Corps Systems Command Program Manager for Training Systems, Orlando, FL). Fairborn, OH: Klein Associates.

Ross, K. G., and Pierce, L. G. (2000). Cognitive engineering of training for adaptive battlefield thinking. *Proceedings of the International Ergonomics Association 2000/Human Factors and Ergonomics Society 2000 Conference*, July 2000, San Diego, California.

U.S. Marine Corps [USMC]. (n.d.-a). *Marine Corps Vision & Strategy 2025*. Retrieved from http://www.onr.navy.mil/~/media/Files/About%20ONR/usmc_vision_strategy_2025_0809.ashx.

U.S. Marine Corps [USMC]. (n.d.-b). *Marine Corps Vision and Strategy 2025: Implementation Planning Guidance*. Quantico, VA: Author.

U.S. Marine Corps [USMC]. (2010). *The 35th Commandant of the Marine Corps Commandant Planning Guidance*. Washington, D. C.: U.S. Department of the Navy.

U.S. Marine Corps Training and Education Command [TECOM]. (2011). *U.S. Marine Corps small unit decision making workshop: Version 4.0* (Final Technical Report). Quantico, VA: Author.

Vogel-Walcutt, J.J., Ross, K.G., & Knarr, K.A. (2013). Small unit decision making assessment battery. *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference.* Orlando, FL: National Training Systems Association.

Vogel-Walcutt, J. J., Ross, K. G., Smith, K. M., & Brown, T. F. (2012, 4 October; Revised 14 December). *Preliminary assessment battery report: Literature review results*. Report submitted under TECOM contract M00264-12-P-0249.

Appendix A: Small Unit Decision Making Behaviorally Anchored Rating Scale

SUDM BEHAVIORALLY ANCHORED RATING SCALE

Comments

Character, Initiative, and Command Presence: The mental, physical, and character traits of an effective leader who demonstrates confidence, sets a positive example, garners respect and trust from his subordinates, takes full responsibility for his own actions, and accomplishes tasks and goals autonomously within intent.

Not observed:	✓	**1**	✓	**2**	✓	**3**	✓	**4**	✓	**5**	✓
Maturity and Command Presence		Demonstrates a lack of confidence in his ability or may be over-confident to make up for the lack of ability. Interacts with his subordinates as if they		Demonstrates an increased confidence in his abilities. Begins to separate himself professionally from the other squad members.		Consistently displays leadership skills. Confidence, not arrogance, is portrayed in his behavior and answers to questions. Behaves professionally.		Demonstrates command presence and a high level of confidence. Marines look up to him.		Sets leadership example for men; Marines want to emulate him, and they would follow him anywhere. Morals never waiver.	
Initiative		Waits for taskings instead of taking the initiative. Requires confirmation or assurance before he makes a decision.		Exhibits willingness to try things on his own and take charge. Shows more initiative, but his judgment has not kept pace; is over-confident when attempting to take initiative.		Starts taking initiative effectively for standard or routine tasks. Knows the tasks for which he is responsible within his squad, and doesn't need to be told what to do (e.g., inspections and counts).		Operates off intent without being being told how to accomplish the task or mission. Seeks guidance from Platoon Commander only when the situation is a potential ROE or international incident		Consistently and automatically takes initiative. Executes tasks with little to no guidance, even when faced with complexity or adversity.	
Leader Responsibili-ties and Interactions		Fails to take charge of whole squad. Often neglects to correct squad members' mistakes. Hesitant to take a stand on issues that are not clear cut. The squad is not responsive to him because he has not earned their trust.		Competent in his duties, but closely monitors and reigns in his team leaders (i.e., doesn't give them lattitude to make mistakes).		Effectively counsels Marines. Marines seek him for assistance. Leadership in the field is stronger than in garrison because he sees the role differently in different contexts.		Respectful towards his Marines but comfortable reprimanding them. Provides criticism tactfully. Gives credit to others when appropriate.		Stands up for the welfare of his Marines in all contexts. Voices opinions professionally and freely, even in the company of senior Marines.	

Tactical Skills/Tactical Thinking: The cognition required to apply tactical, technical, and team knowledge to analyze mission requirements, plan, solve tactical problems, and execute the mission decisively, within the big picture and Commander's intent.

Not observed:	✓	1	✓	2	✓	3	✓	4	✓	5	✓
Understand-ing the Situation		Does not readily notice changes (i.e., anomalies) in the environment. Demonstrates difficulty in filtering information and determing what is relevant.		Can identify anomalies but doesn't grasp the meaning. Exhibits an understanding of the capabilities of enemy weapons and assets but unable to think through what enemy is going to do with them.		Understands what indicators and changes in the environment mean, but cannot make adjustments in time to be useful. Spends a lot of time dissecting the situation to make sure it makes sense to him.		Notices anomalies in the environment and can make changes in time to react to their meaning. Conducts in-depth analysis of the enemy even under stress to understand what enemy is doing.		Recognizes subtle changes or anomalies and knows automatically how to respond. Maintains constant awareness and anticipates problems. Manages large amounts of information effectively.	
Planning		Knows basics of planning but relies heavily on guidance and prompts to plan. Has difficulty connecting all the pieces into a coherent picture during mission analysis. Does not consider advantages and disadvantages associated with employing squad.		Demonstrates a basic understanding of the planning process and can plan with some oversight. Understands what his assets are but not how to employ them. Fails to plan for contingencies.		Able to plan independently. Understands how to operationalize Commander's intent. Thinks ahead to what will be required to complete mission. Plans for standard or common contingencies.		Can think of several COAs and match the best one for minimal loss of life and damage based on his experience. Able to identify a problem, come up with a course of action, and task Marines to execute it. Fully comprehends second and third order consequences.		Operates effectively with broad intent. Critiques COAs and makes suggestions to address gaps. Anticipates issues, even uncommon ones. Plans include contingencies for standard and non-standard events that reflect an integrated METT-TC analysis.	
Executing the Mission		Competent in performing small scale fire team level tactics. Focuses on fire team he is with but fails to maximize use of entire squad. Exhibits tunnel vision by being drawn into his weapon and shooting targets rather than directing the squad. More likely to respond with emotion (e.g., anger) than rational thought, resulting in vulnerability. Does not consider range of responses based on situation; instead leans towards kinetic actions.		Exhibits basic understanding of tactical employment of weapons and people. May fail to think the problem through initially and be forced to adjust placement of assets. Improves at employing the squad, but still gets drawn personally into fighting the fight and distracted from leading his Marines. Focuses on squad mission, not big picture.		Knows and employs his assets correctly. Employs tactical principles effectively in his operations. Effectively combines weapons systems and Marines to maneuver, and can reorganize squad to be more effective. Knows what to do to carry out a plan but not how to implement it smoothly and accurately. Completes mission with little to no guidance.		Understands how to employ a squad against different situational problems. Understands how the mission set impacts how he operates. Understands and uses combined arms effectively, but with coordination assistance. Maneuvers forces smoothly and effectively. Accurately anticipates what adjacent units will do.		Fluidly and effectively conducts combined arms missions without assistance. Uses all available assets appropriately. Stays a step ahead of the enemy. Maintains offensive mindset so that his actions, not the enemy's, dictate the situation.	

Tactical Skills/Tactical Thinking: The cognition required to apply tactical, technical, and team knowledge to analyze mission requirements, plan, solve tactical problems, and execute the mission decisively, within the big picture and Commander's intent.

Not observed:	✓	1	✓	2	✓	3	✓	4	✓	5	✓
Decision Making		Uncertain or afraid to make decisions without approval of guidance from Higher. Completes tactical tasks but continually asks Higher for direction.		Hesitates on his decisions because he doesn't want to make a mistake. Loses some momentum as a result of second guessing himself. Considers consequences in his problem solving activities, but is slow to make a decision.		Generates acceptable, but not optimal, solutions to problems under pressure. Confident in his decision making and does not require constant approval from Higher. Decisive and quick under pressure, but reactive to the situation.		Quickly and effectively assesses situation and makes decisions. Takes decisive actions and understands second and third order implications of decisions.		Proactively identifies and solves problems, regardless of the task, quickly and intuitively. Decides based on enemy's projected actions and capabilities. Makes timely and effective decisions that work without fear of failure.	
Applying Knowledge		Competent in domain knowledge but requires guidance to think through a solution in order to apply knowledge. Requires direct taskings and supervision. Follows procedures and processes but does not understand the "why" of what he is doing.		Identifies the problem and applies the textbook answer, but is also beginning to come up with original solutions. Begins to rely more on SOPs and past experience, but not analytical thinking.		Starts thinking "outside-the-box" for planning and problem solving. Proficient at assessing terrain based on more training, experience, and knowledge of what to look for in the environment.		Demostrates mastery in coordinating use of non-squad assets. Terrain assessment is ongoing and integrated with overall situational awarness based on experience in order to meet mission objectives.		Applies knowledge and principles fluidly across contexts. Mastery of the domain supports effective improvisation and problem solving.	

Adaptability/Flexibility: The ability to fluidly apply knowledge and tactical principles across situations, or alter ones' plans, actions, or decisions when the situation, environment, or circumstances has changed, while still accomplishing the mission or intent.

	✓	1	✓	2	✓	3	✓	4	✓	5	✓
Recognize Need to Adapt		Fails to detect things in the environment that are out of place (changes and anomalies). Driven by the plan and does not see reason to or is hesitant to adapt.		Starts to recognize the need to adapt the plan. Begins to notice things in the environment that are out of place but doesn't know what to do with that information.		Develops and applies contingencies during planning because he knows the squad will have to adapt.		Anticipates there will always be adaptations to the plan during execution based on unexpected events. Attributes meaning to changes and knows which changes and anomalies are important.		Anticipates changes before action is required. Visualizes what to do continually as the situation unfolds. Immediately identifies and understand the meaning of anomalies and changes in the environment.	
Ability to Adapt		Fears that changes will lead to mistakes. Implements TTP and mission tasking without adapting or tailoring it to the situation.		Cannot come up with a new decision fast enough to respond to changes. Cannot analyze what needs to be done to support changes. Reacts to the situation without a sound rationale for his actions.		Able to identify a problem and starts figuring out how to employ squad in response. Starts to consider 2nd and 3rd order effects of changes he may make.		Proven to quickly adapt to any situation. Effectively uses new information from the situation to make a decision. Consistently adjusts well due to squad's level of training.		Maintains a backup plan in hip pocket, or can quickly generate a new plan. Transitions seamlessly from the primary plan to a contingency plan. Displays seamless decision making and action in the face of change.	

Administration: The coordination and supervision of people, processes, and equipment in conjunction with the abilities to multitask and delegate assignments.

Not observed:	✓	1	✓	2	✓	3	✓	4	✓	5	✓
Multitasking and Workload Management		Demonstrates inability to prioritize tasks; treats tasks as having equal weight and time pressures. Exhibits difficulty multitasking.		Shows initial signs of being able to priortize and organize his efforts to meet several task requirements.		Analytically organizes his time to accomplish tasks. Meets goals within the time allotted.		Plans squad events with appropriate consideration of time requirements. Manages time proactively by anticipating tasks that will need to be completed and by using short windows of time to work on upcoming tasks like paperwork.		Consistently and automatically considers time management, personnel management, and equipment as they relate to the mission.	
Monitoring and Supervising		Monitors progress inefficiently due to focus at too fine a level of detail. Demonstrates inflexible management style (e.g., cannot manage more than once personality type).		Demonstrates initial signs of focusing attention on and directing his three team leaders vice all 12 Marines. Fails to supervise delegated tasks to ensure task completion.		Skillfully tasks and manages three team leaders rather than 12 Marines. Effectively assesses the skills of his individual subordinates to understand how to best use their strengths and compensate for their weaknesses.		Matches individual Marines to teams based on their skills and personalities. Manages a reinforced squad well.		Demonstrates willingness to work alongside others to assess, motivate, and teach rather than just direct actions. Effectively controls fires, maneuver, and movement.	
Administration and Organization		Competent in accomplishing basic administrative duties associated with tracking Marines, materials, and equipment. Fails to execute timing as planned.		Understands requirements for mission preparation. Maintains routine tasks effectively.		Effectively accomplishes garrison tasks for which he is accountable. Effectively tracks weapons and people.		Steps back to see big picture and employs squad to succeed in the situation.		Systematically approaches tasks. Effectively applies job aids (e.g., laminated spreadsheets to keep track of weapons).	
Delegation and Direction		Exhibits difficulty delegating to team leaders to accomplish tasks required of squad. Attempts to conduct or contribute to all squad tasks himself, including garrison adminstrative tasks and combat fighting tasks.		Conceptually understands the role of a manager. Leads more efficiently, but performance is variable and choppy. Functions without micromanagement to achieve procedural tasks but needs support for complex tasks.		Effectively assesses the skills of his individual subordinates to design tasks and delegate authority regardless of rank. Understands he's delegating authority, but not ultimate responsibility.		Delegates to team leaders immediately without hesitation or second guessing.		Demonstrates comfort tasking Platoon Sergeant when he is with the patrol. Fights the squad as a weapon system; directs the squad rather than firing his own weapon.	

Communication: Effectively obtaining, relaying, and explaining information to subordinates, superiors, and adjacent squad or section leaders in order to direct actions or maintain shared understanding.

Not observed:	✓	1	✓	2	✓	3	✓	4	✓	5	✓
Knowing When and What to Communicate		Does not prioritize what to report to Higher. Fails to report important events or information to Higher and/or reports too often (such as when trivial events occur). Fails to brief back or seek clarification to ensure understanding.		Relies on Platoon Commander for direction of tasking and next steps. Often appears too overwhelmed to contact Higher when the situation is going badly.		Fully understands the common types of information Higher wants to know and reports accordingly. Provides situation reports to the Platoon Commander without prompting.		Comprehends that one objective of communication to Higher is to ensure the squad receives support if needed. Communicates with Platoon Commander only when the information is significant (e.g., assets he needs or situations that have strategic implications).		Provides timely and relevant information to Higher (i.e., knows what is significant and how urgent the information is). Selects times to communicate to Higher that enable him to focus attention on the squad fight when needed, yet ensure Higher receives needed information.	
Understanding What is Communicated		Difficulty understanding all that is communicated to him (e.g., may miss key details or fill in gaps in understanding with erroneous information). Mistakes the intent of Higher by conducting tasks in the wrong sequence or doing an additional but undesired task.		Can repeat to his Marines what Platoon Commander has communicated, but still needs to go back to him for guidance.		Understands and effectively applies direction from the Platoon Commander or Platoon Sergeant. Understands the value of "cross-talk" among different people to increase understanding of ideas.		Comprehends the rationale behind the Platoon Commander's direction or tasking in the context of the situational demands. Asks questions about appropraite responses to specific contingencies to confirm his understanding of the Platoon Commander's intent.		Comprehends what is communicated, and integrates and compares it to his own fully developed mental model of the situation to deepen his understanding.	
Communicating to Others		Uses negative communication with subordinates in the form of swearing and yelling as part of his direction and discipline to his Marines. Not able to clearly and concisely articulate information. Fails to report an accurate and complete status; may be unaware of situational requirements, or may be trying to hide squad's deficiencies.		Clearly conveys knowledge to Marines. Starting to understand the perspective of other Marines. No longer uses the negative communication.		Begins to adapt communication for different individuals. Clearly tasks team leader and delegates actions to them. Explains the situation accurately.		Provides guidance that is easily understood by his Marines. Understands a variety of people and how to adapt his communication style to different people. Knows what people are thinking before, during, and after an event as a means of shaping his communication.		Paints a picture of the battlefield that enables others to visualize the situation. Gives clear intent and task statement that team leaders can respond to effectively. Refrains from swearing in order to demonstrate professionalism.	

Job Knowledge: The comprehension of procedures, processes, and asset capabilities required to effectively perform the maneuver squad leader role.

Not observed:	✓	1	✓	2	✓	3	✓	4	✓	5	✓
Squad Leader and Responsibil-ities		Knows generally what is expected of him as a Squad Leader but doesn't have a full grasp of job and responsibilities; needs guidance from Higher. Organizes and carries out low level tasks independently. Applies the exact textbook solution (i.e., rule-based).		Understands the job requirements better, but hesitant to make decisions without consent from Higher. Tries to apply knowledge beyond simple rule-based application, but is easily overwhelmed by situational elements in context. References checklists during performance.		Understands doctrine and begins to think about what it means in application. Understands checklists and references them after the fact as a check on performance.		Fully understands job and responsibilities and how to organize tasks, but may still analyze actions in a deliberate manner. Able to act almost without the aid of reference such as in conducting reporting.		Knows domain and job very well (encyclopedic knowledge) and applies it based on an overall grasp of each situation to organize and carry out tasks. Sees solutions quickly and easily in a range of situations.	
Assets and Their Employment		Able to employ all individual squad weapons, but cannot effectively employ all together across three fire teams (e.g., needs experienced Squad Leader supervision or specific orders).		Understands the capabilities and limitations of squad weapons systems, but has difficulty applying the knowledge in the context of a situation.		Performance in applying squad weapons is effective and more natural, based on greater knowledge of how they operate. Understands supporting arms and has knowledge of company and battalion weapon systems, but effective employment requires guidance.		Effectively plans for supporting arms. Demonstrates mastery level understanding of employment of all squad weapons. Orchestrates combined arms missions with coordination assistance.		Effectively employs indirect fire and aviation assets as supporting arms. Teaches others about non-organic assets and how to use them.	
Procedures and Technical Infantry Skills		Has basic infantry skills and MOS knowledge. Can recite procedures for call for fire and close air support but not use them. Requires detailed guidance to complete standard reports.		Knows call for fire and close air support procedures and assets that can be used for support, but does not use them appropriately given a mission context. Goes to the field with checklists in hand.		Conducts key Squad Leader tasks with technical proficiency. Demonstrates weaknesses in some areas (i.e., has not mastered all technical skills equally well). Uses notecards or job aids intermittently to ensure procedures are carried out correctly.		Demonstrates technical and tactical proficiency in his MOS. Proficient with the Orders process.		Anticipates what he needs to have ready. In planning, uses five paragraph order format and addresses each element appropriately. Develops templates and laminates for his Marines' ease and effectiveness (e.g., call for fire or casualty evaution).	

Self-Control and Stress Management: Managing and regulating ones' emotional responses, control, and stability in order to prioritize and perform effectively within high stress contexts.

Not observed:	✓	1	✓	2	✓	3	✓	4	✓	5	✓
Self-Control and Stress Management		Freezes up in a stressful and/or challenging situation; may stop performing. Unable to think clearly under stress. Occassionally reports conflicting information and issues commands that are irrational and illogical. Employs squad ineffectively during stressful situations.		Recognizes he is overwhelmed and must take measures to manage his stress. Is challenged to manage stress because he is unable to priortize tasks. Exhibits anxiety but not panic.		Able to calm himself after an initial reaction of fear or worry. Able to mask underlying anxiety with a calm exterior. Comfortable not knowing when the mission will end; continues to move forward despite uncertainty.		Calmly maintains his position of authority under stress. Demonstrates an even keel and emotional stability.		Performance does not degrade under stress, chaos, or time pressure. Demonstrates logical thinking in the middle of a firefight; thinks on his feet and continues to orchestrate the squad. Diverts attention to assist his Marines with coping under stressful situations; uses humor or assurances to calm them.	

Self-Development: The motivation to continuously acquire and apply new knowledge, skills, and lessons learned to current role requirements and future professional development goals, as a result of an attentiveness to the nature of one's self, personal strengths, limitations, and work styles.

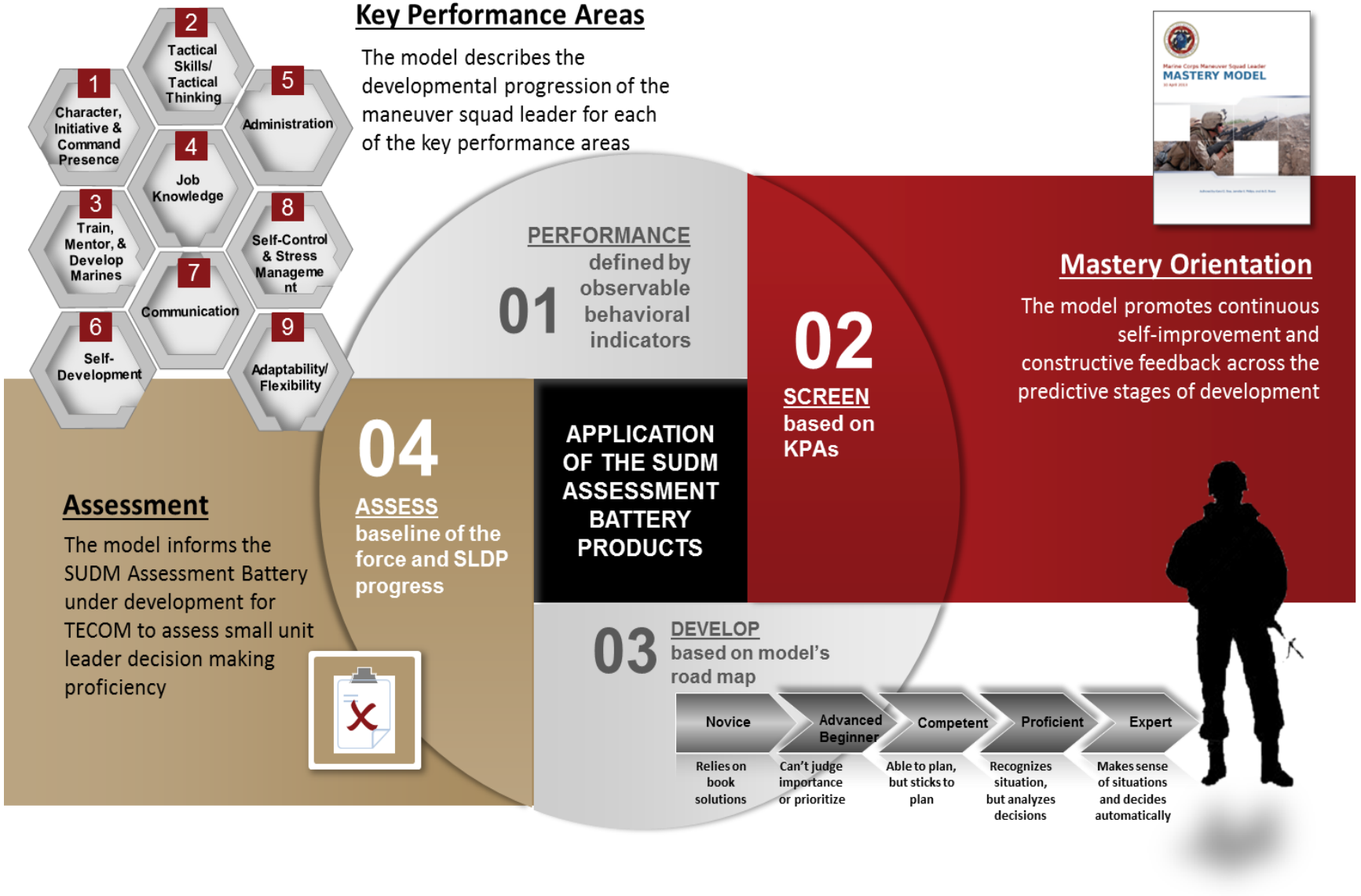
Not observed:	✓	1	✓	2	✓	3	✓	4	✓	5	✓
Self-Awareness		Seeks to understand concepts and learn to achieve goals, big and small. Requires a coach or mentor to tell him what he needs to work on and how. Does not yet know his physical or mental potential. Does not think outside of his job role.		Makes initial self-assessments. Uses standards or feedback to figure out if he is doing new skills the right way (i.e., concerned with being correct as his standard). Shows improvement on his limitations.		Recognizes and comprehends the implications of his gaps in knowledge and understanding. Focuses on developing each skill as a Squad Leader; follows the T&R Manual and Company Commander guidance to understand what is expected of Squad Leaders.		Recognizes his weaknesses and focuses on improvement in those areas. Shows awareness of his limitations.		Demonstrates awareness of most or all of his limitations. Embraces self-improvement by continuously seeking out a variety of learning opportunities. Purposefully interacts with a larger community to learn from other perspective and experiences.	
Strategies and Capabilities for Knowledge Acquisition and Skill Development		Requires a long time to think things through and accomplish new tasks. Unaware of the full range of resources to whom he can take his questions. Unfamiliar with strategies to obtain information or knowledge.		Able to ask good questions on the spot. Engages in self-study of tactical knowledge, weapons knowledge, and how to apply that knowledge. Reads books about military tactics.		Accepts criticisms well and constructively, and learns from it. Knows where to find information and how to seek out knowledge. Collects information proactively to expand his knowledge.		Maintains awareness of his limitations after performance and feedback as a means for improvement. Actively seeks knowledge through courses, peer interactions, and Marines outside his unit or MOS.		Keeps up with new domain knowledge to constantly better himself. Reads constantly; current events, pubs/doctrine, military history. Asks fewer questions and figures things out for himself.	

Train, Mentor, and Develop Marines: Continously caring about and fostering the professional and personal development of subordinates by teaching, training, coaching, building trust, assessing skills and personalities, and providing guidance.

Not observed:	✓	1	✓	2	✓	3	✓	4	✓	5	✓
Focus and Motivation		Focuses on his squad and how its performance looks to others. Not focused on developing each individual. Unaware of the capabilities of his Marines. Uses a "check-in-the-box" system and is less concerned about the proficiency of his squad.		Understands that looking out for others is key to his role. Begins to demonstrate an outward focus, from thinking of himself to now thinking of his Marines. Begins to put himself in other Marines' shoes.		Identifies and understands the individual strengths and weaknesses of each Marine; focuses on captializing on the strengths of each Marine. Comfortable with the abilities of his personnel and does not view Marines as interchangeable.		Takes on more a mentorship role not just a trainer. Squad members trust him and his expertise.		Leads by example all the time. Focuses on building squad's confidence. Uses his own free time to train his Marines. Can separate himself from his work life and talk to Marines as men/friends in an appropriate manner.	
Strategies and Methods		Lacks the ability to direct and develop his squad. Directs Marines to complete their jobs but without insuring proper training or in-process guidance. Waits for company to tell him what to train.		Starts to train basic skills to his Marines. Shows few ideas for how to carry out training. Keeps tasks for Marines simple. Conducts simple, low-detail AARs.		Teaches skills he knows, but can't explain the rationale behind the task (i.e., why something is important or needed). Breaks down tasks and communicates them in a way his Marines will understand. Knows which tasks are relevant to a mission and practices them to proficiency.		Provides rationale for tasks. Conducts better AARs that focus on why things happened, what caused events, steps they can take to be proactive or prevent; asks others what they think. Training style improves because he caters training to different learning styles. Trains to skill gaps he identifies in his Marines.		Optimizes time by never missing an opportunity to teach his Marines. Shows initiative to prepare classes/training on training requirements he has identified. Interacts with peers and seeks out how they handle different situations to then use each other's ideas for training.	

Additional Comments:	

Appendix B: Application Opportunities for the Maneuver Squad Leader Mastery Model


Key Performance Areas
The model describes the developmental progression of the maneuver squad leader for each of the key performance areas
1 Character, Initiative & Command Presence
2 Tactical Skills/ Tactical Thinking
3 Train, Mentor, & Develop Marines
4 Job Knowledge
5 Administration
6 Self-Development
7 Communication
8 Self-Control & Stress Management
9 Adaptability/ Flexibility
Marine Corps Maneuver Squad Leader
MASTERY MODEL
PERFORMANCE defined by observable behavioral indicators
01
02
SCREEN based on KPAs
Mastery Orientation
The model promotes continuous self-improvement and constructive feedback across the predictive stages of development
APPLICATION OF THE SUDM ASSESSMENT BATTERY PRODUCTS
04
ASSESS baseline of the force and SLDP progress
Assessment
The model informs the SUDM Assessment Battery under development for TECOM to assess small unit leader decision making proficiency
03
DEVELOP based on model's road map
Novice
Advanced Beginner
Competent
Proficient
Expert
Relies on book solutions
Can't judge importance or prioritize
Able to plan, but sticks to plan
Recognizes situation, but analyzes decisions
Makes sense of situations and decides automatically

Appendix C: Items Removed from Each Version of the Decision Requirements Interview

Bangui Offensive				
Item #	**Percent Responded**	**Item Description**	**Subscale**	**Score**
Item2A7	0.4%	Notes that indirect fire support (artillery) will take too long to be effective.	Asset Employment	5
Item4C2	8%	Employs a squad member to communicate with Higher.	Asset Employment	3
Item2C16	5%	Changes his mind about what COA he will pursue.	Recognition	-3
		Recognition Scale (Remove all)		15
Item1C3	0%	Identifies need to engage or choose to bypass enemy based on higher's mission	Second and Third Order Effects	5

**Item highlighted in green refers to items that should be further evaluated in the future with a different population. These do not count towards the final score.*

Jafarani Patrol				
Item #	**% achieved**	**Item Description**	**Subscale**	**Score**
Item4A3	0%	Notes this is poor time of day for going into the households of the locals.	First Order Effects	5
Item4A19	0%	Is aware of team's potential to be angry.	First Order Effects	3
Item4A18	0%	Is aware of own feeling of anger for not getting information from Omar.	First Order Effects	3
Item4A27	0%	Changes his mind about his assessment of Omar.	Recognition	-3
Item5A11	0%	Changes mind about situational assessment.	Cue Recognition	-3
Item4A4	0.4%	Notes that the term "insurgent" would not be used during the conversation with Omar.	Perspective	3
Item2A12	0.4%	Changes his mind about whether the farmers are a threat.	Recognition	-3
Item2B17	1%	Changes mind about how to deal with the farmers.	Recognition	-3
Item2A1	0.8%	Notes that he wouldn't be patrolling along roads.	First Order Effects	3
Item1C10	1%	Changes mind about how to conduct actions at Omar's house.	Recognition	-3
Item4A22	1.6%	Is aware of limited egress opportunity.	First Order Effects	3
Item3A8	3%	Changes mind about the threat or hostility associated with the man on the roof.	Recognition	-3
Item1C1	5%	Considers the time of day that patrol will	First Order	5

		arrive in village; knows that mid-day is a bad time to talk to HVI.	Effects	
Item1A9	3%	Changes mind about route selection after initial plan	Recognition	-3
Item5B21	3%	Changes mind about course of action.	Recognition	-3
Item5B3	6.5%	Approaches the vehicle to collect more information.	First Order Effects	-1
Item3A9	4%	Notes questionability of LCpl Richards' judgment.	Cue Recognition	3
Item5B7	66%	Reports the vehicle to higher.	First Order Effects	1
Item5A7	77%	States an assessment that the vehicle is possibly an IED.	Cue Recognition	1
Item5A3	85%	Notes as an indicator the positioning of the car to channelize the patrol.	Cue Recognition	1
Item4A10	79%	Notes Omar’s body language (e.g., crossed arms, distracted) as an indicator.	Cue Recognition	1
		Recognition Scale (remove all)		

Appendix D: Demographics of the Group Used to Establish the Reference Group

Rank	Years in Service	# of Deployments	Tasks in Deployments
NCO	9.0	4	Conducted contact and security patrols, cross trained to machine gun section, designated marksman (DM).
NCO	6.0	1	Part of a joint operation with Jordanian command. Tasked with teaching proper radio procedures before, during, and after the patrols.
NCO	6.0	2	Conducted contact and security patrols as well as Quick Reaction Force. Helped Afghan police during patrol.
NCO	12.0	6	Conducted route clearance, security patrols, and cache searches. Operated IDD DOG for IED searches.
Lt	7.5	3	Enlisted Intelligence Analysts. Participated in Counter-Maritime Drug Interdiction. Electronic Warfare Supervisor and part of the carrier strike group.
NCO	5.0	2	Conducted security patrols in Afghanistan. Patrols consisted of engaging with locals and searching for weapons caches.
NCO	5.0	1	Driver for LEAO vehicle sweep team, swept for IEO, and part of a build team that built bridges. Conducted security and clearance patrols.
NCO	7.0	4	Conducted census and security patrols. Trained Afghan soldiers.
NCO	6.0	1	Information not available
NCO	8.0	2	Conducted foot patrols, census operations, clearing operations, mobile patrolling, and COIN interdiction operations. Participated in COIN raids.
NCO	5.0	2	SAW gunner and lead driver for VC
NCO	7.0	1	Information not available
NCO	6.5	1	Conducted security patrols, served as turret gunner and as artillery (howitzer)
NCO	11.0	3	Conducted combat security patrols, reconnaissance, and zone reconnaissance.
NCO	5.0	2	Information not available
NCO	7.0	2	EOD escort security patrols. Conducted key leader engagements.
NCO	5.0	2	Conducted dismounted patrols in Afghanistan farm fields.
NCO	10.0	3	Information not available

NCO	12.0	4	Conducted clearance security patrols. Served as admin logistics and tactical advisor.
NCO	5.0	2	Conducted security patrols for the purpose of peace-keeping.
NCO	5.0	2	Information not available
NCO	5.0	2	Information not available