NORTH ATLANTIC TREATY ORGANIZATION

SCIENCE AND TECHNOLOGY ORGANIZATION

AC/323(SAS-110)TP/899

www.sto.nato.int

STO TECHNICAL REPORT **TR-SAS-110**

Operations Assessment in Complex Environments: Theory and Practice

(Évaluation des opérations dans les environnements complexes : théorie et pratique)

Final Report of RTG SAS-110.

Published December 2019

Distribution and Availability on Back Cover

NORTH ATLANTIC TREATY ORGANIZATION

SCIENCE AND TECHNOLOGY ORGANIZATION

AC/323(SAS-110)TP/899

www.sto.nato.int

STO TECHNICAL REPORT **TR-SAS-110**

Operations Assessment in Complex Environments: Theory and Practice

(Évaluation des opérations dans les environnements complexes : théorie et pratique)

Final report of RTG SAS-110.

Senior Editor, Dr. Adam Shilling.

The NATO Science and Technology Organization

Science & Technology (S&T) in the NATO context is defined as the selective and rigorous generation and application of state-of-the-art, validated knowledge for defence and security purposes. S&T activities embrace scientific research, technology development, transition, application and field-testing, experimentation and a range of related scientific activities that include systems engineering, operational research and analysis, synthesis, integration and validation of knowledge derived through the scientific method.

In NATO, S&T is addressed using different business models, namely a collaborative business model where NATO provides a forum where NATO Nations and partner Nations elect to use their national resources to define, conduct and promote cooperative research and information exchange, and secondly an in-house delivery business model where S&T activities are conducted in a NATO dedicated executive body, having its own personnel, capabilities and infrastructure.

The mission of the NATO Science & Technology Organization (STO) is to help position the Nations' and NATO's S&T investments as a strategic enabler of the knowledge and technology advantage for the defence and security posture of NATO Nations and partner Nations, by conducting and promoting S&T activities that augment and leverage the capabilities and programmes of the Alliance, of the NATO Nations and the partner Nations, in support of NATO's objectives, and contributing to NATO's ability to enable and influence security and defence related capability development and threat mitigation in NATO Nations and partner Nations, in accordance with NATO policies.

The total spectrum of this collaborative effort is addressed by six Technical Panels who manage a wide range of scientific research activities, a Group specialising in modelling and simulation, plus a Committee dedicated to supporting the information management needs of the organization.

- AVT Applied Vehicle Technology Panel
- HFM Human Factors and Medicine Panel
- IST Information Systems Technology Panel
- NMSG NATO Modelling and Simulation Group
- SAS System Analysis and Studies Panel
- SCI Systems Concepts and Integration Panel
- SET Sensors and Electronics Technology Panel

These Panels and Group are the power-house of the collaborative model and are made up of national representatives as well as recognised world-class scientists, engineers and information specialists. In addition to providing critical technical oversight, they also provide a communication link to military users and other NATO bodies.

The scientific and technological work is carried out by Technical Teams, created under one or more of these eight bodies, for specific research activities which have a defined duration. These research activities can take a variety of forms, including Task Groups, Workshops, Symposia, Specialists' Meetings, Lecture Series and Technical Courses.

The content of this publication has been reproduced directly from material supplied by STO or the authors.

Published December 2019

Copyright © STO/NATO 2019
All Rights Reserved

ISBN 978-92-837-2224-3

Single copies of this publication or of a part of it may be made for individual use only by those organisations or individuals in NATO Nations defined by the limitation notice printed on the front cover. The approval of the STO Information Management Systems Branch is required for more than one copy to be made or an extract included in another publication. Requests to do so should be sent to the address on the back cover.

Table of Contents

List of Figures

List of Tables

List of Acronyms

ACLED	Armed Conflict Location and Event Data Project
AOR	Area of Responsibility
APAN (EAMPG)	All Partners Action Network (East Africa Multilateral Planning Group)
CIDNE	Combined Information Data Network Exchange
CJ-2	Intelligence Section
CJ-3	Operations Section
G-TSCMIS	Global Theatre Security Management Information System
KLE	Key Leader Engagements
LOE	Line of Effort
MCC	Military Coordination Cell
OHASIS	Overseas Humanitarian Assistance Shared Information System

SAS-110 Membership List

CHAIR

Dr. Adam SHILLING
US Army
UNITED STATES
Email: adam.p.shilling.civ@mail.mil

MEMBERS

Dr. Ben CONNABLE
RAND Corporation
UNITED STATES
Email: connable@rand.org

LTC Vaughn DELONG
Supreme Allied Commander Transformation
UNITED STATES
Email: Vaughn.delong@act.nato.int

Dr. Karsten ENGELMANN
US Army
UNITED STATES
Email: Karsten.g.engelmann.civ@mail.mil

Mr. Jan FRELIN
Swedish Defence Research Agency (FOI)
SWEDEN
Email: jan.frelin@foi.se

Ms. Danielle FENNING
Allied Command Transformation (ACT)
UNITED KINDOM
Email: Danielle.Fenning@act.nato.int

Lt. Col. Marko GANGI COE
Command Control (C2)
GERMANY
Email: Marko.Gangi@c2coe.org

Dr. Anton MINKOV
Defence Research & Development-CORA
CANADA
Email: Anton.minkov@forces.gc.ca

Maj Olli-Pekka PAJU
Finnish Defence Forces, Army Research
FINLAND
Email: Olli-pekka.paju@mi.fi

Maj. Carsten SCHULZE
Bundeswehr
GERMANY
Email: Carsten1schulze@bundeswehr.org

ADDITIONAL CONTRIBUTORS

Alexander William BEADLE
Norwegian Defence Research Establishment (FFI)
NORWAY
Email: Alexander.beadle@ffi.no

LTC Vincent BONCICH
US ARMY
UNITED STATES
Email: Vincent.j.boncich.mil@mail.mil

Robert L. BOVEY
Institute of Defense Analyses (IDA)
UNITED STATES
Email: Rbovey@ida.org

LTC Natalie CASEY
US ARMY
UNITED STATES
Email: Natalie.k.casey.mil@mail.mil

Kathleen M. CONLEY
Institute of Defense Analyses (IDA)
UNITED STATES
Email: kconley@ida.org

Paul DICKSON
Canadian Joint Warfare Centre
CANADA
Email: *Email Unavailable*

Aletta EIKELBOOM
Netherlands Organisation for Applied Scientific Research (TNO)
NETHERLANDS
Email: Aletta.eikelboom@tno.nl

Philip ELES
NATO Communications and Information Agency
CANADA
Email: Philip.eles@ncia.nato.int

M. Anthony FAINBERG (ret.)
Institute of Defense Analyses (IDA)
UNITED STATES
Email: *Email Unavailable*

R.G.W. GOUWELEEUW
Netherlands Organisation for Applied Scientific Research (TNO)
NETHERLANDS
Email: Rudi.gouweleeuw@tno.nl

Chris JORDAN
Niteworks
UNITED KINGDOM
Email: Chris.jordan@niteworks.net

LTC Kevin LARRABEE
US ARMY
UNITED STATES
Email: Kevin.s.larrabee.mil@mail.mil

Mark LENO
US Army War College
UNITED STATES
Email: Mark.a.leno2.civ@mail.mil

Guro LIEN
Norwegian Defence Research Establishment (FFI)
NORWAY
Email: Guro.lien@ffi.no

Elin MARTHINUSSEN GUSTAVSEN
Norwegian Defence Research Establishment (FFI)
NORWAY
Email: Marthinussen.gustavsen@ffi.no

Sergio MILLER
Defence Consultant
UNITED KINGDOM
Email: SNMILLER1@qinetiq.com

Jim NORTH
US ARMY
UNITED STATES
Email: James.north.ctr@usmc.mil

Joseph P. NOWAK
Combined Joint Task Force
UNITED STATES
Email: Joseph.p.nowak2.civ@mail.mil

Bruce PENNELL
NATO Communications and Information Agency
UNITED KINGDOM
Email: Bruce.pennell@ncia.nato.int

Valentin POPONETE
SHAPE Strategic Planning
ROMANIA
Email: Valentin.Poponette@shape.nato.int

Geert ROSEBOOM
Netherlands Organisation for Applied Scientific Research (TNO)
NETHERLANDS
Email: Geert.roseboom@tno.nl

Jeffrey SCHWERZEL
Royal Netherlands Army
NETHERLANDS
Email: J.Schwerzel@mindef.nl

LTC Andrew Swedberg
US ARMY
UNITED STATES
Email: Andrew.d.swedberg.mil@mail.mil

Sarah THAMBIDURAI
US ARMY
UNITED STATES
Email: sarah.thambidurai@dhs.gov

James S. THOMASON
Institute of Defense Analyses (IDA)
UNITED STATES
Email: Jthomaso@ida.org

Mark E. TILLMAN
Institute of Defense Analyses (IDA)
UNITED STATES
Email: Mtillman@ida.org

Karla WAYMAN
US ARMY
UNITED STATES
Email: *Email Unavailable*

Operations Assessment in Complex Environments: Theory and Practice

(STO-TR-SAS-110)

Executive Summary

The complexity of the environments in which NATO and its member nations conduct military operations is increasing. The rise of powerful non-state actors and the increasing likelihood of conflict taking place among civilian populations makes operations more difficult and the progress of operations toward their objectives more difficult to perceive. Hence, the increased attention we are paying to operations assessment, which is our attempt to determine the effectiveness of operations in these complex environments.

The study underlying this report is a symptom of the concern among members of the assessment community about their ability to provide useful, timely, and empirically true assessment findings, first, to support decision making in pursuit of more effective operations, but also, to account for the resources that national governments or higher military authorities provide to subordinate military organisations. The Operations Research and Analysis community, which is the *de facto* assessment community, thought they could be more effective at designing and leading assessment processes, and providing assessment and assessment products to the commanders and organisations they support.

The key finding of this activity is that a clear, explicit focus on improving the effectiveness of the organisation performing the assessment of its activities, via a focus on accounting for resources, improves the quality of the assessment findings and products. These, in turn, lead to more effective operations.

In the past, assessment processes often focused on providing accountability of the resources consumed during operations to higher military or political authority in accordance with a set of legal or regulatory requirements. This meant that an assessment product was a bureaucratic requirement, required by higher authority, produced by a group of assessors whose job it was to write this report. Staff, and frequently command, focused their attention on “real work” dealing with actual operations – rather than bureaucratic requirements – while the assessors laboured alone, without the expertise of other staff members, to provide the accountability higher authority required. This also meant that assessment efforts focused on the appearance of the report more so than on the findings the report communicated and the effectiveness gains that could be realised if the findings were acted upon. The report also had the character of a history report that looked backward over the preceding period. Recommendations contained within were almost an afterthought.

In contrast, an assessment process focused on finding ways to be more effective is of most value to the organisation’s commander, and is no longer a bureaucratic requirement. A commander is now greatly interested in assessment findings, rather than the report’s appearance, and the staff responds by providing the expertise that individual assessors lack, which makes the product much better. The report, properly understood as merely a device for communicating findings, looks forward to future operations, and contains recommendations for improving the effectiveness of operations by design.

This report also contains a variety of works on various theoretical topics: the assessment of the protection of civilians, containment of nuclear materials, strategic communication, and command and

control systems; the potential of red teaming and surveys for assessment; the selection of indicators; and several other models or techniques for making sense of complex environments. Practical examples include real world assessment processes from the Pacific region, Afghanistan, Iraq, the Horn of Africa, Western Africa, and major combat operations exercises.

Évaluation des opérations dans les environnements complexes : théorie et pratique

(STO-TR-SAS-110)

Synthèse

Les environnements dans lesquels l'OTAN et ses pays membres mènent des opérations militaires sont de plus en plus complexes. La montée de puissants acteurs non étatiques et la probabilité croissante de conflits au sein des populations civiles rend les opérations plus difficiles et la progression vers leurs objectifs, plus délicate à percevoir. D'où l'attention accrue que nous portons à l'évaluation des opérations, afin de déterminer l'efficacité de ces dernières dans ces environnements complexes.

L'étude qui donne lieu au présent rapport est un symptôme de l'inquiétude partagée par la communauté de l'évaluation au sujet de sa capacité à fournir des conclusions d'évaluation utiles, opportunes et empiriquement vraies, et ce, pour faciliter la prise de décision et améliorer l'efficacité des opérations, mais également pour justifier les ressources que les gouvernements nationaux ou les autorités militaires supérieures fournissent aux organisations militaires subordonnées. La communauté de la recherche et de l'analyse opérationnelles, qui constitue la communauté d'évaluation de facto, a estimé qu'elle pouvait faire preuve de plus d'efficacité dans la conception et la direction des processus d'évaluation, ainsi que dans la prestation de l'évaluation et la fourniture des produits d'évaluation aux commandants et organisations qu'elle soutient.

La conclusion essentielle de cette activité est que le fait d'insister clairement et explicitement sur l'amélioration de l'efficacité de l'organisation qui évalue ses activités, au lieu de se concentrer sur la justification des ressources, améliore la qualité des conclusions et des produits de l'évaluation. Les opérations sont alors plus efficaces.

Par le passé, les processus d'évaluation se focalisaient souvent sur la justification des ressources consommées pendant les opérations auprès de l'autorité militaire ou politique supérieure, conformément à un ensemble d'exigences légales ou réglementaires. Le rapport d'évaluation était donc une exigence bureaucratique, demandé par une autorité supérieure et rédigé par un groupe d'évaluateurs dont c'était le travail. Le personnel, et fréquemment le commandement, concentrait son attention sur le « travail réel » de gestion des opérations – plutôt que sur les exigences bureaucratiques – tandis que les évaluateurs travaillaient seuls – sans l'expertise des autres membres du personnel – pour fournir les justifications exigées par l'autorité supérieure. En conséquence, le travail portait davantage sur la forme du rapport que sur les conclusions qu'il contenait et sur les gains d'efficacité possibles au vu des conclusions. Par ailleurs, le rapport était un compte rendu historique jetant un regard rétrospectif sur la période précédente. Les recommandations qu'il présentait étaient presque une réflexion après coup.

Par opposition, un processus d'évaluation focalisé sur la découverte de moyens d'améliorer l'efficacité a plus de valeur pour le commandant de l'organisation et n'est plus une exigence bureaucratique. Le commandant s'intéresse désormais fortement aux conclusions de l'évaluation plutôt qu'à la forme du rapport et le personnel y répond en apportant l'expertise qui fait défaut aux évaluateurs, ce qui rend le rapport d'évaluation beaucoup plus utile. Le rapport, considéré à juste titre comme un simple dispositif de communication des conclusions, s'intéresse aux opérations futures et contient des recommandations visant à améliorer l'efficacité des opérations dès la conception.

Le présent rapport contient également divers travaux sur divers sujets théoriques : l'évaluation de la protection des civils, le confinement des matières nucléaires, la communication stratégique et les systèmes de commandement et contrôle, le potentiel de recherche des failles et les études d'évaluation, la sélection des indicateurs et plusieurs autres modèles ou techniques donnant un sens aux environnements complexes. Les exemples pratiques incluent des processus d'évaluation du monde réel, issus de la région Pacifique, d'Afghanistan, d'Irak, de la Corne de l'Afrique, d'Afrique occidentale et de grands exercices d'opérations de combat.

INTRODUCTION TO OPERATIONS ASSESSMENT IN COMPLEX ENVIRONMENTS

The volume that follows is divided into a theory section and a practice section. The former explores situations that may become new problems for practitioners of operations assessment; promising methods or tools, including qualitative ones; thoughts on using assessment tools; some thoughts on the history of the assessment problem; and an annotated bibliography. The latter provides examples of assessors solving real world problems, and allows them to explain the strengths and weaknesses of their methodologies, and perhaps, most importantly, they explain their lessons learned or provide advice to new assessors. Of course, there is considerable cross over where a theoretical point is illustrated with an example, or a practical problem leads to a theoretical insight.

Each chapter begins with a short abstract to provide detail to readers on the topic of the chapter.

The team that edited the work (and wrote about a third of it) failed to find a single overarching construct that can be applied to all operations assessment problems, but the reader will find that variations of the word "effective" appear frequently throughout the work, in every article, leading to some conclusions about the purpose of operations assessment. Given that higher authority frequently mandates assessment products, in the past, assessment sometimes devolved into a mechanism for the accountability of the resources that higher authority invested in the reporting command's mission. However, given the prevalence of the use of "effective" in the selections in this anthology, we see that a mere accountability exercise squanders the potential of assessment processes to illuminate systematically and empirically the nature of the operational environment and the command's place in it. This superior understanding, in turn, results in better leadership decisions, and thus more effective operations. Indeed, to perceive assessment's primary purpose as "making operations more effective" has the effect of improving the assessment process for both effectiveness and accountability. Adam Shilling, the senior editor of this volume, explores the philosophical implications of that realisation for the practice of operations assessment in Chapter 1, and briefly discusses similar notions in the lessons learned portion of his article on the assessment of security cooperation events in the Horn of Africa in Chapter 13.

In the theory section, readers will find several articles that grapple with new problems for the assessment community. Beadle, Lien, and Gustavsen (Chapter 3) discuss the assessment of missions to protect civilians under threat in unstable nations, and the manner in which such assessment contributes to the effectiveness of these missions. It outlines five generic approaches to measuring protection of civilians in military operations (casualty figures, civilian behaviour, perception of security, territorial control, and perpetrator capabilities), and uses a series of scenarios to illustrate the relative value of each approach. Bovey, Fainberg, and Thomason (Chapter 4) examine the problem of loose nuclear materials and how assessment can assist in the containment and recovery of these materials. They argue that assessment techniques performed in other missions can be borrowed for the containment of, and the search for, stolen nuclear materials. Their analysis touches on planning, command and control, logistics, and equipping issues. Tillman and Conley (Chapter 10) explore different Command and Control (C2) techniques, and how assessment can help select the technique most appropriate for the situation, confirm the effectiveness of the selected technique as the operation progresses, and vary the technique as the situation changes in order to maintain effective control. They introduce the concept of C2 Agility which is the ability of organisations to adapt C2 techniques rapidly and efficiently to keep pace with a rapidly changing operational environment, and maintain the organisation's effectiveness. Pennell and Miller (Chapter 11) examine the difficult relationship between strategic goal setting, analysis, and operations assessment. This requires some consideration of how strategic goals are developed, a look at the difficulties in honestly assessing progress in any mission, and a reconsideration of the decision-making model which underpins current analysis and assessment practices.

Other authors introduce promising methods or tools. Minkov and Dickson (Chapter 2) are veterans of "red teams", and suggest ways in which Red Teaming can be used to assess operations and make them more effective. A Red Team is a deliberate attempt to challenge the organisation by providing alternatives through critical thinking in order to improve decision making and achieve the endstate most effectively. Red teams will attempt to look at the conflict environment from the perspective of various actors other than the NATO force, such as partners, adversaries, or important third parties. They explicitly attack the validity of friendly assumptions and look for biases or errors that lead to poor planning and to lesser effectiveness. Frelin (Chapter 5) grapples with the use of qualitative methods to fill in the inevitable gaps where quantitative methods are inadequate. His work suggests ways to use qualitative techniques properly in order to reach appropriate conclusions. Jordan (Chapter 6) describes an approach to complex systems which focuses on exploration and evaluation that is qualitative rather than quantitative, and that will support evaluation of complex systems both at strategic and lower levels, and explores the effect of innovative improvements, whilst providing a framework for the incorporation of other analytical techniques. Poponete (Chapter 7) discusses NATO strategic communication requirements and techniques, and how to assess their effectiveness. He recommends a method that uses systematic and objective analysis of texts via computer-based content analysis that help communication practitioners improve their understanding of how messages impact audiences from people in remote villages to the most internet savvy populations. Shilling (Chapter 13) provides a model that embodies a common-sense approach to the identification of appropriate indicators given a set of desired outcomes. By preserving the logical links connecting outcomes to indicators, assessors can isolate the most valuable information and focus collection assets appropriately.

Still other authors describe the use of assessment tools and provide advice on how to make the most of them. Eles (Chapter 8) examines how a programme of public opinion polling supported NATO operations in Afghanistan. This programme provided situational awareness of the strategic operating environment; a direct measure of the effectiveness of strategic effects related to population perceptions; an indirect measure of the impact of other operational tasks; a direct measure of those operational tasks which are conducted to influence perceptions; and an estimate of population demographics. Eles continues (Chapter 9) by providing practical examples of survey analysis in complex environments. They highlight ways in which such analysis can be challenging, such as the constraints associated with a poor security environment, cultural aspects of data collection, or limitations in knowledge of population distributions, and population displacement. Therefore, survey methods may stray from textbook examples.

Finally, in the theory section, Frelin (Chapter 12) looks at several historical cases to investigate how the formal methods for creating effective feedback have assisted military forces over the last quarter century. He finds that these have generally fallen short, which creates an important gap between military practice and doctrine. He then attempts to explain this failure, using theoretical ideas from cognitive research, organisational learning, and the research on complexity and chaos, and provides some recommendations.

In the practice section, authors provide examples of solutions to real world problems that will reveal and criticise methods that future assessors might adapt to solve future problems. Examples come from Afghanistan, which is the locus of so much NATO experience in the last two decades, from Iraq, where NATO members have been engaged, from security cooperation in Africa and the Pacific, and also from a U.S. exercise featuring what Americans call "major combat operations".

Larrabee (Chapter 16) describes an assessment process and framework developed for the advisory group supporting the Afghan Special Operations Command. The framework tracked indicators on command and control, personnel and equipment management, planning, and logistics, and provided analysis on these indicators to assist Afghans to become more effective. Larrabee describes the initial operational design conference, assessment framework development, and implementation of the assessment. Eikelboom, Gouweleeuw, Roseboom, and Schwerzel (Chapter 21) relate the experiences of Dutch assessors in Uruzgan Province, Afghanistan. Dutch-led Task Force Uruzgan assisted the Afghans to set conditions for a secure and stable province, to assist the provincial government in building its capacity, and to synchronise the overall

reconstruction and development effort. Its assessment section worked hard to demonstrate the usefulness of assessment in making operations more effective. Their activities included helping planners specify and prioritise specific and measurable effects and to measure the progress made toward them.

Nowak (Chapter 17) provides insight on maximising the assessment benefit of limited data from the early days of Operation Inherent Resolve in Iraq. When the Assessments Cell was formed, the Inherent Resolve Coalition was not yet fully manned, and the availability of relevant data was significantly limited and often lagged behind events by several weeks. Nowak describes the Assessment Team's effort to overcome data obstacles and provide timely, relevant, and credible analysis to inform senior leaders in a Coalition conducting operations in a complex environment.

Three practical examples come from the Combined Joint Task Force – Horn of Arica (CJTF-HOA). The chapters are arranged in chronological order. Shilling (Chapter 18) describes in some detail how the command planned and evaluated the security cooperation "events" that formed a significant portion of the command's activities. The point was to prepare mission commanders, relatively junior soldiers (OR-6 to OF-4), to be most effective and to capture their lessons learned so that the next mission would be more effective also. The methodology also captured observations on the capacity of African partner militaries. The lessons learned section of the chapter presents a brief discussion of what might be called principles of assessment that overlap well and complement the ideas in Chapter 1 of this volume. Boncich and Casey (Chapter 19) describe the assessment processes in place in CJTF-HOA a year after Shilling left the command. Their description includes methods and processes to help the command prioritise their activities to maximise their effectiveness given limited resources. Their assessment focused at the CJTF-level rather than at the level of the security cooperation event. Analytic efforts included: integration of operations assessments into the command decision-making process; development of an operations assessment framework for the campaign plan; analysis of the level of impact from activities; analysis of significant activities data from the African Union Mission in Somalia (AMISOM) Force Headquarters; and analysis of Public Perception Surveys. Leno (Chapter 20) served at CJTF-HOA a year after Boncich left the command, and describes the processes in use at the CJTF-level at that time. His recommendations for new analysts were very practical and focused on how new assessors can integrate themselves into the staff. In addition to the chapters on CJTF-HOA, Engelmann and North (Chapter 23) provide an approach for evaluating Security Sector Reform (SSR) and Defence Sector Reform (DSR) in Sierra Leone. They introduce a Defence/Security Sector Reform Concept Model to organise activities and provide a model for assessment of the effectiveness of these activities.

Thambidurai and Wayman (Chapter 15) describe the assessment process and products of U.S. Special Operations Command Pacific (SOCPAC). The chapter covers assessment of all activities of the command, and talks in some detail about the development of appropriate indicators and the collection, processing, and integration of qualitative data to answer the questions that the commander, staff, and assessment cell have posed. Finally, they show some examples of the products they used to communicate their findings.

Swedberg (Chapter 22) departs from the norm for this volume with a discussion of the assessment challenges in operations with "decisive action focus". His chapter presents the processes and products he employed during a U.S. Warfighter exercise characterised by intense conventional combat. He found the primary question that drove his analysis for the commander was "Is the plan still valid?" or had enemy action or unexpected opportunity required a change of plan. He details the process required for assessment under severe time constraints of major combat operations.

Connable (Chapter 14) rounds out the volume with an annotated bibliography to the most important assessment literature. He will help new assessors to get started and old assessors to increase their knowledge.

This volume represents a significant amount of work from some experienced assessors. Their desire was to make the readers' learning quicker and less painful than their own. We hope we have succeeded.

Part 1: THEORY OF OPERATIONS ASSESSMENT

Chapter 1 – SOME PHILOSOPHY: A NEW WAY OF LOOKING AT OPERATIONS ASSESSMENT

Adam Shilling
US Army
UNITED STATES

ABSTRACT

This chapter introduces a "new way" of thinking about operations assessment for complex environments. It answers some disquiet that American assessors experienced in the past several years by thinking about assessment in a new way. It finishes with a paragraph for each of the new things that an assessor must know and do to maximize their effectiveness as a member of the command's staff.

The report that follows is necessary because there is concern about the assessment community's ability to provide empirically-backed understandings of complex environments that would aid commanders in deciding which activities would lead to the achievement of their desired outcomes most effectively. A great deal of thought and discussion has gone into the attempt to assuage this concern – including this effort – and a new way of looking at the operations assessment problem is emerging.

NATO's definition of operations assessment, found in version 3.0 of the NATO Operations Assessment Handbook [1], is "The activity that enables the *measurement of progress and results* of operations, and the subsequent development of conclusions and recommendations that *support decision making*" (emphasis is mine).

In thinking about operations assessment in a new way, the first lesson is that the *primary purpose of assessment is to make operations more effective*. This statement sounds hardly earth shattering, until I point out that assessment's purpose, the definition above notwithstanding, is not primarily to measure progress or to support decision making, although it does both those things. These things are only instrumental to its true purpose: making operations more effective.

Most experienced assessors do not need an explicit use of any form of the word "effective" in describing assessment activities because they understand the point of their efforts has always been to produce "the subsequent development of conclusions and recommendations that support decision making." Nevertheless, I have found that use of the word "effectiveness" has a somewhat magical effect upon senior leaders (because they are all about effectiveness) and upon other members of the staff, both junior and senior, who have mistakenly believed that the point of assessment was to produce a colourful set of graphics to project on the wall or send off to higher headquarters. The use of "effectiveness" generates an epiphany in their minds that makes clear that the *product of assessment is an effective operation, not a slideshow*. It also makes clear that assessment done right requires their participation and their expertise.

All of that said, the *secondary purpose of assessment is, in fact, to provide accountability*. Relegating accountability to secondary status does not mean that it is not important, just of lesser importance than making operations more effective. It is absolutely critical to keep this order straight because it affects the priority that the assessment process receives within the headquarters, the behaviour of the staff relative to the assessment process, the form of relevant written products, and the focus of the assessment section as it leads the staff through the assessment process. Effectiveness over accountability also affects senior leaders' attitudes toward the assessment process because they begin to perceive it not as a bureaucratic requirement imposed by their bosses, but as an aid to their own personal effectiveness and useful in making their organisations more effective.

Indeed, in my first assessment assignment, I wrote several recurring reports for General David Petraeus, then commander of the International Security Assistance Force (ISAF), and was responsible for the assessment of the Afghan National Security Forces line of effort, but I never spoke to the gentleman personally. So I provided accountability for the headquarters, but cannot truly say that I supported decision making. Of course, the empirical information that I reported to the NATO, United States, or other member states' hierarchies, which was gathered by other people, was shared with General Petraeus, by still others, and the information was, in fact, supporting decision making.

From this, I learned that *assessment is everybody's business*. A host of individuals on the staff or throughout an organisation should be involved in assessment – determining what information requires collection, who will collect it and how, how it will be reported and stored to facilitate analysis, and what sort of analysis is required. Assessors, usually operations researchers, cannot do all of these things, without assistance, for all functional areas on the staff, and should not need to. They simply do not have the requisite expertise. Experts throughout the staff collect information, make sense of it, and report the results to decision makers. The value added by operations researchers is in helping to structure information collection and storage, in increasing the rigour and sophistication of analytic or data visualisation techniques (when appropriate), and in the quality assurance of the analysis of empirical data.

This also implies that *operations assessment is not new*. The fact is we have always collected data, and always made sense of it. We have personnel reports, logistics reports, operations reports, training evaluations, and a host of other things that gather and report data, answer leaders' questions, and provide the basis of sound decision making that leads to more effective operations. What is different is that the environments NATO conducts operations in are more complex today than in the past. Back then, knowing the front line trace, the loss-exchange ratio, and the ownership of various pieces of key terrain was sufficient to say we are winning or we are not. This is no longer true.

Therefore, the daily brief that many commanders receive, though not a product of the assessment section, is still an assessment product as it is a monitoring of key information that answers key questions decision makers have posed. The idea that assessment is larger than the written products of the assessment section shows that *assessment is a process more so than a product*. Indeed, undue focus on the production of a product or set of products distracts from the process of gaining an improved understanding of the operational environment, and our command's place in it, that we need in order to select the activities for friendly forces which will make us most effective.

One way to begin an appreciation of the new way is to discipline our vocabulary when discussing assessment processes and products. As we circulate about a headquarters, we will hear the words, "the assessment". Typically, these words refer to the report, or the slideshow, or some other tangible deliverable, which would thump if we dropped it on the table. This, I argue, is a cognitive problem. We should consider "the assessment" to be *the improved understanding of the operational environment, and our place in it, that we get from working our way through a deliberate assessment process*.

This understanding is a cognitive product. It is too large, too complex, too nuanced, too changing to be written down readily. It resides in the minds of multiple members of the staff, because it requires more varied expertise than any one staff member possesses. It is analogous to what the U.S. Army calls a "running staff estimate", and if we did (or could do) our business perfectly, "the assessment" would be indistinguishable from the common operating picture.

The written product, whatever form it may take, is merely an executive summary of this understanding at a point in time. Its purpose is to communicate to leaders the important things that the staff has discovered while working through the assessment process, answer questions that leaders have posed, or provide information that decision makers have requested. *The product is not the point of the assessment process; it is merely a communication device*.

In the past, we believed this product to be the deliverable, and as good staff officers will, we focused our efforts on this deliverable rather than on the understanding it is designed to communicate. The appearance of the product was frequently of greater consequence that the understanding it contained. We compounded our error by using phrases such as "measure progress" to describe the purpose of the assessment process, which reduced the product to a sort of scoreboard we could use to show ourselves that indeed we are winning.

We further compounded our error in that many products were designed to account for the resources we were expending to higher headquarters and national governments. Accountability is critical, to be sure, but this focus gave our products the character of history reports which looked backward over the preceding period. For all but the best assessors and the wisest senior leaders, the appendage of recommendations for improving the effectiveness of operations was virtually an afterthought, rather than the purpose of the assessment process.

Now we understand that the deliverable of operations assessment is not a report or slideshow, but an effective operation. By focusing efforts on understanding our environment, we hopefully can understand it better, and therefore can design activities for friendly forces that achieve our desired outcomes more effectively. Then our effectiveness-focused products are forward looking, and the recommendations for more effective operations are produced by design, and not as an afterthought.

The new way of looking at operations assessment recognises a number of things:

- **Effectiveness is more important than accountability.** If higher authority has provided resources for us to achieve a set of outcomes, the people in authority would prefer that we succeed in the most effective manner possible rather than they would receive a colourful, aesthetically-pleasing report. This is not to say that providing accountability is not a completely legitimate use of the staff's time. Accountability is critical to many human endeavours, military operations included. But the focus of the operations assessment process should be on discovering the truths we need to select appropriate operations for our forces that will make them most effective. Moreover, usually it is the case that *if we are collecting the information that is critical to understanding the operational environment, and making our operations more effective, it requires little additional effort to provide the accountability that higher authority requires*. It may be as simple as packaging the information in another way or a different format.
- **Assessment is not about writing a report.** The report is merely a communication of the true assessment – the improved understanding of the operational environment that we get from working through the assessment process. This report is not our deliverable, and it is not the focus of our efforts. It is merely instrumental to sharing what the staff has learned with leaders so that they can make decisions and issue orders to make operations more effective.
- **Assessment requires a whole-of-staff approach.** Although assessment billets often go to operations researchers, no group of several operations researchers huddled together in an assessment section has the expertise they need to identify the most important information requirements for every functional area, collect the indicators which answer them, and make sense of the trends of those indicators. Assessors need the expertise of functional experts on the staff – in operations, intelligence, sustainment, civil-military cooperation, legal matters, and so on – to identify the indicators, to aid in their collection, and to make best sense of them.

  In the past, when assessment products provided accountability only, the staff often behaved in a manner that assigned any tasking that included the term "assessment" to the assessment section, while the rest of the staff got on with the "real work" of the headquarters. Now that the assessment process is focused on effectiveness, the commander and staff understand the value added of an effective assessment process, other staff members contribute their expertise freely, and key leaders receive the benefits of broad, staff-wide expertise that is now contained in assessment products.

- **A focus on effectiveness means products look forward. A focus on accountability features products that look backward.** In the past, written products had the character of history reports. "We did these things, and the results were…" This is important surely, but not nearly as important as saying "the next thing we need to do to be more effective is…"
- **A focus on effectiveness also means products are for key leaders and members of the staff within your organisation (in and down) while a focus on accountability means products are for other people (out and up).** Effectiveness products communicate information and recommendations to the lowest leader who can effect the needed change. Accountability products often go to someone who is unlikely to effect any change or will only do so by the reallocation of resources at a higher level.
- **Aggregation schemes hide the problems we are looking for.** A number of assessment products in the past relied upon aggregation schemes to make things digestible to decision makers or to quantify a diverse set of indicators into a score that allegedly demonstrates progress. There are a couple of problems with this approach. Some things are just not quantifiable; diverse quantifiable things should not be combined into a single, meaningless measure; and inept aggregations to not help anyone understand the problems we are facing or the effective solutions to them. Rather, given that assessment is about isolating problems for solution, hiding them in a complex aggregation scheme does not help.
- **Conceptually, questions are more important than answers.** A lot of staff members mistake the mere collection and display of indicators for the meat of assessment work. This is a problem because new assessors, myself included, often attempt to know all that can be known – to collect as many "relevant" indicators as they can. The problem with this approach is that collection resources are limited, and some information just is not that important. The way we can determine what information is important is to ask a series of questions about our environment, the changes we wish to make in it which are specified as outcomes, and the empirical evidence of change we expect to see as we succeed (or fail). These questions form the logical links between the desired outcomes we are seeking and the empirical evidence of their achievement. The most specific of these questions, that are answered by empirical evidence, are information requirements, and the evidence which answers them are indicators.

The list of questions is easier to prioritise than an exhaustive list of potential indicators. This helps us to isolate the most important information, because if the question is not important, then the answer is not either. The list of questions also helps us understand the logical relationships of indicators to each other, which we might miss if we only have a list of indicators.

Ultimately, the New Way of looking at operations can be summed up as: *operations assessment is a systematic approach to understanding the operational environment, and our place in it, so that we choose activities for our forces that make them most effective in pursuit of our desired outcomes.*

1.1 AUTHOR'S BIOGRAPHY

Adam Shilling, PhD, is a veteran of two operations assessment tours abroad – one to the Afghan Assessment Group at Headquarters, International Security Assistance Force (ISAF), and one at the Combined Joint Task Force – Horn of Africa. He spoke about his experiences and lessons learned at an international meeting. His ideas made sense to analysts and decision makers, and so his superiors selected him to represent his organisation, the Centre for Army Analysis, in the community that was writing operations assessment doctrine for the United States military, and to act as senior editor for this report.

1.2 REFERENCES

[1] NATO. 2015. NATO Operations Assessment Handbook, version 3.0. Norfolk, VA: Supreme Allied Commander Transformation. July 2015.

Chapter 2 – RED TEAMING AND OPERATIONS ASSESSMENTS

Anton Minkov
Department of National Defence of Canada (DND)
CANADA

Paul Dickson
Canadian Joint Warfare Centre
CANADA

ABSTRACT

This chapter highlights the potential of a Red Teaming analytical capability to inform or contribute to operations assessments, drawing on the authors' experience at ISAF. The chapter starts with a general overview of how the ISAF campaign was analysed and assessed in order to illustrate the disparity of assessment sources available to the commanders in theatre. It demonstrates that operations assessments are not confined to formal assessment staff groups, nor they are as centralised as previously thought. Commanders realized the limitations of centralised assessment products, based primarily on quantitative metrics, and sought other sources to augment these products. Decision support Red Teaming is one such capability that can complement operations assessments. The chapter outlines the similarities between operations assessments and Red Teaming theory. Since the main challenge to any assessment process is whether it is measuring the right things in the right ways, Red Teaming could act as the independent reviewer of critical elements of the assessment process. Methodologically, it offers a range of qualitative and contextual analytical methods to decision makers. The chapter also provides examples of Red Teaming products that served as or informed assessment.

2.1 INTRODUCTION

Operations assessments, i.e., measuring the progress of the campaign, have again become a prominent element in the recent counterinsurgency campaigns in Iraq and Afghanistan. As in all complex operations, Counterinsurgency (COIN) is very analysis and intelligence intensive. COIN doctrine emphasises the importance of ongoing analysis and assessment. It defines assessment as the "continuing monitoring and evaluation of the current situation and progress of an operation against established criteria" [1]. Performance measures address task accomplishment against intent – are we doing things right? – while effectiveness measures addressed change against expectations – are we doing the right things?

Simultaneously with the progression of both campaigns, and following them, a significant body of literature on the topic emerged criticising elements of the assessment process, the metrics being used, approaches (i.e., quantitative vs. qualitative), or the frameworks being utilized for assessment [2], [3], [4], [5]. It has been also noted that the assessment process as a whole, as well as data repositories feeding that process, eventually become centralised and institutionalised as part of the force structure [6]. Thus, the products of the staff charged with formal assessment have been most visible and had certain "official" bearing on the chain of command, and often go beyond internal consumption – e.g., public and congressional briefings. However, the reality is that even after the establishment of formal centralised assessment bodies, other groups in theatre continued to provide analytical products, and commanders informed their opinions by a multitude of sources. In fact, US doctrine also suggests that a significant portion of a commander's assessment may come from other sources than staff assessment [7]. These additional sources may have not necessarily been called assessment products and thus remain obscure in the overall discourse on operations assessments.

It is important to acknowledge that analysis and assessment are distinct in form, even if they can both perform the same function. Analysis is best characterised as a focused process that uses an identified method. Assessment is a decision support function that might or might not include one or more analytic processes and methods. An assessment can include analysis, and the latter can serve as the assessment. For example, a military commander can provide a campaign assessment based on personal reading of the data with no

formal analysis, or they can incorporate a time series analysis of similar data over time, which requires the focused application of a method. The latter could serve as the assessment, or could be one of a series of analyses that inform an assessment of the campaign [8].

This chapter will highlight the potential of a Red Teaming analytical capability to inform or contribute to operations assessment drawing on the experience of the authors in the International Security Assistance Force (ISAF) Red Team during 2012–13 in Afghanistan. The Red Team was one of three headquarters analytical groups providing the Commander, ISAF; the Deputy Chief of Staff, Intelligence (CJ2); and other senior commanders with in-depth analytical products to support decision making and avoid strategic surprise. The establishment of the Red Team was part of the evolving analytical assessment architecture that reflected refined ideas for executing the operational concepts and strategy.

First, the chapter will provide a general overview of how the ISAF campaign was analysed and assessed in order to illustrate the disparity of assessment sources available to the commanders in theatre. Then, it will demonstrate the similarities between operations assessments and Red Teaming theory and finally, provide examples of Red Teaming products that served as or informed assessment.

2.2 ASSESSMENT AND ANALYSIS IN ISAF

In the first eight years of the campaign (2001 – 08), individual contributing nations had developed their own assessment models and methods resulting in, by one analyst's estimate, over twenty different campaign assessments [3]. From early 2009, a new campaign analysis architecture was established, both to address the shortfalls highlighted later that year by General Stanley McChrystal [9], and as part of the effort to bring coherence to the campaign assessments. However, there was no consistent or comprehensive capture of the required data before 2009. The stand-up of the Afghan Assessment Group (AAG) at Headquarters (HQ), ISAF in early 2009 was recognition that analysis was an essential element of counterinsurgency. It was also the first step towards a centralised strategic and operational assessment process. McChrystal refined the analytical architecture, expanding the AAG's mandate to include qualitative assessments, and changed the reporting structure, minimising charts and graphs, and calling for more interpretative reports. He also established a COIN Advisory and Assistance Team (CAAT) to "operationalize intent in the ongoing counterinsurgency campaign" to support the effort outside the capital [10]. Data and qualitative assessments were organized along three Lines Of Operation (LOO) that reflected McChrystal's population-centric COIN: protecting the population, building the Afghan National Security Forces (ANSF), and increasing the capacity of the Afghan government [6]. In addition, McChrystal leaned heavily on outside experts to identify metrics and assess their significance. Primary among those shaping the COIN debate was David Kilkullen, an Australian officer who was instrumental in promoting COIN, and then influencing the debate about metrics as well. His standard brief – "United States Counter-insurgency: An Australian View" – was considered the catalyst for a U.S. reconsideration of the campaign [11]. McChrystal also pressed, unsuccessfully, to have the narrative assessments unclassified, and subject to scrutiny by academics, journalists, and other organizations.

General John Allen's assumption of command in July 2011 marked perhaps the most significant overhaul of the ISAF assessment process. The need for change was necessitated by the start of the drawdown of the U.S. surge that same month, the announcement of formal timelines for the transition to Afghan-led security, and the consequent debates over the scale and timing of troop withdrawals. All that increased the complexity of assessing the campaign as well as the variables to be considered. There was increased pressure to understand how the campaigns were progressing relative to established timelines, decreasing resources, and ever-changing threats to success. The decision to declare an end to the mission reshaped the idea and evaluation of progress. Victory was not simply a function of transitioning security to the Afghans, but how successful the latter would be at providing security. Identifying conditions for the transition exacerbated the challenges of measuring success [12].

Within a few months of assuming command, Allen changed the assessment architecture. AAG initiated a review with the aim of creating a system of assessment that was both holistic and comprehensive [13]. Two levels of assessment were developed: a strategic and a campaign assessment. The former was designed to ensure NATO and U.S. policy and strategic goals were being met. The latter was a refinement on existing metrics, with the notable introduction of "regional relations", i.e., how neighbouring countries can impact the campaign. These changes prompted a rebalance between quantitative and qualitative analysis to inform the assessment process.

The new Commander, ISAF (COMISAF) also tried to introduce a more systematic approach to coordinating the analytical efforts across the headquarters, as well as enhancing the capacity to provide contextual and qualitative analysis. The renewal of the Red Team (originally stood-up in 2011) was one example. Allen also changed the remit of the CAAT, renamed the COMISAF Advisory and Assistance Team, with a new mission to provide directed observations and reporting to COMISAF on strategic priority areas. In practice, the CAAT was focused on the development of Security Force Advisory Teams, the aim of which was to advise, mentor, and train Afghan security forces at the unit level [10]. The third team charged with providing context and assessments was the Commander's Action Group (CAG), a small multinational staff working directly to support COMISAF. The CAG was formerly known as the Commander's Initiative Group (CIG). In 2011, it had acted as a support staff preparing files and briefing books for COMISAF, but its role was expanded in 2012 to include some analysis.

Over the course of 2012, an equilibrium was established that assigned the CAG more immediate and short-term assessments, and the CAAT specific tasks related to campaign execution. The Red Team focused on contextual and narrative analysis. There were other groups of analysts providing analysis on similar topics and issues. The CJ2 intelligence officers provided "deep dives" and daily assessments. Some analysts worked to support their national contingents; others worked for other layers of command, notably the ISAF Joint Command (IJC) and the Regional Commands (RCs). Desk officers in the various divisions also produced analysis and assessments. The NATO Senior Civilian Representative (SCR) had its own analytical pools. Commanders also brought their own advisors or brought in outside experts and specialists. There was no shortage of analysis being undertaken, representing a wide range of opinions and conclusions, and of varied quality and impact.

U.S. Marine Corps General Joseph Dunford assumed command in February 2013. He established headquarters priorities as managing the ISAF drawdown and the transition to the post-2014 Resolute Support Mission [14]. The change of command also brought about a change in decision-making culture and a new analytical support paradigm. The Red Team's analytical contributions, for example, gradually shifted from papers to briefings as well as direct support to staff and operational planning.

2.3 RED TEAMING AND OPERATIONS ASSESSMENT

Red Teaming is a fairly well-established concept in U.S. Army and Navy. It provides an opportunity to challenge the organization by providing alternatives through critical thinking in order to improve decision making and achieve the end state most effectively [15], [16], [17]. In 2005, a Red Teaming training centre was established at the University of Foreign Military and Cultural Studies at Fort Leavenworth, Kansas, which produced the Red Teaming Handbook. It was recently re-titled *The Applied Critical Thinking Handbook* [18]. While there are a variety of definitions, the Handbook defines Red Teaming as "a function that provides commanders an independent capability to fully explore alternatives in plans, operations, concepts, organizations and capabilities in the context of the Operational Environment (OE) and from the perspectives of partners, adversaries and others" [18]. To distinguish Red Teaming as an adversarial perspective, typically used in threat assessments in intelligence and cyber-security, from that used in military operations planning, NATO introduced the term "alternative analysis", which it defined as "the application of critical thought from an independent perspective" [19].

Since the main challenge to any assessment process is whether it is measuring the right things in the right ways, Red Teaming could act as the independent reviewer of critical elements of the assessment process. According to the latest edition of *The Applied Critical Thinking Handbook*, key assessment questions for Red Teaming include:

- Are the proposed measurements of effectiveness clearly linked to the strategy, mission, or endstate?
- Does the measurement have a clear start point (baseline) in which to measure progress?
- Does the measurement system compatible with higher headquarters metrics? Are the unit's tasks developed to local conditions?
- What are the assumptions in the measurements and the definitions?
- What is the level of coalition or interagency agreement to the assessment measures? If no agreement, what are the implications?
- How do we validate the data collection? Is it comprehensive or representative?
- Who has primary responsibility for assessment? Has the task (who, what, when, where) been established?
- Do the metrics reflect a cultural sensitivity, whereby important things are measured? From the civilian population perspective, does the MOE matter?
- What are the expectations of the people in terms of patience for process?
- From the enemy's perspective, what are their measures of effectiveness? Does our MOEs measure what is important to him?

Furthermore, recent studies in operations assessments have identified need for more qualitative analysis in assessments [20], a function that could be fulfilled by Red Teaming analysis. Decision support Red Teaming applies a range of methods to offer decision makers qualitative and contextual analysis. The following elements can be provided by such an analysis:

- Broaden understanding of the operational, strategic, and political environment.
- Assist in defining end states.
- Challenge planning assumptions and definitions.
- Offer alternative perspectives.
- Ensure the consideration of adversarial perspectives, e.g., how an adversary might assess their operations – current and future enemy courses of action.
- Identify friendly and enemy vulnerabilities.
- Ensures staffs are assessing the right things (indicators).
- Help the staff determine the right things to do.
- Anticipate 2nd and 3rd order effects of operations.
- Anticipate the strategic and operational-level implications of actions.

2.4 PRACTICAL EXAMPLES OF RED TEAMING AT HQ ISAF

The COMISAF Red Team's (RT) mandate was decision support Red Teaming as distinct from Red Team (or Red Force) analysis that emulates the enemy or threat (see Appendix 2-1). The Team was also built to provide a range of perspectives as a multinational team of military and civilian analysts. Their products (papers and briefs) provided alternative and strategic perspectives and recommendations on key issues (as identified by

COMISAF), examined underlying assumptions, adversary motivations, regional dynamics, and social, cultural and economic factors surrounding critical issues. Most notably, its key responsibility included "researching and writing alternative assessments using sound analysis and critical thinking techniques" (see Appendix 2-1).

As mentioned earlier, the Red Team worked closely with the AAG and the other analytical groups at HQ, ISAF. On occasion, the Red Team hosted outside analysts brought into theatre for ad hoc assessments. For example, during the Spring of 2013, the Red Team hosted a senior analyst from the Congressional Research Office tasked to prepare an independent assessment for the Commander ISAF Joint Command (IJC). What is significant from an organizational point of view was that, although formally under CJ2 in the organizational chart, the Red Team had a direct link of communication with COMISAF and the list of future projects for the team was approved personally by the Commander. The team's products had a very limited distribution, having primarily COMISAF and the other top commanders as its main clients. As a result, the Red Team production was not influenced by bureaucratic competition or turf war with other assessment and staff organizations, and was able to fulfil its obligation of providing independent analysis. Another critical element for success was the team's composition. During the authors' tenure, the team members were all experienced analysts with multiple deployments in Iraq and Afghanistan, with deep knowledge of Afghan history and culture, and considerable experience with ISAF operations.

The examples below illustrate how Red Team products were able to address some of the assessment questions or provide critical context to campaign objectives. The cases are drawn from the work of the two authors at the Red Team in 2012–2013. Since the papers were classified as Secret, Releasable to U.S., NATO, and ISAF (paper titles and section headings were unclassified), the vignettes only describe the nature of the topics and not any actual conclusions or recommendations.

2.4.1 Reframing the Outcome of the 2013 Fighting Season

The Red Team paper on the assessment framework for the 2013 fighting season represents a good example of challenging assessment metrics. A desire to produce impressive metrics of success was probably on the mind of every COMISAF because an important part of his job was to handle pressure for results from Troop Contributing Nations (TCN). In that sense, assessments were also considered by senior commanders to be part of or contributing to the strategic narrative emanating from ISAF to the TCNs. In 2013, positive developments in the security environment were deemed particularly important for the continuous support of TCN in 2014, as the U.S. forces commenced their drawdown. When developing a strategic narrative to demonstrate progress, however, an emphasis was placed on producing better quantitative measurements. The Red Team challenged the excessive use of quantitative metrics and a number of assumptions, including those focusing on the insurgency and the levels of violence, and suggested a more narrative approach to documenting progress.

2.4.2 Hedging Among Afghans

In 2012, ISAF identified "hedging" among Afghans, i.e., individuals providing qualified support to a number of options to reduce personal risk, as a major threat to the transition to Afghan security. Due to hedging, Afghanistan was losing some of its best educated and wealthy members of society, while at the same time it was sending a negative signal to the international community, which was being asked to make commitments post-2014. A Red Team study on Afghan hedging was a good example of the progress made in taking a more nuanced approach to measuring and understanding Afghan attitudes. The study was a multi-part analysis using Afghan regions as the unit of study to understand Afghan attitudes and responses to the course of the campaign and the looming departure of the ISAF coalition. COMISAF characterised Afghan hedging activities as direct evidence of uncertainty regarding the future, but increasingly believed that the short-term solution was political [21]. Hedging became shorthand in media and even headquarters' discourse to suggest Afghans' lack of faith in their country's future following the withdrawal of coalition forces. However, the concept was not well understood, nor measured. The Red Team studies of the concept and regional case

studies helped to clarify data collection, definitions, and assumptions, the importance of cultural sensitivity, need to avoid simple linear relationships, as well as the impact on definitions of indicators [22], [23], [24].

2.4.3 Understanding the Impact of Factionalism

This project assessed and challenged the prevailing view that ethnic factions, particularly those with military forces, were a threat to the Afghan government's survival. The Red Team argued that the nature of factions and their level of participation in existing political system were better indicators than the mere fact of their existence. Indeed, how to engage diplomatically and politically the insurgents proved a challenge throughout this period. The theory behind this engagement was that the insurgents represented a part of the body politic. This was incrementally realized through 2012 as the headquarters examined insurgent responses to the campaign as well as the announcements to end the mission. For example, an assessment to determine the likelihood that the insurgency might stop operations in response to ISAF 2014 drawdown challenged perceptions of the insurgent leadership's potential response. Fundamental assumptions about the utility of trying to divide the Taliban while concurrently trying to negotiate were also addressed [25], [26].

2.4.4 Comprehensive Cross-Border Strategy Responses

The Red Team also contributed to campaign assessment by helping COMISAF understand better the impact of cross-border incidents. One of the main issues through 2012 was the growing number of incidents of cross-border artillery fire between Pakistan and Afghanistan, as well as the question of border regulation and interdiction activities. During the period of 2012–2013, the Red Team produced several evaluations of possible unintended consequences of attempts to better police the border. The analysis showed that corruption and patronage were complicating factors in coalition efforts to improve control of the border, particularly given patronage networks that extended from border guards to government officials. Red Team and NATO SCR analyses, among others, suggested that, among many factors in enforcing border regulations, these complicated patron-client relationships needed to be understood as a political as well as a legal issue [27]. It took time to fully understand the traditional social affiliations, the dynamics of political and economic relationships, and balances of power within them, as well as the respective roles of individuals in the context of the broader national political settlement. Nevertheless, the shape of proposed border strategies in general remained a divisive issue between military and civilian leadership [28].

Cross-border fires also reflected divisions between ISAF and Afghan political and military leadership as well as a situation where General Allen's command of U.S. Forces-Afghanistan (USFOR-A) was in potential conflict with his command of coalition forces (while the U.S. has officially recognized the Durand line, the border between Afghanistan and Pakistan, protection of Afghanistan's border was not part of ISAF's mandate). However, Afghan refusal to recognize the Durand line, which divided the Pashtun people between Afghanistan and Pakistan, was an easy way for the Afghan government to generate popular support among Pashtuns and any border incident would be handled aggressively. Since ISAF retrograde operations were vulnerable to tensions with Pakistan, the issue continued to unsettle senior decision makers through 2013 as they were forced to mediate Afghan-Pakistani border relations on multiple occasions [29]. The Red Team provided a historical perspective of the border dispute and recommendations of how to manage it [30].

2.4.5 Regional Analysis

One area, where the Red Team was able to complement significantly AAG's assessment work was regional analysis. In fact, regional dynamics was a major consideration for HQ, ISAF planners. Through the course of 2012–13, the increased focus on Afghanistan's geo-strategic situation was also evident in the growing demand for analysis of regional relationships, their potential impact on Afghanistan, and what, if anything, ISAF or USFOR-A could do to influence those relationships. While President Obama had specifically directed that engagement with Pakistan become a U.S. priority, it was under Allen that HQ, ISAF turned to regional relations in real terms and relative to

other priorities. Assessments of regional powers and their impact on the mission were included in the campaign plans and there was a significant analytical requirement on regional dynamics issues. Since neither the AAG nor the other analytical groups were structured to provide deep, qualitative products on regional analysis, a substantial part of the burden fell on the Red Team.

One troublesome issue, for example, was Iran-Afghanistan relations. ISAF leadership and western policy analysts in general perceived Iran's objectives in the region as malign, therefore the focus of the Iran's engagement in Afghanistan was most often assessed through the prism of its support for the Taliban and other activities that may have presented a direct threat to the military campaign, or were damaging the relations between ISAF and the Afghan government [31]. COMISAF directed his analysts to focus on evidence of the latter [32]. Red Team analysis attempted to address what might be an evolution of Iranian interests and offered more comprehensive evaluation in view of the changing regional environment, in particular the Syrian conflict, and the broader Iranian objectives in Afghanistan, not simply those threatening ISAF efforts [33]. On the other hand, the growing cooperation between Russia and Iran in Syria raised concerns whether they would attempt to duplicate this approach in Afghanistan, which was also an issue addressed by the Red Team. The Red Team examined issues such as the security environment, the energy sector, and foreign trade to understand the effects of joint policies in Russia's and Iran's overlapping areas of interest and possible areas of common influence in Afghanistan [34].

Analysis of Pakistan, technically considered a major U.S. ally in the region and a recipient of significant military and economic aid from Washington, was another challenge. Pakistan played a double-sided game since the Taliban was potentially one of its instruments in post-ISAF Afghanistan and did little to prevent the insurgents from using its territory as a safe-heaven. From 2009, the U.S. administration began to consider the Afghan-Pakistani border area as a single theatre of operation and introduced the term "Af-Pak" [35]. Given the new strategic focus on Pakistan's role in solving the Afghan quagmire, Pakistan was increasingly prominently in HQ, ISAF analytical efforts. It is notable that as late as 2013, HQ, ISAF formulated its Af-Pak strategy by asking for assessments of Pakistan's views towards Afghanistan [36]. One of the weaknesses of the Af-Pak strategy was it failed to appreciate the larger regional context for Pakistan, namely the role of India in formulating Islamabad's policy towards Afghanistan [37], thus ignoring the wider regional context. The Red Team paper "Indian Military Aid to Afghanistan" tried to address this gap in analysis [38].

2.4.6 Measuring Success

On occasions the Red Team was asked to construct strategic assessment scenarios of the campaign using qualitative analysis. One of the products was dedicated to "what winning looks like" and the other to "what losing looks like." The Team used NATO and U.S. strategic documents for Afghanistan [39], [40] as a starting point, and timelines looking at one year, five years and ten years respectively to create three distinct scenarios for success and failure. The products were grid-like tables describing the different conditions of winning and losing for each strategic objective in the established timeframes. The products facilitated mission planning and suggested appropriate indicators for tracking the different scenarios.

2.5 CONCLUSION

This chapter demonstrates that operation assessments are not confined to formal assessment staff groups, nor they are as centralised as previously thought. Commanders realized the limitations of centralised assessment products, based primarily on quantitative metrics, and sought other sources to augment these products. Decision support Red Teaming is one such capability that can complement operations assessments. Based on the authors' practical experience, Red Teaming, specifically in its role in supporting operations assessments, could be defined as: an independent analytical capability in direct support to the commander to challenge

planning assumptions and provide qualitative alternative analysis in the context of the strategic and operational environments from multiple perspectives. Such capability is complementary, but autonomous to planning and assessment staff.

It is clear that for Red Teaming to be successful, the team needs to be organizationally positioned not to compete with other analytical groups, and be as independent as possible from other staff organizations. On the other hand, the limited distribution list prevented other assessment organization benefitting from the Red Team products. The latter were only integrated into the decision-making process at the higher echelons of command. Whether this is a successful model is debatable. Sometimes even a physical co-location of different analytical groups can help facilitate better alignment between assessment products. For example, a small Red Team cell, or an individual team representative co-located with dedicated assessment staff can provide a linkage and better awareness of efforts to measure campaign progress at a lower level and better provide for effective campaigning.

2.6 AUTHORS' BIOGRAPHIES

Dr. Anton Minkov holds a PhD in Islamic history from McGill University, and has written extensively on the Soviet experience in Afghanistan, indicators for political instability in the Middle East and North Africa (MENA) and other security issues in the MENA region. In 2012, he received the NATO Scientific Achievement Award for developing security transition metrics for Afghanistan. In 2013, he served in the COMISAF Red Team. Currently, he is a member of the Advisory Cell to the Assistant Deputy Minister (Science and Technology) at the Department of National Defence, Ottawa.

Dr. Paul Dickson is a Strategic Analyst with Defence Research and Development Canada's Centre for Operational Research and Analysis. He holds a PhD in military history. His book *A Thoroughly Canadian General: A Biography of General H.D.G. Crerar* (University of Toronto Press, 2007) was awarded the 2008 C.P. Stacey Award for the best book on the study of conflict and society. He served as a senior analyst, and then Director of the COMISAF Red Team, 2012 – 13, and was awarded a Meritorious Service Decoration for his service in Afghanistan. Currently, he is posted to the Canadian Joint Warfare Centre, Ottawa.

DISCLAIMER: The opinions expressed in this article are those of the authors and do not represent, or otherwise reflect, any official opinion or position of the Government of Canada, or any of its departments and agencies.

2.7 REFERENCES

[1] The U.S. Army and Marine Corps. 2007. Field Manual (FM) 3-24, Counterinsurgency. Chicago: The University of Chicago Press.

[2] Clancy, J. and Crossett, C. 2007. Measuring Effectiveness in Irregular Warfare. Parameters, 37:2 Summer.

[3] Schroden, J. 2011. Why Operations Assessments Fail: It's Not Just the Metrics. Naval War College Review, 64:4 Autumn.

[4] Downes-Martin, S. 2011. Operations Assessment in Afghanistan is Broken: What is to be Done? Naval War College Review, 64:4,9 Autumn. Available from: https://digital-commons.usnwc.edu/nwc-review/vol64/iss4/9.

[5] Connable, B. 2012. Embracing the Fog of War: Assessment and Metrics in Counterinsurgency. Santa Monica, CA: RAND Corporation.

[6] Mushen, E. and Schroden, J. 2014. Are We Winning? A Brief History of Military Operations Assessment. CAN, Occasional Paper.

[7] Joint Staff. 2011. Commander's Handbook for Assessment Planning and Execution, Version 1.0. Suffolk, VA.

[8] Connable, B., Perry, W.L., Doll, A., Lander, N. and Madden, D. 2014. Modeling, Simulation, and Operations Analysis in Afghanistan and Iraq Operational Vignettes, Lessons Learned, and a Survey of Selected Efforts. Washington, D.C.: RAND Corporation.

[9] Commander NATO International Security Assistance Force (COMISAF). 2009. Commander's Initial Assessment.

[10] International Security Assistance Force (ISAF). 2009. ISAF COIN Advisory & Assistance Team (CAAT). Power Point Brief to NATO HQ: Brussels, 29 October.

[11] Kilkullen, D. 2010. Counterinsurgency. Oxford University Press.

[12] North Atlantic Treaty Organization (NATO). 2014. Chicago Summit Declaration on Afghanistan, Issued by the Heads of State and Government of Afghanistan and Nations contributing to the NATO-led International Security Assistance Force (ISAF).

[13] Schroden, J., Thomasson, R., Foster, R., Lukens, M. and Bell, R. 2013. Operations Assessments at ISAF: Changing Paradigm. Innovation in Operations Assessment: Recent Developments in Measuring Results in Conflict Environments. Norfolk, VA: NATO ACT.

[14] Dunford, J. 2014. Statement of General Joseph F. Dunford Commander U.S. Forces-Afghanistan before the Senate Armed Services Committee on the Situation in Afghanistan. United States Senate, 12 March.

[15] Sandoz, J. 2001. Red Teaming: A Means to Military Transformation. Alexandria, VA: Institute for Defense Analysis.

[16] Gold, T. and Hermann, B. 2003. The Role and Status of Department of Defense Red Teaming Activities. Washington, D.C.: Defense Science Task Force.

[17] U.S. Department of the Army. 2005. Field Manual (FM) 5-0 Army Planning and Orders Production. Washington, D.C.: Government Printing Office.

[18] TRADOC G2. 2015. The Applied Critical Thinking Handbook.

[19] North Atlantic Treaty Organization (NATO). 2011. NATO Alternative Analysis (Red Teaming) Concept: Conceptual Framework for the Implementation of a Bi-Strategic Command Decision Support Tool.

[20] Schroden, J., Thomasson, R., Foster, R., Lukens, M. and Bell, R. 2013. A New Paradigm for Assessment in Counterinsurgency. Military Operations Research 18:3.

[21] Allen, J. 2014. General Allen Speaks about the Day the Afghan Campaign Almost Ended. Foreign Policy. 21 February.

[22] Dickson, P., Abbott, L. and O'Neill, T. 2012. Understanding the Impact of Hedging. Red Team Paper.

[23] Hall, W., Ivory, J. and Dickson, P. 2013. Hedging and Uncertainty in RC-North. Red Team Paper.

[24] Commander NATO International Security Assistance Force (COMISAF). 2014. Commander ISAF's Afghanistan Update, Winter 2014. ISAF.

[25] Red Team group authorship. 2012. Taliban Strategy: Alternative Courses of Action. Red Team Paper.

[26] Red Team group authorship. 2013. Should We Divide the Taliban? Red Team Paper.

[27] Dickson, P. and Berg, H. (CAPT(N). 2012. Comprehensive Cross-Border Strategy Responses. Red Team Paper.

[28] NATO Senior Civilian Representative (SCR). 2012. Discussion with the Red Team by NATO Senior Civilian Representative Ambassador Simon Gass.

[29] TOLO News. 2012. ISAF Mandate Does Not Allow Cross-Border Action. Katz, 29 October.

[30] Hall, W., Minkov, A. and Ivory, J. 2013. Afghanistan, Pakistan, and the Durand Line. Red Team Paper.

[31] Shanker, T., Schmitt, E. and Rubin, A.J. 2012. U.S. Sees Iran in Bids to Stir Unrest in Afghanistan. New York Times, April.

[32] Department of Defense (DoD). 2012. Report on Progress Toward Security and Stability in Afghanistan. 18 December.

[33] Minkov, A. 2013a. Reconsidering Iranian Engagement in Afghanistan. Red Team Paper.

[34] Minkov, A. 2013b. Russia-Iran Cooperation in Afghanistan: Myth or Reality. Red Team Paper, July.

[35] Obama, B. 2009. Remarks by the President in Address to the Nation on the Way Forward in Afghanistan and Pakistan. The White House: The Press Office.

[36] Hall, N. 2013. Pakistan's Strategic Goals in Afghanistan. Red Team Paper.

[37] Cohen, S. 2011. Failure in Af-Pak: How the U.S. Got It Wrong. The National Interest, 15 July.

[38] Terrell, J. 2013. Indian Military Aid to Afghanistan. Red Team Paper, August.

[39] North Atlantic Treaty Organization (NATO). 2008. NATO Afghanistan Briefing: Helping Secure Afghanistan's Future. Brussels.

[40] United States Government Accountability Office (GAO). 2010. The Strategic Framework for U.S. Efforts in Afghanistan. Washington, D.C.

APPENDIX 2-1: COMISAF RED TEAM MISSION

The following mission description was extracted from the Red Team Director Position Description and the Deployed Assignment Agreement between Canadian Joint Operational Command (CJOC) and the Defence Research and Development Canada (DRDC) strategic analysts.

COMISAF Red Team is expected to provide Commander ISAF (COMISAF), ISAF Chief of Staff (CoS) and ISAF Deputy DCOS Intelligence (and other senior NATO general officers) with an independent capability to fully challenge assumptions and provide alternative analysis of strategic-level and theatre-strategic-level policies, strategies, plans and operations in the context of the strategic and operational environment from multiple perspectives, including adversaries and others. These perspectives include those of the insurgency, Government of the Islamic Republic of Afghanistan, NATO Coalition partners and the Government of Pakistan, among others. The Red Team's core mission consists of examining policy, strategy and planning assumptions, adversary motivations, regional dynamics and broad underlying social, cultural and economic factors that might affect the outcome of operations. The Red Team will conduct this analysis in collaboration with the planning and intelligence staffs, subordinate headquarters, multinational intelligence agencies, as well as military and external experts.

The Director Red Team reports to the Deputy DCOS Intelligence and is accountable to COMISAF. The Red Team's specific responsibilities include:

1) Researching and writing alternative assessments using sound analysis and critical thinking techniques;
2) Developing and maintaining relations with Afghan and International, governmental and non-governmental entities and individuals to enhance the Team's analysis and to support the exchange of ideas, which may at times result in engagements and related activities;
3) Sharing expertise with other Red Team and CJ2 members to ensure the team has a strong grounding in Afghanistan and regional issues;
4) Participating in and contribute to relevant ISAF forums;
5) Guarding against groupthink, using sound analytical methodologies; and
6) Performing other related duties as directed by the Director COMISAF Red Team.

Chapter 3 – ASSESSING PROTECTION OF CIVILIANS IN MILITARY OPERATIONS

Alexander William Beadle, Guro Lien and Elin Marthinussen Gustavsen
Norwegian Defence Research Establishment (FFI)
NORWAY

ABSTRACT

Protection of civilians has emerged as an increasingly important task in NATO operations. Yet, there exists only limited guidance on how the achievement of this task may be measured. In this chapter, we explore how protection of civilians can be assessed more effectively and comprehensively. It outlines five generic approaches to measuring protection of civilians in military operations (casualty figures, civilian behaviour, perception of security, territorial control, and perpetrator capabilities). All are relevant to assess in any conflict, but the relevance of each approach will vary according to the particular type of threat civilians are faced with. Thus, the chapter uses seven different scenarios, ranging from genocide to insurgency, to suggest which of these generic approaches will be most relevant to assess protection of civilians in different situations.

3.1 INTRODUCTION

Assessing changes in an armed conflict's dynamics is critical in order to track the progress and effectiveness of any military operation. The purpose of assessment is to understand the situation on the ground in order to act, do the right things, and thereby achieve the operation's goals most effectively. This is not, however, an easy task. Ideally, the assessment should take into account all relevant factors that will reveal any changes in the environment relevant to the achievement of the objectives of the operation. Ultimately, it should also detect the causes of the change, or at least identify likely causes of change.

One consideration that has become increasingly important in contemporary military operations, including NATO's, is the protection of civilians. Successful protection of civilians is viewed as central to ensure local and international legitimacy and often necessary to accomplish the operation's overall objectives. Some of NATO's previous operations, such as in Kosovo (1999) and Libya (2011), had protection of civilians as the primary goal, whilst in others, such as in Afghanistan, it has been considered one of several important considerations. It has therefore become necessary to assess the degree to which the military operation is able to protect civilians and how it can be done more effectively.

There is no unified definition of protection of civilians. Protection of civilians means different things to different actors on the ground, including military forces, humanitarian actors and human rights groups. Protection of civilians can therefore be broadly understood as everything from physical safety and access to basic human needs to the enjoyment of human rights in a protective environment preserved by the host nation [1]. Protecting civilians from threats of physical violence is usually considered to be the main responsibility of military forces, as few other actors will be able to provide it and it may be a prerequisite for achieving other types of protection.

In 2016, NATO adopted its first policy on protection of civilians. Here, protection of civilians includes "all efforts taken to avoid, minimise and mitigate the negative effects that might arise from NATO and NATO-led military operations on the civilian population and, when applicable, to protect civilians from conflict-related physical violence or threats of physical violence by other actors" [2]. This also includes "the use of force, as appropriate, to prevent, deter, pre-empt, and respond to situations in which civilians suffer physical violence or are under threat of physical violence" [2].

In light of this policy, NATO developed a concept in 2017, which seeks to "translate" the policy on protection of civilians into a framework that will support the planning and conduct of operations. In doing so, it seeks to identify "what" needs to be considered, but not "how" such matters should be addressed. The concept divides protection of civilians into three different aspects: mitigating threats of physical violence, facilitating access to basic needs, and activities that contribute to a safe and secure environment in the long term. Although all of these aspects are essential to the protection of civilians, the primary role of NATO lies in mitigating and minimising threats of *physical* violence, as this is viewed as the most basic type of protection where military forces have the primary role to play. The physical aspect of protecting civilians is two-fold: It includes efforts to protect civilians from NATO's own actions ("collateral damage"), but also efforts to protect civilians from other actors who deliberately target civilians [3].

In this chapter, we explore the question of "how" physical protection of civilians can be assessed and how it can be addressed in different situations NATO may find itself in. The next section of this chapter describes five generic approaches to measuring protection of civilians from a military point of view. Next, we use seven different scenarios, ranging from genocide to insurgency, to suggest which of these approaches are most relevant to assess protection of civilians from fundamentally different types of physical threats. Examples of indicators are provided for each scenario.[1]

3.2 HOW TO MEASURE PROTECTION OF CIVILIANS

Assessing the degree of civilian protection is an important part of a comprehensive assessment of an operation. In order to measure the degree to which protection of civilians is being achieved, we have developed five approaches that can be used in any mission:

1) Civilian casualty figures;
2) Civilian behaviour;
3) Perception of security;
4) Territorial control; and
5) Perpetrator capabilities.

These approaches can be used to develop more specific metrics to fit the assessment of the different military operations.

3.2.1 Civilian Casualty Figures

In most operations, it is an explicit or implicit objective to minimise the number of civilian casualties. Recording civilian casualties will therefore be a useful starting point for measuring the degree to which an operation is successful in terms of protection of the civilians. This includes not only the number of civilian deaths, but also the number of civilians displaced and harmed in other ways (e.g., injured, arrested, tortured, or abducted). In many conflicts, perpetrator groups target civilians deliberately as part of their strategies, and assessors should track the numbers of civilians they injure or kill. Additionally, NATO forces must track the numbers of civilians injured or killed by all armed actors, including those by NATO forces. So, to get a complete picture of the casualties, we must categorise the numbers by the type of casualty and the actors responsible.

However, the issues of data reliability raise considerable concern when it comes to using casualty data for assessment. Accurate casualty figures are often difficult to obtain in an area of active conflict. Collecting accurate data are even more difficult because – in order to be truly valuable as a basis for analysis – the

[1] This chapter is based on Ref. [4].

data should include the location, time, and description of how the violence was perpetrated [5]. Assessors or headquarters can ensure necessary data collection, for instance, by publishing report formats that operators must use when encountering evidence of civilian casualties.

Casualty figures also tend to be highly contentious and politicised. During the Libyan civil war in 2011, the rebels initially estimated 50,000 fatalities, a figure eventually adjusted downwards to 4,700 rebel supporters and a similar figure for government supporters [6]. Alan Kuperman argues that during the initial days of the uprising Western media reported more than 2,000 deaths, while the Human Rights Watch has documented only 233 deaths in the same period [7]. In other conflicts, such as in Afghanistan, relatively reliable civilian casualty figures are provided by the United Nations Assistance Mission in Afghanistan (UNAMA), whose estimates has largely corresponded with other assessments, such as NATO's own. In Syria, however, the UN stopped counting the death toll in 2014 due to a lack of confidence in its own data and has only provided rough estimates in the order of hundreds of thousands since [8]. In sum, reliable casualty figures are often unavailable in the midst of ongoing conflicts, but estimates can provide an initial indication of the level of violence facing civilians. Thus, it is important to present the figures and assessment alongside with the uncertainty of the data.

Another relevant question is: Which particular group of civilians is under threat? When assessing protection of civilians on basis of civilian casualty figures, it is important to measure the number of casualties against the number of people in *the relevant targeted group*, and not the entire population in the area of operation. For example, prior to the genocide in 1994, Rwanda's population stood at close to 7 million. The two largest ethnic groups were the Hutu (who made up roughly 85% of the population) and the Tutsi (14%). By 1995, the population was reduced to around 5 million people. Around 800,000 people had been killed, which constituted 11% of the original population. However, the vast majority of these victims were Tutsi. In fact, it is estimated that as many as three-quarters of the minority Tutsi population were killed [9].

Therefore, there can be significant differences between the proportion of civilians killed in general and the proportion of victims from the targeted group of civilians. In other conflicts, such as tribal wars in South Sudan, even a relatively small number of people (e.g., hundreds) can represent a substantial percentage of the group's total population, and therefore, constitute a significant failure to protect civilians for a military operation. Both relative and absolute numbers should be taken into account, but must be carefully used in the assessment report.

3.2.2 Civilian Behaviour

The choices people make in everyday life tend to reflect the opportunities and restraints of their immediate environment, and violence against civilians is likely to alter the pattern of civilian behaviour. Observing civilian behavioural patterns over time can offer an effective and relatively cost-efficient approach to assessment. This can be done by obtaining information from after action reviews from patrolling units, statistics from aid organisations, or opinion surveys. If civilian activity is showing signs of normalisation, this could indicate a reduction in the threat against civilians. However, there are many factors which can alter the pattern of civilian behaviour, like social or religious events, and so on. The assessment must therefore be put into context and include such factors.

The presence of rebel groups who, for instance, ambush civilians along the road is likely to prevent people from daily activities, e.g., going to the market or church and children from going to school. In addition, rebel presence in the vicinity of rural villages may hamper the planting and harvesting of agricultural products, thereby restricting local supply and driving prices up. As such, the local population's attendance at social events (like in the market, church or school), and the cost of food, its availability, and the cost and selection of various goods may provide useful indicators of civilian security in the area.

The availability of relevant economic and social information is likely to vary considerably between cases. Sometimes national statistics or statistics from international organisations are available, but may not be timely. Therefore, to assess local conditions in rural conflict affected areas, information will most likely have to be collected from in-theatre NGOs, aid organisations, and the communities themselves.

3.2.3 Perception of Security

A third way of assessing the civilian security situation is to monitor the civilian population's perception of their own security, for instance by conducting surveys, focus groups, or field interviews. The perception of security could be monitored through direct questions like "How satisfied are you with the security situation in your area?" or indirect questions like "What is the biggest problem in your area?" In any case – direct or indirect questions, surveys, or focus groups – the findings should be interpreted with caution and be presented together with the limitations of the results.

Many factors will influence people's perception of the situation, such as the visibility and activities of military forces. In 2011, Oxfam International conducted an opinion survey in conflict affected areas of Province Orientale, North Kivu, and South Kivu in eastern Democratic Republic of the Congo (DRC) [10]. Civilians in these areas considered the UN peacekeepers to be a positive force provided they were present and engaged in active patrolling. Conversely, when peacekeepers did not patrol, the local civilians found them ineffective.

Civilians may also perceive security itself differently in certain operational environments, e.g., security regarding physical violence and security regarding economic and health issues. Careful wording of the questions is essential.

3.2.4 Territorial Control

Territorial control may be another indicator of civilian security. First, armed actors who have territorial control can more easily harm and exploit the population, especially when perpetrators deliberately target civilians in order to control them. In these situations, perpetrators may use violence to eliminate opponents and deter cooperation with their enemies.

Second, the risk to civilians may also increase when territory changes hands. For instance, when the largely Hutu rebel group Democratic Forces for the Liberation of Rwanda (FDLR) retook areas lost during an offensive by Congolese and Rwandese security forces in 2009, it violently retaliated against civilians it considered government collaborators [11]. Since then, numerous shifts in territorial control have been associated with increased violence against civilians by many different actors in eastern DRC.

Usually, the intelligence branch will have a good overview of territorial control. Data can also be collected through an opinion survey.

3.2.5 Perpetrator Capabilities

Perpetrator capabilities refer to the ability of armed actors to inflict the type and level of violence that serves their goals. For instance, those who intend to commit genocide will require substantial preparation, coordination, and sufficient numbers of killers and weapons to achieve high death rates, while insurgents may only require simple explosives to achieve the intended destabilising effect. Thus, NATO's policy for the protection of civilians emphasises the importance of identifying the particular threats perpetrators pose, including "their motivation, strategies and tactics, capabilities, and the expected outcome for civilians" [2].

An assessment of the perpetrators' capabilities will not only measure the ability of a perpetrator to attack civilians today, but also estimate his *potential* to attack civilians in the future. A reduction in perpetrator capabilities will therefore involve a reduction in both the actual and potential threats to civilians. However, such assessments do not capture the *will* of the perpetrator, which derives from their motivation for targeting civilians. Comprehensive assessments should monitor both, especially as motivations may change and violence may escalate if perpetrators also possess the ability to do so.

Again, the operation's intelligence branch will usually have good information when it comes to a perpetrator's capabilities. However, enemy capabilities are often assessed against one's own forces, strengths and vulnerabilities, not those of the civilians.

The Norwegian Defence Research Establishment (FFI) has therefore identified five main categories of capabilities that may be critical for perpetrators *to attack civilians* and provide opportunities for an intervening force to degrade that ability [12], [13]:

- Freedom of movement. This relates to the means to move from one place to another with a low risk from enemy attacks. It is an operational requirement that all perpetrators require to attack civilians.
- Armed units and weapons. The required armed units and weapons will vary considerably depending on the perpetrator's strategy of violence.
- Advance planning. Preparations are especially important for actors who intend to initiate systematic and widespread violence against civilians. It is less critical to perpetrators of criminal, opportunistic, or indiscriminate violence.
- Top-down coordination. Political and military leaders often play a leading role in instigating and organising mass violence. For that reason, targeting Command and Control (C2) nodes and the responsible leaderships is likely to have a protective effect in many situations.
- Ambiguity. This capability concerns how necessary it is for perpetrators to maintain the support required from other actors in order to execute violence (e.g., members of the armed forces, ethnic groups), while concealing their criminal actions from those who may stop them (e.g., moderate politicians, the international community).

3.3 A SCENARIO-BASED APPROACH TO ASSESSING PROTECTION OF CIVILIANS

All of the five approaches described above are relevant to measure protection of civilians in most operations. However, civilians can be faced with fundamentally different types of physical threats depending on the conflict. Therefore, some aspects of information will be more important to assess depending on the *specific* threat to the civilians in one's area of operations.

In order to determine which of the five approaches are most relevant to assess protection of civilians in a particular conflict, it is necessary to understand *how* various perpetrators attack civilians. FFI has developed seven generic scenarios based on a number of historical conflicts, which seek to capture the range of situations where civilians face fundamentally different types of physical threats [12]. The scenarios are based on why perpetrators decide to attack civilians, the types of actors involved, the strategies and tactics they employ, the capabilities they require in doing so, and the expected outcome in each scenario, measured in terms of violence against civilians if perpetrators actually succeed.

If the violence against civilians typical of a particular scenario is being reduced, it suggests that a certain degree of success in terms of protecting civilians is achieved. Below follows a short summary of each scenario, with a description of the five approaches and their relevance to the specific scenario in order to assess whether efforts to protect civilians are succeeding.

3.3.1 Genocide

The gravest threat to civilians occurs when actors seek to *exterminate* a certain national, ethnic, racial, or religious group. Only states or militarily superior actors are likely to possess the means required to kill enough people to achieve this objective. Based on previous cases, such as the Herero, Jewish, and Rwandan genocides, this is the only scenario where the majority of potential victims are likely to be killed (>50%), if perpetrators succeed (see Ref. [12], p. 28).

The perpetrator's objective is to kill as many people as possible in the shortest amount of time. Therefore, the number of large-scale massacres of civilians will be the primary indicator, where a reduction would indicate a degree of success. Since a reduction might also be a result of diminishing number of potential victims, assessors must measure the casualty figures in relation to the remaining targeted population in the area.

By contrast, collecting data on changes in civilian behaviour or perceptions of security will have less relevance when harder evidence exists that shows civilians are already faced with an imminent, existential threat.

The perpetrator will depend on territorial control. Preventing or stopping attacks in areas where potential victims congregate in large numbers is likely to save many lives. Therefore, the perpetrator's territorial control and any shifts in the territorial control of these areas could be a useful indicator. Disrupting a perpetrator's capability to plan, execute, and coordinate multiple attacks on civilians is also critical to protecting civilians. Targeting the perpetrator's C2 infrastructure and armed units used to execute the violence on the ground could therefore reduce their capability to exterminate civilians in large numbers. The relevant units in this scenario will usually be militias, paramilitaries or special units rather than regular armed forces.

3.3.2 Ethnic Cleansing

A less deadly, but more frequent situation – ethnic cleansing – occurs when actors seek to *expel* a certain group from a specific territory. These perpetrators are also likely to be states or militarily superior actors, as territorial control and significant capabilities are preconditions for conquering and cleansing new territory. Here, perpetrators use demonstrative violence to coerce the targeted group into leaving and to prevent their return. Relatively fewer people will be killed in this scenario compared to genocide, but the proportion of victims who are displaced has historically been very high (~90%) (see Ref. [12], p. 33). Examples of this scenario include the violent breakup of Yugoslavia (1992–1995), as well as the ethnic cleansing of Albanians in Kosovo (1999) and the reverse cleansing of remaining Serbs in 2004.

As fewer people die, the number of civilians killed is a less useful indicator. The large numbers of civilians who flee or prepare to flee from specific areas will be the defining characteristic of the expected outcome in this scenario. Therefore, the expected civilian behaviour would be indicated by figures which measures the number of civilians displaced. Another relevant metric could be the number of homes destroyed in victim areas, as perpetrators tend to destroy property belonging to the targeted group to permanently prevent them returning. Perceptions of security may be a relevant measure during ethnic cleansing, especially early on. However, surveys and opinion polling may be impractical as events often unfold very quickly.

Territorial control through military superiority on the ground is a prerequisite for conducting ethnic cleansing, which makes it an important aspect to measure. Areas where the victim group is in the majority and in areas sandwiched between or surrounded by the perpetrator's ethnic brethren will be important to monitor, as this is where violence is most likely to occur immediately following the seizure of new territory [14].
Like in genocide, the units conducting the ethnic cleansing are most likely to be irregular forces or special units, such as interior police. Disrupting the freedom of movement for these units and their ability to coordinate attacks will be critical to prevent ethnic cleansing.

3.3.3 Regime Crackdown

A third scenario – regime crackdown – occurs when a government or de facto authority responds to threats against its own survival with *violent repression* of the population. Here, perpetrators do not primarily target civilians on basis of their ethnic or sectarian identity, but according to presumed or real affiliation with any opposition. The number of people killed or displaced will therefore vary according to the local level of fighting, with many of the dead being combatants rather than civilians (see Ref. [12], pp. 38-39). Recent cases of this scenario include the early phases of the Libyan and Syrian civil wars from 2011 onwards.

The numbers of civilians killed, displaced and injured may gradually reach high numbers because of the conventional heavy weaponry that government forces are likely to use against population centres. A key objective in protecting civilians in this scenario is therefore to reduce the number of civilian casualties in the most exposed areas, which will be the most basic indicator of the assessment.

During regime crackdowns, civilian behaviour will be a more ambiguous indicator with greater local variations, but it can offer some clues as to where in the area of operations civilians are under most immediate threat. Most violence will occur in areas where resistance is strongest. Similarly, civilian perceptions of security could potentially be a valuable indicator, whose trends over time may indicate the potential for escalation into more violent scenarios, such as ethnic cleansing or genocide, especially locally.

There is a real risk of retribution from all sides during shifts in territorial control in this scenario. In situations where government forces and rebels alternately gain control of a town or village, the civilian population may be seen as having collaborated with the enemy and be subject to revenge attacks from both sides.

Unlike genocide and ethnic cleansing, the units responsible for the majority of civilian casualties will be regular forces and heavy weapons, such as air forces, heavy armour, and artillery. Denying perpetrators access to and monitoring the status of these units and weapons will be highly relevant in order to assess both the current and future potential for violence against civilians.

3.3.4 Post-Conflict Revenge

A far less violent situation occurs in most post-conflict environments as former victims take *revenge* against previous perpetrators. During post-conflict revenge, the perpetrators will usually be individuals or loosely organised mobs seeking to settle scores on a personal basis [15]. This violence is more criminal than strategic in nature. Perpetrators select targets on basis of previous culpability, and violence is most likely in areas where most abuses have occurred before. Previous examples include targeting of Serbs in Kosovo post-1999 following the Serbian withdrawal and government officials in Iraq and Libya after the falls of Saddam Hussein and Muammar Gaddafi.

The absolute number of people killed will be comparatively few, due to the very selective targeting of individuals and the limited scale of violence involved (see Ref. [12], p. 43). However, even low levels of violence can prompt many to flee if the former perpetrators are associated with a specific group of people. Thus, the number of civilians subjected to criminal acts of violence (murder, arson, kidnapping, looting) and number of civilians that have fled will be important indicators. Changes in civilian behaviour and perceptions of security will also be relevant aspects to monitor in the absence of significant casualty figures.

Perpetrator capabilities are not critical in this scenario. During post-conflict revenge, violence is usually very low-scale. It may not extend beyond hand-held weaponry conducted by mobs, criminals, or mere individuals. Therefore, the freedom of movement for individuals or mobs seeking revenge will be the principal (and perhaps only) indicator looking at the perpetrator's capabilities. Furthermore, the danger of retribution increases with shifts in territorial control and in chaotic post-conflict environments where no one is in control, as it is what provides the opportunity for individuals and mobs to settle personal scores.

3.3.5 Communal Conflict

A potentially violent situation occurs when whole communities engage in seemingly endless cycles of violence, motivated by a combination of *revenge* and *self-protection*. Because both sides organise themselves along shared communal identities rather than as organised armed actors, they are unlikely to possess the means required to settle conflicts permanently. However, they cannot afford *not* to retaliate, as they believe this will invite further attacks upon themselves [16]. Civilians are primary targets for both sides, as the both factions assume the role of perpetrator and see themselves as victims. A recent example of this scenario is the conflict between Christian and Muslim communities in the Central Africa Republic (CAR) since 2013 [17].

As in post-conflict revenge, the total number of people killed during communal conflicts may not be very high, but it may constitute a relatively high proportion of each community's population size. The number of retaliatory attacks and casualties during each cycle may therefore be a more useful indicator of civilian protection than the absolute number of civilians killed.

Civilian behaviour is another important indicator in this scenario. Although the level of violence may be low, it may persist for long periods. Thus, a useful indicator of the security situation and civilian protection could for instance be the prevalence of trade at local markets.

The perception of threat amongst both communal groups is also important to monitor, as fear of being attacked is a principal motivation for both sides for launching attacks. If both the number of retaliatory attacks and the perception of threat among both communities are declining, a primary cause of communal conflict is being mitigated. This could be achieved through military presence that denies both sides the opportunity to attack the other. However, if threat perceptions only decline for *one* of the communities, this may indicate a lack of protection of the other community or that one community is gaining the superiority required for more decisive violence. There are several previous examples of communal conflicts escalating into ethnic cleansing, as happened when Christian communal militias gained military superiority in predominately Muslim areas during the aforementioned conflict in CAR.

Shifts in territorial control may occur during larger attacks. Permanent seizure of neighbourhoods populated by the other community is more common in mixed urban areas where communities reside closely together, as happened with Muslims areas in the capital of Bangui, CAR. In rural areas, where communities often reside further apart, a group's withdrawal to own areas following an attack is more common, as seen following large-scale attacks by tribal militias in South Sudan. In either case, the community that loses territorial control is likely to be at severe risk in these areas. Thus, monitoring territorial control and shifts in territorial control is relevant in both situations.

Perpetrator capabilities will be critical during the most deadly forms of communal conflict. Here, political and military leaders often play a principal role in instigating and organising violence. Monitoring their statements may provide important indicators of the potential for escalation. This sort of violence is usually conducted with basic weaponry and by loosely organised militias, which can make it difficult to monitor their military capabilities.

3.3.6 Predatory Violence

In weak states, actors may prey on the local population simply to ensure their own survival or for economic profits. In this scenario, these actors are typically rogue security forces or rebels who have failed to achieve their political objectives, but refuse to demobilize or disarm [18], [19]. Often physically removed from the geographic areas where they may gain popular support, they have few incentives to limit violence against civilians. Instead, they rely on pillage, forced recruitment, illegal taxation, lootable resources, and labour exploitation [20]. All civilians and humanitarian actors are potential victims. A

classic example of a predatory perpetrator is the Revolutionary United Front (RUF), which ravaged Sierra Leone from 1991 to 2001.

Predatory violence will not necessarily lead to many civilian deaths. However, large, sudden refugee flows can occur due to the brutality and unpredictability of attacks. Indicators in this scenario could be the numbers of civilians affected by violence (killed, injured, displaced, and abducted).

Civilian behaviour and their general perception of threat can also provide valuable information in this scenario. If civilians seem to go about their daily life without fear of attacks or harassment, protection efforts are probably successful. In addition, geographical analysis of the perception of threat could provide a scope of the threat.

There is a real danger to civilians associated with changes in territorial control. This is because predatory rebels are unlikely to hold any ambitions of eventually gaining popular support from the population. This may be reflected in dramatic peaks in the number of people killed (e.g., during the RUF's operation "No Living Thing" in Freetown in 1999) or in terms of the brutality used to scare the population (e.g., the practice of cutting off ears, lips, and noses employed by the Lord's Resistance Army in the central and eastern African borderlands).

In this scenario, perpetrator capabilities are less relevant to measure. In some cases, however, perpetrators conducting predatory violence rely on orders from centralised leaderships. In these cases, targeting rebel leaders and their Command and Control (C2) nodes can degrade their ability to attack civilians on a large scale.

3.3.7 Insurgency

The final scenario – insurgency – describes the situation where armed groups, fighting over political power, target civilians as a tactic. Although government forces or rival armed groups are the primary targets, insurgent groups tend to employ a combination of indiscriminate attacks against civilians to destabilise the security situation, as well as selective violence to prevent civilians from collaborating with the enemy [21]. An example of this type of situation is the violence perpetrated by the Taliban in Afghanistan today.

The proportion of people killed or displaced during insurgency is generally lower than in other scenarios, due to the less destructive means involved and because insurgents do not seek to exterminate or expel, but to govern the population. However, it could still be useful to assess the overall number of civilians killed, injured and displaced over time. For instance, in Afghanistan, NATO successfully reduced the share of civilians killed by international and Afghan government forces from 39% in 2008 to 11% in 2013, with a subsequent increase to 17% in 2014 [22]. Yet, the total number of civilian deaths actually increased from 1,523 in 2007 to 3,699 in 2014 due to increased targeting by the Taliban and other armed groups [23]. This "gap" in protecting civilians in Afghanistan was one of the reasons why NATO's policy includes protection of civilians from other actors' actions in addition to efforts to protect civilians from NATO's own actions.

As in regime crackdowns, the physical threats to civilians in insurgencies are likely to vary significantly at local levels. For instance, in the Faryab province of Afghanistan, surveys have shown that unemployment has consistently been the primary concern of the local population, with lack of security mentioned alongside lack of electricity and water, poverty, illiteracy, and poor roads [24]. At the same time, on the national level, insecurity was the biggest problem [25]. There were geographical differences *within* the province as well. For instance, the Faryab Survey showed that in the more insecure areas 72% were dissatisfied with the security situation while in the calmer areas 41% were dissatisfied. Thus, civilian perceptions of the security situation in this scenario will likely vary in the different geographical areas and should be assessed in conjunction with other possible concerns that may be equally important.

As insurgents primarily target civilians for purposes of population control, there is also a real threat of retribution when territory changes hands. A common situation occurs when government forces capture a village or city only to withdraw, allowing rebels to retake it. If the insurgents view the civilian population as having collaborated with the government, retaliatory attacks are likely.

Another common characteristic of an insurgency is the use of explosive weapons, which causes a high proportion of civilian casualties in urban areas [26]. Improvised Explosive Devices (IEDs) are important killers of both government forces and civilians in this scenario. However, the specific types of weapons responsible for the majority of civilian casualties will be very context specific and likely to change over time as the insurgents adapt to the protector's counter-measures.

3.4 CONCLUSION

Assessing the degree to which a military force is able to protect civilians is an important part of a comprehensive assessment of any NATO operation. We have presented five approaches to do so:

1) Civilian casualty figures;
2) Civilian behaviour;
3) Perception of security;
4) Territorial control; and
5) Perpetrator capabilities.

All are relevant to assess in any conflict, but the relevance of each approach will vary according to the particular type of threat civilians are faced with. The approaches can be used to develop more specific indicators to fit the assessment of the different military operations.

We have discussed the relevance of the five approaches to seven different scenarios that vary from genocide to insurgency. If the conflict is comparable to one of these scenarios, the scenario can serve as a reference for evaluating what constitutes "success". For instance, protecting civilians from ethnic cleansing requires efforts that successfully limit the number of victims being displaced by perpetrators, who, in previous conflicts where protection has been unsuccessful, has been able to displace 90% of the targeted population. Success, however, depends on what is "good enough" and must be defined in relation to the specific conflict and its level of threat. In some cases, such as if one is faced with an imminent genocide, allowing or even facilitating all victims to escape may be the best option available, given the very deadly outcome of previous genocides, even though it in practice may amount to ethnic cleansing on the ground.

Importantly, the relevant aspects of information and the relevant criteria for determining success in protecting civilians will change in line with the threats on the ground. So, if the scenario changes, the type of threat civilians are faced with will also change, and thus also the approaches to assessing protection of the civilians. For instance, if the situation de-escalates from one-sided ethnic cleansing to a more even communal conflict, civilian perceptions of security will become more important to monitor, as this is a key driver of continued communal violence motivated by a revenge and self-protection.

Monitoring *civilian casualty figures* is a starting point for assessing protection of civilians in all operations. An important distinction can be made between scenarios where the civilians are killed or displaced intentionally by the perpetrators (genocide, ethnic cleansing) and scenarios where civilians are injured or displaced incidentally due to the presence of conflict (regime crackdown, insurgency). However, the important thing is not necessarily always to establish exact figures, which are often unobtainable, but to establish *trends* that help the decision makers understand the nature and direction of the conflict, so that efforts to protect civilians become more effective.

Civilian behaviour and *perceptions of security* are relevant aspects to monitor during post-conflict revenge, communal conflict, predatory violence, and insurgency, as these scenarios usually involve protracted, but lower levels of violence. By contrast, civilian behaviour is a less useful metric during genocide and ethnic cleansing, as these scenarios often unfold very quickly and with easily observable effects when most people flee or die.

Monitoring *territorial control* (and shifts in territorial control) is also relevant in most conflicts. Armed actors who have territorial control can more easily harm and exploit the population, and the risk to civilians also tend to increase when territory changes hands.

Perpetrator capabilities, especially the ability to coordinate attacks against civilians, are most important indicators to monitor during genocide, ethnic cleansing, regime crackdown, and the most deadly cases of communal conflict. In general, the more violent a perpetrator plans to be, the more capabilities will he require to implement that violence. Maintaining ambiguity will be critical for perpetrators who plan to commit extensive violence against civilians, but need to conceal their actions so that the victims will not escape (during genocide and communal conflict) or to prevent outside intervention in most instances. For non-state actors who are not able to formally oppose outside intervention, ambiguity is generally less important.

The ability to provide credible assessments of effectiveness is important in an age of global information warfare where military operations must manage the expectations of the local population and where a broad spectrum of media commentators will actively participate in forming public perception. The political leadership, international and domestic interest groups, and the general public will expect to see the effective protection of civilians during the operation.

The most important reason for conducting operations assessment, however, is that it can be an invaluable tool for commanders to adjust operations in light of developments on the ground to be more effective. This is critical when it comes to protecting civilians from violence, because unlike territory or systems of government, once a human life has been lost, it cannot be recovered.

3.5 AUTHORS' BIOGRAPHIES

Alexander William Beadle is a researcher at the Norwegian Defence Research Establishment (FFI). He received his BA in War Studies and MA in Conflict, Security & Development from King's College London. His research has primarily focused on how military force can be used to protect civilians from perpetrators of violence. He was the main author of a military planning and assessment guide for the protection of civilians in NATO operations, which has since been adapted for use in UN peace operations. He has also been involved in the development of NATO's policy and operational concept for the protection of civilians.

Guro Lien has been working as a senior scientist at the Norwegian Defence Research Establishment (FFI) since 2008. She has an M.Sc. in Comparative Politics from the London School of Economics and Political Science and a BA in Comparative Politics from the University of Bergen, where she also studied English and History. Her research has been on international operations, especially Security Sector Reform and Military Assistance. Lien has also worked in operations research, and has among other things, developed a methodological framework on how to measure effects from security sector reform initiatives. She has published a number of reports, articles, and book chapters on these topics.

Elin Marthinussen Gustavsen is a principal scientist at the Norwegian Defence Research Establishment (FFI). She has a Cand. Scient. degree in astronomy from the University of Oslo, and has worked as an operational analyst at FFI since 2006. She is a veteran with one deployment as an operational analyst to Afghanistan. After the deployment, her research has mainly been on assessment, in particular on a semi-annually opinion polling conducted in Faryab province in Afghanistan. She has amongst others

developed a framework for progress assessment in Faryab province. She is currently supporting the Norwegian Defence Staff in the assessment process.

3.6 REFERENCES

[1] Beadle, A.W. and Kjeksrud, K. 2014. Military planning and assessment guide for the protection of civilians. FFI-rapport 2014/00965. Kjeller: Norwegian Defence Research Establishment.

[2] NATO. 2016. NATO Policy for the Protection of Civilians. Endorsed by the Heads of State and Government participating in the meeting of the North Atlantic Council in Warsaw 8–9, July 2016. www.nato.int/cps/en/natohq/official_texts_133945.htm?selectedLocale=en.

[3] Keenan, M. and Beadle, A.W. 2015. Operationalizing Protection of Civilians in NATO Operations. Stability Journal: International Journal of Security & Development, 4(1):1-13.

[4] Våge, A.S. and Beadle, A.W. 2014. Assessing protection of civilians in military operations. FFI-rapport 2014/00966. Kjeller: Norwegian Defence Research Establishment.

[5] Beswick, J. and Minor, E. 2013. Casualty Recording as an Evaluative Capability: Libya and the Protection of Civilians. In Hitting the Target? How New Capabilities are Shaping International Intervention, edited by Michael Aaronson and Adrian Johnson, pp. 65-78. RUSI: Whitehall Report.

[6] Black, I. 2013. Libyan revolution casualties lower than expected, says new government. The Guardian, January 8. http://www.theguardian.com/world/2013/jan/08/libyan-revolution-casualties-lower-expected-government [Accessed date: January 26, 2016].

[7] Kuperman, A.J. 2013. Lessons from Libya: How Not to Intervene. Harvard Kennedy School Belfer Center Policy Brief. http://belfercenter.ksg.harvard.edu/files/Kupermanpolicybriefpublishedversion2.pdf [Accessed date: February 17, 2016].

[8] Foreign Policy. 2016. U.N. Envoy Revises Syria Death Toll to 400,000. 22 April 2016. http://foreignpolicy.com/2016/04/22/u-n-envoy-revises-syria-death-toll-to-400000/.

[9] Meredith, M. 2005. The State of Africa: A History of Fifty Years of Independence. London and New York: Free Press.

[10] Oxfam International. 2011. We are entirely exploitable: The lack of protection for civilians in eastern DRC. Oxfam Briefing Note. https://www.oxfam.org/sites/www.oxfam.org/files/bn-protection-civilians-eastern-drc-2011-08-02-en.pdf [Accessed date: February 17, 2016].

[11] Clark, M. 2009. Hutu rebels in Congo strike back against joint offensive. The Christian Science Monitor, March 23. http://www.csmonitor.com/World/Africa/2009/0323/p04s01-woaf.html [Accessed date: February 28, 2014].

[12] Beadle, A.W. 2014. Protection of civilians – military planning scenarios and implications. FFI-rapport 2014/00519. Kjeller: Norwegian Defence Research Establishment.

[13] Våge, A.S. 2014. Violence against civilians – case-studies of perpetrators. FFI-rapport 2014/00520. Kjeller: Norwegian Defence Research Establishment.

[14] Melander, E. 2007. Ethnic Cleansing in Bosnia-Herzegovina, 1992–1995. Conference on Disaggregating the Study of Civil War and Transnational Violence. University of Essex.

[15] Boyle, M.J. 2010. Revenge and reprisal violence in Kosovo. Conflict, Security & Development, 10(2):189-216.

[16] Chagnon, N.A. 1988. Life histories, blood revenge, and warfare in a tribal population. Science, Vol. 239(4843):985-992.

[17] Øen, U.H. 2014. Protection of civilians in practice – emerging lessons from the Central African Republic. FFI-rapport 2014/01918. Kjeller: Norwegian Defence Research Establishment.

[18] Tranekaer, T. 2012. When Passion Dries Out, Reason Takes Control: A Temporal Study of Rebels' Motivations in Fighting Civil Wars. Working Paper Series 2012, No. 12-127. London: London School of Economics, Development Studies Institute.

[19] Metelits, C. 2007. Short-term Behavior Effects of Rebel Groups in Weak States: The Setting and the Model. Working Paper, No. 07-002, January 2007. Northwestern University: Buffett Center for International and Comparative Studies.

[20] Weinstein, J. 2007. Inside Rebellion: The Politics of Insurgent Violence. New York: Cambridge University Press.

[21] Kalyvas, S. 2006. The Logic of Violence in Civil War. Cambridge University Press.

[22] United Nations Assistance Mission to Afghanistan (UNAMA). 2014. Annual Report on Protection of Civilians in Armed Conflict, 2013.

[23] United Nations Assistance Mission to Afghanistan (UNAMA). 2015. Annual Report on Protection of Civilians in Armed Conflict, 2014.

[24] Marthinussen, E., Nordli, D. and Eggereide, B. 2014. Faryab Survey wave 8 – a year after the redeployment of Norwegian forces from Faryab. FFI-rapport 2014/00064. Kjeller: Norwegian Defence Research Establishment.

[25] The Asia Foundation. 2014. Afghanistan in 2014: A Survey of the Afghan People.

[26] Dodd, H. and Perkins, R. 2013. An Explosive Situation: Monitoring explosive violence in 2012 Action on Armed Violence. http://aoav.org.uk/wp-content/uploads/2013/06/An-Explosive-Situation-Explosive-Violence-in-2012.pdf [Accessed date: February 17, 2016].

Chapter 4 – OPERATIONS ASSESSMENTS FOR COUNTERING CLANDESTINE NUCLEAR THREATS

Robert L. Bovey, M. Anthony Fainberg (ret.) and James S. Thomason
Institute of Defense Analyses (IDA)
UNITED STATES

ABSTRACT

The authors argue that over the next decade the North Atlantic Treaty Organisation (NATO) must improve its ability to interdict a nuclear weapon or enough nuclear explosive material to construct an improvised nuclear device that is stolen from a Russian facility. They argue further that, in conjunction with analyses, simulations, and exercises and experiments within the mission area, operations assessments performed in other mission areas can contribute in important ways to learning how to make improvements in NATO's interdiction capabilities. Operations assessments can provide important insights to be used in working out the deployment of troops to barriers or search areas, equipping those troops in part from national equipment pools and in part specialised equipment from centralised caches, and then providing necessary support and command and control.

4.1 BACKGROUND

Operations assessments have been used to assist in search operations since at least World War II when they were applied to anti-submarine warfare [1]. The results have evolved to handbooks for conducting a wide variety of searches including those to rescue distressed boaters and locate forest fires. This chapter proposes the use of operations assessment to assist in the mission area of Countering Clandestine Nuclear Threats (CCNT), which has become more acute recently and continues to evolve as discussed below.

Since 2005, one of the authors has led structured interviews of serving and former senior United States (U.S.) national security officials [2]. Consistently, in their formal risk assessments, these officials judged nuclear threats to the U.S. homeland to be of low probability, but of such great consequence as to present the greatest risk to the nation among all physical risks assessed, including conventional war. Informal comments by the officials focused on the dangers posed by clandestine nuclear threats, such as, the covert delivery of nuclear devices by non-state groups or agents of rogue states to a U.S. city. To our knowledge, no similar series of interviews have been conducted of the leadership of other NATO allies. However, this can and should be done, and it would be no surprise if CCNT is also high on their lists of concerns.

The primary means to address this threat has been fourfold. First – and the main focus at the head-of-state level – has been to eliminate, as much as possible, Weapons-Usable Nuclear Material (WUNM) and tightly secure whatever is left. Second, there have been consistent efforts to monitor commerce to detect smuggled WUNM, typically, by civil agencies worldwide at sea and air ports or border crossings. Third, police and intelligence services have probed the underworld for smugglers through so-called "sting operations" among other tactics. For example, in November 2015, the Centre for Public Integrity published a report [3] demonstrating that a significant amount of WUNM was stolen from a Russian facility in the early to mid-1990s and that samples of the material were captured in sting operations in 1999, 2001, and 2011. Fourth, NATO member government agencies, including their military special forces, have developed capabilities to use radiation detectors to locate WUNM in small areas and seize it.

The cessation of U.S. funding and removal of the peer pressure from on-scene U.S. personnel after 2013 has let Russian personnel slip into habits of lax nuclear security, making it very likely that the security of Russian WUNM (RWUNM) – nuclear devices not actually attached to a delivery vehicle or quantities of highly enriched uranium or plutonium sufficient to make an improvised nuclear device – will decline over the next decade. As a consequence, NATO members will gradually lose confidence in Russian security [4].[1] Assuming Russian government intransigence continues, the obvious way to maintain even the current overall level of global security is to strengthen defensive capabilities outside Russia. Building the requisite capability over the next decade will involve the actions of military as well as civilian agencies in each NATO country, and coordinating actions across several NATO nations and with Nordic, Caucasian, and Central Asian partners.

4.2 MISSION

Dealing with RWUNM that is out of government control, perhaps in the hands of criminals or terrorists, is the focus of CCNT. Each NATO country and NATO collectively, in their collective preparations for CCNT, must make themselves ready to accomplish the following tasks:

- Discover that RWUNM is, or very likely is, loose. This can happen as a result of an intelligence operation or a warning from a station in the steady-state global nuclear detection system. For example, a border guard may recognise that RWUNM is passing his post but be unable to stop it. In either case, from the moment of warning that RWUNM may well be loose, its location becomes increasingly uncertain as time passes.
- Quickly strengthen existing barriers and establish new ones to limit the further movement of the loose RWUNM. These could involve strengthening border surveillance and control capabilities, establishing defensive rings around capitals and ports, and deploying tactical barriers across countries to capitalise on terrain features that might channelize smuggler movements.
- Search for a loose RWUNM in broad areas where the material may have been contained by barriers. “Broad areas” are those that are large enough that radiation detection plays only a small role in searching. For example, these areas could be cities, towns, fields, forests, mountains, and other kinds of terrain in different kinds of weather. The searchers will be using a wide range of Intelligence, Surveillance and Reconnaissance (ISR) and law enforcement tools to locate individuals, vehicles, and buildings that merit close examination through small area search techniques.
- Search “small areas”, which are those where radiation detection can play a major role in finding hidden RWUNM. Such areas may be identified in the course of searching broad areas or by other means.
- Seize, secure, and render safe the RWUNM and perform all the attendant tasks, such as transportation and forensics. These tasks may begin when a border guard succeeds in stopping RWUNM passing his post or after the other tasks listed above.
- For reasons of brevity – and in some cases classification – this chapter is confined to the second and third tasks: establishing and operating barriers and searching broad areas where the RWUNM may be. The authors submit that lessons from operations assessments in other mission areas can contribute significantly to improving how NATO allies and partners organise and equip themselves to perform these tasks.

[1] This was not an official intelligence conclusion by a NATO member, but rather the consensus of a team of experts convened in 2015 at the Institute for Defense Analyses (IDA) that included two of the authors, Bovey and Fainberg [4].

4.3 ASSUMPTIONS AND CONTEXT

This chapter assumes that intelligence collection and processing systems will have developed that very likely will be able to give prompt warning of a theft involving sufficient RWUNM to create a nuclear explosion [5]. It further assumes that adequate capabilities exist for searching small areas, where radiation detection can play a major role in locating hidden RWUNM, and once it is found, for seizing, securing, and rendering safe the RWUNM, plus all the attendant matters of transportation, forensics, and the like. In short, it assumes that we can focus on developing capabilities for broad area searches within containments, while others address the other components of the whole CCNT mission area.

How precisely NATO CCNT-focused organisations will evolve over the next decade has not yet been decided. Lessons learned from operations assessments in other mission areas will help shape that evolution. As a starting point for discussing operations assessment, it will be useful to think of the following notional elements evolving together [4]:

- In the army of each participating country, a company to battalion-sized unit, in addition to its other duties, trains and practices establishing and operating barriers and conducting broad area searches in cooperation with other government agencies.
- A group of 10 to 20 officers and non-commissioned officers in each country is devoted to planning and exercising CCNT barrier and search operations – and assessing them – in cooperation with similar teams in other countries and NATO headquarters. Building these groups likely will start in a few Eastern European member countries, for example, in those that border Russia, with support from larger member countries. The number of such countries likely will increase over time, based on the assessment of the early adopters.
- A centralised NATO support unit consisting of 50 or more specialists maintains a cache of specialised equipment that is especially useful in operating barriers and searching. The unit distributes the equipment to member state units for exercises and in crises. Small groups of these specialists accompany the equipment to help train and assess units in exercises and provide technical advice and assistance in crises.
- Lodged in a NATO headquarters is a unit of at least 6 to 12 military and civilian people who are responsible for guiding the effort generally, and more specifically, for designing and conducting experiments and operations assessments, performing tactical analysis of the results, and promulgating best practices and CCNT doctrine.

What elements are eventually established will likely be influenced by lessons of operations assessments in other mission areas. Furthermore, one or more of the elements that do evolve will use the knowledge gained through operations assessments to help them operate effectively.

4.4 PROCESS FOR DEVELOPING NEEDED CAPABILITIES

Operations assessments will be essential to developing and increasing the effectiveness of the capabilities needed for this NATO CCNT mission. Even before the mission area is established, operations assessments in other mission areas will help generate a knowledge base on, for example, how to operate barriers, conduct broad area searches, provide the logistics involved, and enable the Command, Control, and Communications (C3) capabilities that will be needed when conducting multi-country operations. These capabilities are all interrelated but also, to some extent, separable. For example, how people and equipment are moved to where a barrier is to be established and how they are to be supported there is separable from how the barrier is operated. Further, each element is decomposable into sub-elements that can be addressed individually, to some extent. The ability to decompose a CCNT operation allows lessons learned from the assessment of non-CCNT operations to be applied to similar activities that are part of CCNT operations. An example of this is searching. Figure 4-1 illustrates one decomposition of the search problem.

To a large degree, study of the subjects in the different sections of Figure 4-1 can be undertaken independently. For example, one can examine the fusion of information from multiple disparate sources, while treating those sources as "black boxes" that produce a particular range of possible outputs. Ref. [6] also shows examples of cases in which the results of such fusion were used to control sensor activities in real time. Potentially, an Intelligence, Reconnaissance and Surveillance (ISR) platform can contribute to a broad area search by identifying anomalies that merit investigation by people on the ground with radiation detectors.

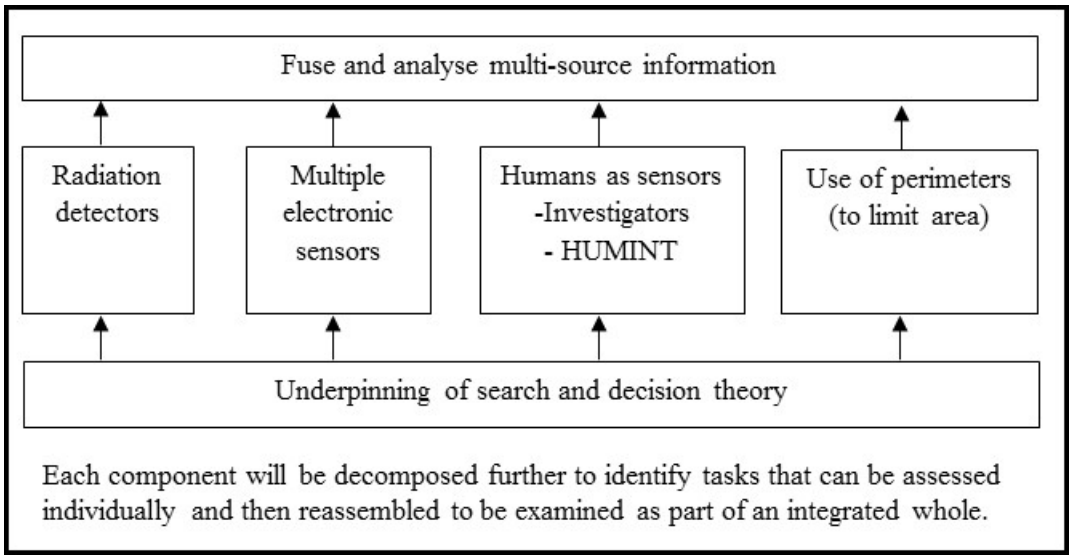


Figure 4-1: Decomposition of Challenges of Broad Area Search.

4.5 OPERATIONS ASSESSMENTS ROLE AND GUIDELINES

Assessment of barrier and search operations that do not involve nuclear materials are useful. What makes the CCNT barrier and search problem unique is that it involves RWUNM rather than some other contraband. Some forms of RWUNM very likely can be carried by one or two men along a sidewalk, across a field, or through woods. Other forms may require machinery to move. This leaves ample room for the smuggling of RWUNM to be like smuggling other contraband or the illicit movement of persons, and so assessments related to anti-smuggling activities or fugitive search activities can help improve the effectiveness of CCNT operations. For example, the June 2015 search for escaped convicts Richard Matt and David Sweat in northern New York State in the United States is a useful case study [7]. The escapees were known to have emerged from a manhole outside Clinton State Prison within a few hours. The subsequent search lasted three weeks and involved about 1,600 searchers and barrier operators. House to house searches were made in villages and a number of types of searches – on foot and by helicopters – were made in hundreds of square kilometres of wilderness. Perimeters were set up at various times to try to limit the travel of the escapees. That operation is sufficiently similar to what would be needed to find smugglers moving RWUNM to provide lessons useful for building a CCNT search capability. For example, although Matt was killed, in-depth interviews of Sweat – who was captured and remains in prison – provided lessons to improve the effectiveness of police actions in searches. In sum, many tasks are similar across mission areas so operations assessment measurements on tasks performed in other mission areas can serve as samples to frame developments of CCNT capabilities.

The discussion of the information to be addressed by operations assessments is divided into four sections: deployment to begin operations, operating a barrier, conducting a search, and providing support during operations.

4.5.1 Deployments

Given that uncertainty as to the location of loose RWUNM increases rapidly with time, it will be important to establish barriers and begin searches very quickly – within hours, not days – after the first warning that RWUNM is loose. Operations assessments will help identify the extraordinary measures needed to deploy troops and their equipment to barriers and searches so quickly, the locations of which will depend on the particular scenario. For example, establishing barriers in the Baltic republics involves great stretches of boreal forest and good transportation infrastructure; doing the same in the Caucasus involves expanses of mountains and modest transport infrastructure; and in Central Asia, it involves mountains and desert, and primitive infrastructure.

To establish barriers to contain loose RWUNM and to search the cordoned area, NATO will want to deploy several hundred people within hours within an eastern NATO country. They would require supporting equipment totalling less than about ten cubic meters and a couple tonnes from a central repository in NATO. In examining non-CCNT deployments – regardless of mission area – that are similar in size and urgency, the assessor should examine:

- Planning. What planning and pre-arrangement for such services as transportation are done and tested before the actual deployment?
- Times to assemble, deploy and begin operations. This should be supported by a narrative addressing distances, numbers of people and equipment, and costs: the how as well as the what. Specific matters to be addressed should include:
  - How do organisations maintain the continual capability to deploy within, say, 4 hours?
  - What formal agreements are necessary for host nations to provide transportation and billeting for arriving specialist support personnel?
  - Time to position and begin operation of supporting communications and ISR.

4.5.2 Barriers/Perimeters on Land

Recognising that the purpose of CCNT barriers is to intercept RWUNM contraband or at least delay it in the region within the barriers to allow searching to succeed there, assessors observing barrier operations in other mission areas, including police operations, should collect information pertaining to planning, balancing different aspects of the operations, tactical adjustments, and assessment of the barriers' effectiveness.

Planning specifies where to place barriers for maximum effectiveness. An example of planning a barrier was an analysis that demonstrated that only 89 check points are required to cover every possible vehicular route from any point more than 45 miles from Times Square in New York City into a 15-mile disk centred on Times Square, compared to the 708 road segments that cross the 15-mile radius boundary [8]. To analyse the road network around New York, the authors obtained a geographic information system database, "JServer", that contains and characterises all road segments in the U.S. The database includes the residential streets, urban alleyways, and other paths that vehicles can navigate; it is derived from hundreds of sources and is maintained and updated frequently. The authors drew the two rings and added a "super source" outside the 45-mile ring connected to all 722 crossings of it and a "super sink" inside the 15-mile radius line connected to all 708 crossings of it. They then assigned a "capacity" of one to each segment in the resulting network and employed the GNET^{TM} proprietary solver to determine a minimum cut set of the network using a maximum flow (from source to sink) linear programming algorithm. They then checked the solution for possible off-road bypasses using overhead imagery. Rail, maritime, and air approaches were addressed separately. The same methodology could be used for any city or port complex. It also could help plan checkpoint placement to limit movement of RWUNM from a known location. More generally, analysis – performed in advance, if possible – can help select effective means of tactical operations at checkpoints or along so-called green borders, which in this chapter mean those portions

of land barriers other than where checkpoints are operating because it is not a foregone conclusion that barriers deployed in Europe to contain RWUNM would conform to national borders; other lines might well be more effective.

Balancing the operation answers such questions as: How are tactical operations of disparate barrier segments tailored to the particular terrains/situations, such as road and railroad check points, possible aircraft flights over the barrier, and various green borders to include terrain passable by individuals walking, riding animals or driving all-terrain vehicles as well as rivers and lakes aligned with or across the direction of expected movement? How does the commander apportion capabilities across these situations in order to present a barrier to a smuggler that is uniform in the sense that any path forward is equally risky to him? Are other criteria used for barrier placement?

Tactical adjustments involve the movement or disestablishment of barriers and the management of this process. For example, knowledge of why barriers were moved or disestablished and insights into the impact of such decisions on effectiveness would be valuable.

The operation assessment of barrier operations would look for indicators of effectiveness, such as evidence that the barrier was penetrated. Such evidence would include a smuggler who is detained while attempting to cross, or a smuggler might be identified, but escape. (Then performance questions arise about a chase, which are not addressed here.) That a smuggler has passed undetected might be revealed if he is detained later, beyond the barrier, and reveals under questioning that he passed. It also would be useful information if a smuggler is detained before the barrier and under questioning reveals that he was deterred from proceeding by the barrier.

Even cases in which there is no specific evidence relative to penetration can assist in bracketing the likelihood of penetration when treated as one sample among many. They also can provide useful information on other aspects of barrier operations effectiveness. In check point operations, the operations assessor should report the speed of processing and experience of delays by members of the public including those who are erroneously detained as suspicious.

Operations assessments of the use of specific pieces of equipment also can be helpful. For example, NATO forces deployed in Afghanistan often used various systems to improve their perimeter defences. Two such systems were the U.S. Battlefield Anti-Intrusion System (AN/PRS-9 BAIS) and Lighting Kit Motion Detection (AN/GAR-2 LKMD) system. Presumably, these systems also will be useful in operating CCNT barriers along the green borders. The use of such systems in CCNT could be much more productive if the design of barriers using them can be informed by operations assessments of the same systems in other operations.

A descriptive narrative should be provided to both provide context and to suggest ideas as to what works in after action reports. Operations assessment professionals observing ongoing operations and interviewing participants in barrier operations can do so at border control points at international ports of entry. Here so-called Red Team tests, attempts to penetrate customs control points, are common. Many of these tests are performed to monitor operations for accountability rather than to measure progress and inform decisions [9], but the information contained in these assessments can also be used to make CCNT barrier operations more effective.

4.5.3 Broad Area Searches

Assessors observing search operations in other mission areas, including police operations, should also collect information such as that described in the following paragraphs.

One of the principal decisions in starting a search is whether to conduct it overtly or covertly. If conducted overtly, the search can take advantage of inputs from the general population. That said, many professionals

fear that information that RWUNM is loose would both cause panic in the general population and alert the smugglers that they are being sought. Proponents counter that the smugglers very likely assume they are being sought anyway; they also argue that the fact that a broad area search is underway will leak out in any case, so it is better to disseminate information proactively so as to control the message.

The correct decision will differ with the situation. Operations assessments in other fields may well contribute insights to creating a draft template for a decision aid that will help make the proper decisions in CCNT crises. Therefore, operations assessors should seek to document the procedures employed and public reactions in other potentially dangerous situations. So far, evidence on the value of involving the public is mixed. For an example arguing for openness, a 2009 contest to find ten 8-foot red balloons moored at randomly selected fixed locations in the U.S. was won in less than nine hours by a team employing internet-enabled crowd sourcing [10]. For another example, in the 2015 search for escaped convicts, Matt and Sweat, the public offered approximately 2,500 tips. In retrospect, following up at least two of these tips led searchers far afield. Two other tips, however, proved to be key to finding one of the fugitives [7].

Quantitative planning has been used to assist in search operations since at least World War II with the applications to anti-submarine warfare [1]. Planning aids have evolved to handbooks for conducting searches to rescue distressed boaters and locate forest fires, among other things. For example, coast guards have come to use simulations employing sophisticated models to plan searches at sea that allow planners to take into account ocean currents, as well as target and searcher features affecting the ability of a searcher to recognise the target [11]. These techniques will provide significant help to naval forces attempting to inhibit sea transport of RWUNM and will be of some help on land. For example, the principles of Bayesian search call for:

1) Overlaying the terrain to be searched with a grid and estimating a “prior” probability distribution of the likelihood that the smuggler is in each cell; and then
2) Searching in such a way as to pay the most attention to the peaks of that distribution;
3) Recalculating a “posterior” probability distribution based on what was not found; and
4) Searching again, and so on [12].

In the more complex, heterogeneous environments in which CCNT broad area searches on land are likely to take place, it also will be necessary to use a grid to select the basic search means and an estimated prior distribution to control their employment. For example, in cells within a town, local police conducting door-to-door inquiries may be the best approach. Aircraft with foliage penetrating sensors may be better initially for searching cells in nearby forests, with any people or vehicles discovered checked individually on the ground using techniques that include radiation detection. Assessments of search operations in other mission areas should document how these decisions are made in order to help develop rules for quickly determining prior probability distributions of loose RWUNM and selecting search modalities based on the specifics of the particular case. Specifically, an operations assessment of a search operation based on the Bayesian approach should document how the grid is laid down, how search means are selected for each cell, how the subjective prior probability estimates are made, and how probabilities are updated.

The fundamental measure of effectiveness of a search is its success in finding the target of the search. A very important secondary measure is the time required to find the target.

4.5.4 Support

As a starting point for operations assessments, the authors suggest assuming that support to maintain a series of barriers and to search will be needed for up to about a month. This includes operational (transportation, communications, command and control, intelligence analysis, etc.) and logistical (feeding, berthing, sanitation, health care, etc.) support. Examples of what support is needed are given below so operations assessors in other mission areas can identify comparable activities to assess.

4.5.4.1 Operational Support

Transportation for establishing searches is much like those for barriers, but there are important differences when an operation is underway. Maintaining a search operation is more demanding because the people being supported are moving over broad areas. Both will require local transportation to move searchers from place to place – and less frequently to redeploy barriers – and to supply both with daily needs. Transportation will be needed to rotate searchers and advisors between stations on barriers or search sub-areas and bivouacs 24 hours per day seven days per week.

Command and Control and Communications (C3) will be challenging. Establishing a series of CCNT barriers and searching areas very likely will involve the land territory of several nations and nearby waters. This will call for extensive information sharing and coordination. It will call for decisions as to how specialised equipment will be allocated among the participating forces. Further, interest in the CCNT barriers and searches will be intense at the highest political levels of NATO and member governments. Dealing with the demands for information will challenge the regional coordination or command centre responsible for facilitating the barrier operation and searches.

Public affairs will be critical to success. People will be curious about operations taking place in their midst, and uncontrolled rumours may inhibit operations. Operations assessors should document how the barrier and/or search commander communicated with the local population, and record judgements as to the reactions of the populace to the basic announcement of a search, tolerance exhibited by the population to the disruption caused by the operations over time, and utility of information provided to the population.

4.5.4.2 Logistical Support

Messing, continually both in the field and in bivouac, and berthing must be provided for local people and foreign advisors and assistants, such as intelligence analysts. Sanitation services will be needed in the field and in bivouacs. Medical care must be available to searchers and barrier operators in the field and medical evacuation services must be available for conditions too serious to be handled locally.

Operations assessments of all these activities in other mission areas – in addition to making them more effective in their own right – can provide valuable lessons and planning factors for the CCNT mission area.

4.5.4.3 Effectiveness of Support

Operations assessor should document how all of the types of support are accomplished including quantitative measures of the support relationships – such as litres of water per searcher per day, numbers of auto trips per day to rotate barrier operators, numbers of communication circuits of what kinds to provide command and control and intelligence to the field and receive support from abroad, and the like – to compare with rules of thumb for other operational situations and to better estimate the support requirements for CCNT barrier and search operations. Similarly, the degree to which each can be provided by local governments without the need to bring in supplies and equipment from abroad will be important to planning CCNT operations.

4.6 SUMMARY

Operations assessments can guide the development of NATO CCNT capabilities and increase the speed with which NATO can develop effective tactics, techniques, and procedures. This chapter focuses on obtaining tactical-level information to make CCNT operations more effective.

In the wrong hands, RWUNM can devastate a city. Therefore, the response to RWUNM known or thought to be loose, will be unlike any other operation. Nonetheless, operations assessments in other mission areas can provide valuable insights into how build the capabilities to operate CCNT barriers and conduct broad area CCNT searches.

Further, if a CCNT barrier and search operation is mounted, operations assessments will be an integral part. At the most basic level, a Bayesian search operation will start with a prior probability distribution and then employ operations assessments of the first searches to calculate a posterior probability distribution to modify future searches. More broadly, operations assessments can continually apprise the NATO senior leadership of the progress of the operation so that the leaders can make better decisions about larger operational and strategic questions.

There must be a small organisation in the NATO system devoted to making CCNT operations more effective. Previous sections of this chapter propose efforts that will generate an enormous amount of information, which, in turn, will guide future experimentation and aid planning and preparation for a crisis.

4.7 AUTHORS' BIOGRAPHIES

Dr. Bob Bovey has been a researcher at IDA since 2000, specialising in homeland defense. Earlier, he engaged in international business, first as a consultant, then as president of a nuclear transportation services company, and finally as an investor's representative starting businesses based on technologies in Russian research institutes. His prior experience was Navy and included serving as military assistant to two Secretaries of Defense, special assistant to a Director of Central Intelligence, and captain of a nuclear submarine. Bob earned a Ph.D. from The Johns Hopkins University in operations research and a BS from the U.S. Naval Academy.

Dr. Tony Fainberg retired in 2017 from the research staff of the Institute for Defense Analyses where he had served for over ten years, specialising in studies of countering terrorist use of weapons of mass destruction. For example, he led a capabilities based assessment of the domestic layer of the Global Nuclear Detection Architecture. Earlier he served in management and analysis positions in the Department of Homeland Security, Defense Threat Reduction Agency, Federal Aviation Administration, and in the Congressional Office of Technology Assessment. He earned his Ph.D. in experimental physics from the University of California, Berkeley, and performed research at Brookhaven National Laboratory before turning to systems and policy studies.

Dr. Jim Thomason leads the Strategy and Risk (S&R) Program at IDA, where he is also the Deputy Director of the Strategy, Forces and Resources Division. S&R builds risk management tools for DoD, DHS and others. Dr. Thomason trained in Political Science and International Relations at Harvard (BA, Honors) and Northwestern (PhD). He has published over 100 research articles, including in Parameters, Joint Forces Quarterly, Armed Forces and Society, Phalanx, and the Military Operations Research Symposium (MORS) Journal. Among recent studies, Dr. Thomason just completed IDA's analytic support for DoD (OSD) concerning Executive Order 13806 (on the U.S. Defense Industrial Base) and leads the strategic materials analysis program for DoD/DLA.

4.8 REFERENCES

[1] Morse, C.P. 1976. In at the Beginnings: A Physicist's Life. 1st ed. Cambridge, MA: MIT Press.

[2] Thomason, J.S. and Bexfield, J.M. 2017. A Method for Improving Strategic Decisions and Senior-level Teamwork in U.S. National Security Organizations, Part 2. Phalanx: The Military Operations Research Society, June.

[3] Birch, D. and Smith, R.J. 2015. The Fuel for a Nuclear Bomb is in the Hands of an Unknown Black Marketeer [sic] from Russia, U.S. Officials Say. Washington, D.C.: Center for Public Integrity, 13 November. www.publicintegrity.org.

[4] Bovey, R.L., Harvey, J.R., Fainberg, M.A., Heady, S.C., Bashaw, V.D., Wagner, R.L., Brooks, L.F., Celec, F.S., Cuda, D.L. and Dubin, R.D. 2015. Force Enhancement Packages for Countering Nuclear Threats in the 2022-2027 Time Frame, IDA Document D-5568. Alexandria, VA: Institute for Defense Analyses, September.

[5] Economist. 2016. Counter-terrorism: Shrinking the Haystack. The Economist, January 16, p. 86.

[6] Wilson, A.G., Cazares, S.M., Papadantonakis, K.M., Avery, M.R., Cartier, J.F., Dolph, P.A., Fregeau, J.M., Holzer, J.R., Morrison, K.A., Nunes, S.M., Renn, S.R., Snyder, J.A. and Spencer, K.M. 2013. Review of Methods and Algorithms for Dynamic Management of CBRNE Collection Assets. IDA Paper P-4995. Alexandria, VA: Institute for Defense Analyses, July.

[7] Clukey, K. 2015. Strategy, Luck Key in Hunt for Richard Matt, David Sweat. Albany, NY: Times Union, 12 July.

[8] Barnett, D., Sean, L., Anderson, B., Atwell, R.J. and Bovey, R.L. 2007. Application of the Maximum Flow Problem to Sensor Placement on Urban Road Networks for Homeland Security. Homeland Security Affairs III, (3):IV, September. www.hsaj.org.

[9] Bexfield, J., de Nijs, J., Farina, F.F. and Williams, A. (editors). 2013. Innovation in Operations Assessment: Recent Developments in Measuring Results in Conflict Environments. Norfolk, VA: Headquarters Supreme Allied Commander Transformation. p. 117.

[10] Ford, C.M. 2011. Twitter, Facebook and Red Balloons: Social Networking and Problem Solving in Homeland Security. Homeland Security Affairs, 7:III, February. www.hsaj.org.

[11] National Search and Rescue Committee. 2000. United States National Search and Rescue Supplement to the International Aeronautical and Maritime Search and Rescue Manual. Washington, D.C.: National Search and Rescue Committee, May. pp. 4-3.

[12] Stone, L.D.1975. The Theory of Optimal Search. New York, NY: Academic Press, Inc.

Chapter 5 – ASSESSMENT AND INTERPRETATION

Jan Frelin
Swedish Defence Research Agency (FOI)
SWEDEN

ABSTRACT

In this chapter, I outline how interpretative, or qualitative, methods may assist effective operations assessment. Initially I describe how hermeneutic interpretation is conducted, and how it may be used in a realist sense to get closer to the true situation in an operation. I then discuss how choice of assessment staff, critical analysis of sources, and triangulation can be used to increase the quality of an interpretive assessment. Finally I discuss what introducing these new approaches to assessment would entail, and make a comparison to a related military field, all-source intelligence.

5.1 INTRODUCTION

As indicated in Chapter 12, the experiences of using operations assessment in the field have been decidedly mixed [1]. In this chapter, I wish to investigate if a change in methods and paradigm would offer a better chance of effective operations assessment.

In the previous study [1], I found that in four cases, (Operation Desert Storm, Operation Allied Force, Operation Iraqi Freedom and ISAF), assessment failed to meet the expectations and requirements of operational commanders. Commanders then were left largely with their intuition; analytical products from assessment added limited value. In addition, I found that then-current methods, mostly based on performance measurement approaches, show limited theoretical hope of ever delivering the expected results in the context of complex operations [1].

The motivation to look for answers outside of quantitative approaches is mainly the observation that "the mathematics of complex systems warns us of the limitations of mathematics itself when considering such systems, de-emphasising the exclusivity of mathematics to analysis" [2].

5.2 QUALITATIVE METHODS

Qualitative inquiry is based on in-depth interviews, direct observations, or written communications. Qualitative methods address how to gather and interpret qualitative data and disseminate its findings. Qualitative inquiry can assist with illuminating meanings, studying how things work, capturing stories, elucidating how systems function, understanding context, identifying unanticipated consequences, and discover important themes and patterns across cases [3]. All of these research purposes may be relevant for operations assessment, but studying how things work, elucidating how systems function, and identifying unanticipated consequences directly addresses well-known challenges in the operations assessment field.

There are many flavours of qualitative methods (ethnography, realism, grounded theory, phenomenology, constructivism, narrative, semiotic, hermeneutic, and so on) (see Ref. [3], pp. 97-99). For the following discussion, I will concentrate on *hermeneutic interpretation* and *realism*.

Interpretation and hermeneutics was chosen because it focuses on the basic act in all qualitative inquiry, the interpretation of text (see Ref. [3], pp.136-138).

Within the qualitative research field there is a strong strand of relativism, essentially claiming that there exists multiple socially constructed realities, and that "truth" is the best informed construction on which there is consensus (see Ref. [4] p. 84).

Realism in this context refers to the assumption that world around us actually exists, and that meaningful observations can be made about it (see Ref. [3] pp., 111-113). Current strands of realist thinking acknowledge that while knowledge is socially and historically constructed, there are other ways of testing the validity of statements than merely consensus (see Ref. [5], pp. 35-36; also Ref. [3] p. 113). Not all statements are equally true, and it is possible to tell the difference. In the remainder of the chapter, I will discuss how to tell the difference when working with text rather than numbers.

5.2.1 Interpretation, Hermeneutics, and the Possibility of Realism

The following discussion on interpretation in the hermeneutic tradition is mainly based on Refs. [6] and [7]. Both are American scholars concerned with defending the capability to make justifiable claims of truth in what Mazlish calls "the uncertain sciences". Mazlish is a historian reflecting on his own profession, while Bernstein has been a professor of philosophy at Haverford College, among other things.

Both Bernstein and Mazlish base their treatments on Gadamer's approach to hermeneutics. The initial tenet of hermeneutics is that in the *Geisteswissenschaften,* the "spiritual sciences" or the human sciences (such as history), interpretation, rather than experimentation, is the basic method of inquiry. Interpretation requires a text, or something that can be treated like a text, and a reader who does the interpreting (see Ref. [7], p. 92). The text does not necessarily have to be in the form of the written word. Works of art, or observations of events (common in assessment), can be the subject of interpretation just as well (see Ref. [6], p. 141).

The purpose of interpretation is to create understanding, specifically understanding of a person, or a culture, or a time, which is different from that of the reader. If the culture of the text is completely alien to the reader, no understanding is possible, while if the culture of the text is similar to that of the reader, no interpretation is necessary. Interpretation thus occupies a middle ground, where there is some commonality between the reader and the text that enables interpretation, while there is still some distance that makes interpretation necessary (see Ref. [7], p. 92).

In the natural sciences, prejudice is seen as something hindering objectivity, something that has to be eliminated from the inquiry as much as possible. In Gadamer's approach to hermeneutics, "prejudice" (or preconceptions) is a necessary prerequisite for any understanding. Without prejudice, there can be no interpretation (see Ref. [6], p. 127; also [7] pp. 94-95). Interpretation thus is a dialogue between the understanding, prejudiced reader and the text.

This dialogue is frequently described as the *hermeneutic circle*, with the idea that the reader is going back and forth between different modes of reading. Characteristic modes of the hermeneutic circle are near-far or self-subject; the reader alternately considering herself and the text (see Ref. [6], p. 134 and p. 136; also Ref. [7] p. 91); the whole and the parts, alternately seeking overview or details; or facts and theory, shifting between realities and concepts of the text (see Ref. [7], p. 91).

Every individual and every culture is said to have a horizon, which limits what one can see and understand from a certain place. This horizon is limited, finite, changing and fluid. As the reader keeps moving in the hermeneutic circle, a *fusion of horizons* is achieved, where the horizon of the reader is enlarged and enriched (see Ref. [6], pp. 143-144).

In addition, Bernstein argues that interpretation forms the basis of human existence as a sensing and sense-making being (see Ref. [6], p. 113). In order to orient ourselves in the world, a constant process of interpretation is necessary. Thus hermeneutics is more than an approach for a specific task. As a

consequence of that argument, an element of interpretation forms part of all sciences, even the "hard" ones. In order for a scientific result to be presented – any kind of scientific result – a narrative interpretation is necessary (see Ref. [7], pp. 112-113). There is no specific algorithm for choosing one scientific theory over another; results will always have to be interpreted (see Ref. [6], p. 172).

As a result of the necessary subjectivity of the prejudiced reader, no two interpretations of the same phenomenon are likely to be similar. It is not possible to speak of *the correct* interpretation. That does not imply, however, that all interpretations are of equal merit. It is possible to *rank* interpretations, although that ranking has to be based on the judgement of other readers/interpreters (see Ref. [6], p. 125). Thus, the quality of an interpretation is basically ensured by other observers/interpreters (see Ref. [7], p. 127).

5.2.2 Making Credible Interpretations and Judgements

In order to increase the credibility of interpretive research, several options are available. In this section, the importance of the quality of the analysts, critical analysis of sources, and triangulation are identified as important strategies for increasing the credibility of interpretations.

5.3 CHOOSE INTERPRETERS WISELY

In all qualitative research, *who* does the research, or the interpreting, means a lot for the outcome of the process. In a qualitative inquiry, the researcher is the instrument. The credibility of the inquiry is closely connected to the credibility of the person or the team conducting the inquiry. As an inquirer conducting operations assessment, it is important to disclose one's world-views and predispositions and reflect over one's own outlook; as a commander, it is important to choose the assessment team for their insight, credibility, training, and variety of outlooks and skills (see Ref. [3], pp. 700-709).

5.4 CRITICAL ANALYSIS OF SOURCES

One way of increasing the credibility of interpretation, from a realist point of view (see above), is to apply the principles of critical analysis of sources (see Ref. [8], pp. 89-90; also Ref [9]). Critical analysis of sources mainly aims to clarify factual questions such as "What happened?" and "How did it happen?" (see Ref. [9], p. 7). Critical analysis of sources involves applying four criteria to any source for analysis: authenticity, concurrence, independence, and bias.

Authenticity means investigating if the source is what it says, that it is not a forgery. Forgeries have been reasonably common throughout history; well-known forgeries include the Turin cloth and the Protocols of Zion (see Ref. [9], pp. 12-14).

Concurrence means investigating if the source is concurrent with the events it recounts. Human memory is weak. Not only do we forget things, but there is also a propensity for adding details and embellishing stories as time goes by. Thus, diaries are usually considered good sources, but even diaries are prone to errors in recollection. Stories told many years after they occurred are even less trustworthy. The process of forgetting is not linear; however, a lot is forgotten immediately after an event, but memories then tend to stabilise. Singular events are usually prone to modification, while repeated activities are remembered long after they have ceased, such as daily working routines (see Ref. [9], pp. 26-33).

Independence means investigating if sources are related to each other. If two independent sources seem to agree on the outline of an event, that increases the credibility of that narrative, but if one of the sources can be shown to be dependent on the other, no such conclusion can be drawn. Critical analysis of sources differentiates between primary sources and secondary sources. A primary source, in Thurén's terminology, is usually an eye-witness account, while a secondary source is dependent on other sources. Sometimes concurrence and independence can

be in conflict though. A secondary source based on a concurrent eye-witness account may be more trustworthy than a primary source recorded long after the fact (see Ref. [9], pp. 35-48).

Other types of dependency are the repetitions of urban legends and folklore, which usually have no basis in reality(see Ref. [9], pp. 44-48), and so-called narrative contagion, which can occur between witnesses to an event, if they have the time to negotiate an interpretation between themselves. Another pattern is normative dependency, retelling a story in a form that it "should" be told. A heroic epic follows certain given elements, regardless of what actually took place (see Ref. [9], pp. 49-62).

Bias means investigating the tendency or interests of a certain source, what the strategic intent is behind a certain statement or text. Bias in a source might manifest itself as the telling of outright lies, as well as the careful selection of supportive evidence. Sometimes an actor may be unaware of his own bias (see Ref. [9], pp. 63-77).

5.5 TRIANGULATION

Another way of increasing the credibility of a qualitative inquiry is known as triangulation. The term is borrowed from land surveying, and is used metaphorically in qualitative inquiry to indicate that no single method ever adequately solves a problem (see Ref. [8] pp. 100-101; also Ref. [3] pp. 661-676).

Patton has identified four kinds of qualitative triangulation (see Ref. [3], p. 661):

- Triangulation of qualitative sources, checking out the consistency of different data sources within the same method. In hermeneutics, relying on more than one source.
- Mixed qualitative-quantitative methods triangulation. Checking out the consistency of findings by different data collection methods.
- Analyst triangulation. Using multiple analysts to review findings.
- Theory/perspective triangulation. Using multiple perspectives or theories to interpret data.

Both discovering consistency and inconsistencies among interpretations are important results of triangulation. Divergent results should not be cut off, but seen as possible cues on issues that needs to be explored (see Ref. [3], pp. 661-662).

5.6 THE ROLE OF INTERPRETATION IN ASSESSMENT

I have proposed new methods with the hope of improving assessment, and discussed their intellectual underpinnings. Former approaches to assessment have, as far as I am able to judge, not provided assessments that have helped operations getting more effective. The various tools I have suggested are all more or less established approaches to qualitative research. What I have proposed in this chapter would entail a change in both doctrine and manning of assessment teams, as well as an acceptance that qualitative approaches adds valuable knowledge to the operations process.

An interpretive approach to assessment has much more in common with all-source intelligence than with the current way of conducting assessment. As all-source intelligence is already an established part of military commands, there is a model to emulate. Assessment staffs would have to learn from intelligence officers, and adapt their methods to suit the assessment problem.

Interpretive assessment does not do away with quantitative methods. Opinion polls have been identified as providing helpful insight for assessment, that practice should form part of an "all-source assessment" approach [10], (see also Ref. [1], p. 54).

Such an all-source assessment process cannot be seen as a tested concept, but I believe that it would be worth testing these ideas for further development of the assessment field.

5.7 AUTHOR'S BIOGRAPHY

Jan Frelin is an Operations Researcher (OR) currently working at the Policy and Plans department of the Swedish Defence Staff. Previous assignments include OR positions at the Swedish Joint Command, the Nordic Battlegroup, the Swedish Command and Control Command, Swedish Ground Forces Command and at KFOR Headquarters in Kosovo. Frelin has also conducted evaluations of the Swedish Provincial Reconstruction Team (PRT) in Afghanistan as well as contributing to a NATO evaluation of ISAF. With a background in computer science, Frelin initially worked with Modelling and Simulation (M&S), but is currently focused on so-called 'soft' OR, current operations, and evaluative activities.

5.8 REFERENCES

[1] Frelin, J. and Norén, A. 2016. Approaches for Monitoring and Evaluation of PSO. FOI-R–4275–SE. Stockholm: Swedish Defence Research Agency.

[2] Richardson, K.A., Mathieson, G.L. and Cilliers, P. 2009. Complexity Thinking and Military Operational Analysis. In Knots, Lace and Tartan. Making Sense of Complex Human Systems in Military Operations Research. Edited by K. A. Richardson. Litchfield Park: ISCE Publishing. pp. 27-69.

[3] Patton, M.Q. 2015. Qualitative Research & Evaluation Methods. Fourth edition. Los Angeles: SAGE.

[4] Guba, E.G. and Lincoln, Y.S. 1989. Fourth Generation Evaluation. Thousand Oaks: SAGE.

[5] Allwood, C.M. and Eriksson, M.G. (editors). 2011. Vetenskapsteori för psykologi och andra samhällsvetenskaper. Lund: Studentlitteratur.

[6] Bernstein, R.J. 1983. Beyond Objectivism and Relativism: Science, Hermeneutics, and Praxis. Oxford: Basil Blackwell.

[7] Mazlish, B. 2007. The Uncertain Sciences. Second Edition. New Brunswick: Transaction Publishers.

[8] NATO. 2012. NATO Guide for Judgement-Based Operational Analysis in Defence Decision-Making: Analyst-Oriented, Volume: Code of Best Practice for 'Soft' Operational Analysis. AC/323 (SAS-087)TP/345. NATO.

[9] Thurén, T. 1997. Källkritik. Stockholm: Almquist and Wiksell.

[10] Banko, K. 2013. Public Opinion Polling in Operations Assessment with Examples from Afghanistan. In Innovation in Operations Assessment. Recent Developments in Measuring Results in Conflict Environments. Edited by A. Williams, J. Bexfield, F.F. Farina and J. de Nijs. Norfolk: NATO. http://www.act.nato.int/volume-1-innovation-in-operations-assessment-recent-developments-in-measuring-progress-in-conflict-environments.

Chapter 6 – APPLICATION OF HOLISTIC COMPLEX SYSTEMS APPROACH TO OPERATIONS ASSESSMENT

Chris Jordan
Niteworks[1]
UNITED KINGDOM

ABSTRACT

This chapter describes an innovative approach to the conduct of assessment in complex operations. The Niteworks partnership (a UK industry/MOD partnership) has conducted analysis involving system dynamic modelling, constructive modelling, virtual environments and controlled experimentation over the past 12 years and concluded that current analytical methods of assessment fall short when employed in the evaluation of the complex systems encountered in military operations.

There are many contexts where controlled experimentation and the application of the scientific approach is valid and appropriate, particularly in relation to hard physical systems and controllable inputs and outputs. However, the biggest challenge with operations assessment is that it involves humans performing activities as part of a complex sociotechnical system.

Niteworks' analysis found that controlled experimentation can be insufficiently open-minded, flexible and agile for complex system evaluation, due to a narrowing of focus to demonstrate cause and effect. This is perhaps a symptom of a broader systemic issue in that the analytical community resorts to shaping the problem to fit the techniques already in use rather than developing techniques better able to address the problem.

An alternative approach to assessment is proposed, which focuses on exploration and evaluation that is qualitative rather than quantitative and is more applicable to operations assessment in complex contexts. It will support evaluation of complex systems both at strategic and lower levels and explore the effect of innovative improvements, whilst providing a framework for the incorporation of other analytical techniques. The foundational thinking on this approach was developed by the Niteworks partnership in a Niteworks White Paper [1] intended to examine systemic issues in defence capability. This chapter describes the application of the approach to operations assessment.

Five ideas are expressed within the approach: explore with loose boundaries; live with uncertainty; be innovative; evaluate the wider consequences; then iterate from the baseline, taking account of political, military, economic and social factors as well as capability components and effect. This holistic approach to evaluating complex systems will provide a multi-faceted examination of the consequences of actions employed and inform those we choose to adopt in the future.

6.1 THE COMPLEX OPERATIONAL CONTEXT

> *"While state-versus-state conflict is still possible, contemporary operations are likely to be more complex and adversaries could be more difficult to identify. Increasingly we live in a world of wicked problems, which are so complex that they defy process-driven, management or scientific approaches. This does not mean that they are unsolvable, but the approach must be open-minded, agile, flexible and adaptable to work through the complexities."* [2]

[1] Niteworks is a UK MOD contract. This chapter was produced under UK MOD Crown Copyright 2016 and authorised for release.

The UK publication was written in 2011 but is no less relevant today. The operational environment becomes more challenging each day with adversaries continually adapting and evolving their strategies. If the analytical community does not adapt and continue to evolve their methods, it risks being irrelevant to problems future commanders will face. Recent events in the Middle East and Africa have highlighted the unpredictability of the operational environment, the agility of the adversary, and the challenges facing defence capability in such complex environments.

While many definitions for complex systems exist, for the purposes of this chapter the properties outlined in the Niteworks White Paper [1] have been adopted:

- Large numbers of and high degrees of variability in the elements, particularly if those elements include people.
- Large numbers of and high degrees of variability in relationships between elements, such that the system cannot easily be reduced to a number of distinct subsystems.
- Emergent properties/behaviours that are novel or unpredicted; i.e., where the whole is greater than the sum of its parts.
- Elements within the complex system change themselves and their relationships in response to their environment – known as adaptation.
- There is considerable uncertainty in relation to outcomes, such that cause and effect relationships within the system are only possible to establish in retrospect.
- Non-additive effects or non-linearities exist, such that the combined effect of two or more factors does not equal the sum of the two individual effects.
- There is sensitivity to initial conditions, where the same system can exhibit significantly different behaviours from near identical starting conditions.
- A sociotechnical system, with military actors performing tasks within an operating environment that includes political, military, economic and social aspects, is a complex system.

6.2 FAILURE OF CURRENT METHODS TO EVALUATE COMPLEX SYSTEMS

Given the highly interdependent nature of elements within a complex system and the emergent behaviour likely to be experienced, the application of methods that adopt a scientific approach, collecting quantitative data to confirm a hypothesis, will yield inappropriate evidence when employed to evaluate complex systems. In the past, while soft methods, such as influence diagrams or system dynamics models, have been used to understand context and focus, there has been a consistent view that the cause and effect relationship that exists between factors can only be determined through quantitative data capture. There is a growing school of thought, however, that suggests that only through retrospection can a cause and effect relationship be identified for a complex system underpinned by qualitative consideration of the context [3]. In addition, each outcome is completely context dependent so the generalisability of the results is often questionable.

6.3 CURRENT APPROACHES TO EVIDENCE GENERATION

Such quantitative, centralised, one-size-fits-all approaches to operations assessment have been used for many decades [4], [5]. The data collection and analysis has focused on quantitative data, e.g., body counts in the Vietnam War, and casualties in conflicts, rather than an approach that would look holistically at the complex system in context. This is not surprising as operations assessment has been driven towards a scientific approach when clearly in COIN contexts the top-down reductionist, aggregation of data approach is inappropriate [4]. In this method, the quantitative data gathered from each component is aggregated inappropriately to present a picture to the Divisional or Regional Commander.

In an attempt to evaluate an action and determine whether it has had the appropriate effect (assuming that the desired outcome had been determined in advance), the analysis of the complex system has typically been overly constrained to a cause and effect understanding of a small number of factors.

This reductionist mechanistic approach is described by Ackoff as machine age thinking [6]. His alternative, systems age thinking, describes a system within an open boundary, that is partially observable but only through holism[2] and synthesis can it be understood [7]. The next section offers some thoughts on how such an holistic approach can be used to *explore* the complex system.

6.4 PROPOSED HOLISTIC APPROACH TO COMPLEX OPERATIONS ASSESSMENT

An approach to operations assessment is proposed to support evaluation of the complex operating environment. The foundational thinking on this approach was developed by the Niteworks partnership in a Niteworks White Paper intended to examine systemic issues in defence capability [1].

Five ideas are expressed within the approach (see Figure 6-1): explore with loose boundaries; live with uncertainty; be innovative; evaluate the wider consequences; then iterate from the baseline, taking account of political, military, economic and social factors as well as capability components and effect. This holistic approach to evaluating complex systems, whilst not providing definitive answers, will provide a multi-faceted examination of the consequences of actions employed, or those we may choose to adopt.

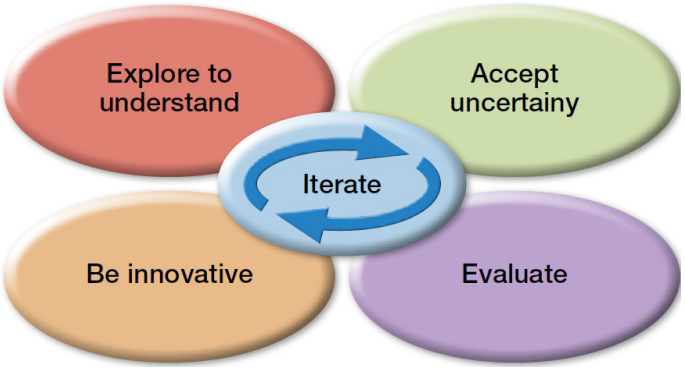


Figure 6-1: Summary of the Approach.

6.4.1 Explore to Understand

The first idea underpinning the approach is the adoption of an *exploratory* mindset [8] that encourages a breadth and curiosity, focusing on recognition of patterns and trends instead of a one-to-one mapping of cause and effect. This does not go as far as the multi-factor system dynamics diagram that described operations in Afghanistan [9], but attempts to capture the key factors of interest whilst establishing the loose boundaries and relationships.

To explore some of the boundaries and to identify some of the potentially relevant factors in the operational environment, we propose the use of the seminar wargame. This technique encourages collaboration and engagement involving participants from the military as well as local representation covering the political, social and economic aspects. This approach will ensure that the complex system is considered from a wide range of perspectives. In Ref. [10], O'Neil argued that when we think about another community or culture's

[2] Holism suggests that there are different levels of explanation for a given problem and there are likely to be emergent properties at each level. Reductionism in contrast tries to break down things into their constituent parts.

problem, we think about it from our perspective. We apply our own experience and pre-defined schema (an ethnocentric perspective) without attempting to appreciate the situation from a completely different perspective (cultural relativity). A collaborative multi-user wargame representing different cultural aspects, ideally with representation from the region of interest or a player with experience of the culture, should facilitate better understanding and exploration. It should support both this stage and subsequent steps to deliver better assessment of the operation and potential future improvements in a culturally relevant context. There are four steps to develop our understanding and define a baseline upon which any changes can be assessed. These are summarized in Figure 6-2.

Figure 6-2: Steps of Understanding.

6.4.1.1 Step 1: Understand the Strategic Context and Intent

Appreciation of the strategic context for any conflict and the strategic intent is a key leg of the operations assessment stool. What is clear from analysis of previous conflicts, particularly COIN campaigns, is that strategic intent has not always been clearly defined, for example, the effort in Afghanistan suffered from the absence of a clear exit strategy in terms of the achievement of a strategic outcome or political endstate. Equally the position and allegiances of neighbouring countries and associated regions were not well understood at the point we began to get involved in Afghanistan. Our ability to leave a campaign or operation feeling that we had achieved our objectives and had satisfied the strategic intent is highly dependent on an appreciation of these broader strategic conditions, for example the relationship between Afghanistan and Pakistan, and the relationship between Muslim factions and the Taliban.

Furthermore, there is the home context – the attitudes of the home nation population, the political perspective, and the social context. The US participation in Vietnam resulted in thousands of veterans returning to the US, disillusioned and suffering from combat stress (700,000 Vietnam veterans suffered psychological after-effects), whilst anti-war opinion changed the political landscape. For example, according to one source, news of the killings at My Lai lost the US the claim to moral superiority and its status as the world's defender of freedom and rights [11]. The implications of any involvement both on the local community and the home population need to be appreciated.

The strategic context shapes the strategic intent and the intent needs to be defined and articulated upfront to help guide the conduct of the operation. Without the clarity this provides, it is very difficult to conduct any form of operations assessment. Strategic intent is one of the key assessment criteria within the complex systems approach proposed in this chapter.

6.4.1.2 Step 2: Understand the Operational Context and Objectives

Having attempted to understand the strategic context and having articulated the strategic intent of the specific operation, the consideration of the political, military, economic, social, information, infrastructure, physical environment and time pressure factors (based on the PMESII-PT framework[3]) captures the operational context. This supports an appreciation of the specific characterisation of factors for a particular operational environment. Alongside the strategic indicators associated with country allegiances and home pressures, the operational indicators help planners to appreciate the tribal rivalries, religious tensions, and the stability of the society into which military elements have been introduced.

An ordinal rating for each of the factors of the PMESII-PT framework, communicated by a Red/Amber/Yellow/Green/Grey (RAYGG) coding, is a way to characterise the operational environment. While Connable has criticised the use of colour coding in operations assessment and highlights the risk that it can carry more weight than it should, especially when commands attempt to aggregate variables, the approach proposed here moves away from aggregation and includes presentation to leadership of the colour coding of each of the sub-variables under each PMESII-PT factor, as well as the assessment of the colour rating of the factor itself [4]. Taking an example using the Political factor, the sub-variables include (with each colour-coded separately): attitude to UK (G); centres of political power (A); government type (A); government effectiveness (A); influential political groups (A). The Political factor overall, taking account of the sub-variable ratings, would be indicated as Amber. The sub-variables and the factors are underpinned by a justification table to ensure transparency by defining the rating levels. This helps preserve the context of a particular situation and the underlying rationale for any assessment. Where colours are aggregated (for example, if numerical values were attributed to the individual colours and then the mean taken) the underlying rationale and context would be lost; this is undesirable. In our method the RAYGG ratings will provide a status dashboard, capturing whether individual PMESII-PT factors and sub-variables, for example the Political factor: influential political groups, has a positive or negative impact on the situation: red and green indicating the negative and positive extremes. So if the rating of the Political factor was indicating red this would highlight a very negative strategic context and the underlying sub-variable ratings would provide some indication of why this might be the case. The approach is not definitive, but it is a communication device that summarizes the staff's complex assessment for decision makers.

Having an adequate understanding of the operational context, and with a clear articulation of our understanding of the strategic context and intent, the operational objectives can be captured. Figure 6-3 provides an example of the dashboard that could be developed in order to communicate the assessment.

6.4.1.3 Step 3: Understand Whether the Military Capability Currently Supports the Operational Objective

Whether there is some form of defence engagement already in the operational theatre or a new team coming in to support the operational objective, some assessment of the capability of an intervention force, including the preparation, command and control, intelligence and protection, is required to determine the extent to which components of military capability are contributing to supporting the operational objective. In a manner similar to that described above in Step 2, each of the components of capability, and the tasks that underpin that capability, are reviewed, and qualitatively assessed using an ordinal scale (communicated by colour codes) based on the extent that the tasks have been conducted satisfactorily. For example, was the team adequately prepared? Did they have enough or appropriate training? Was the intelligence available of sufficient quality and was it provided to the appropriate people?

[3] PMESII-PT is used by US Department of Defence for military planning within a counter-insurgency context (see NATO AJP-2(A)) and promoted in the development of the Decisive Action Training Environment (DATE) for collective training both in the US and UK.

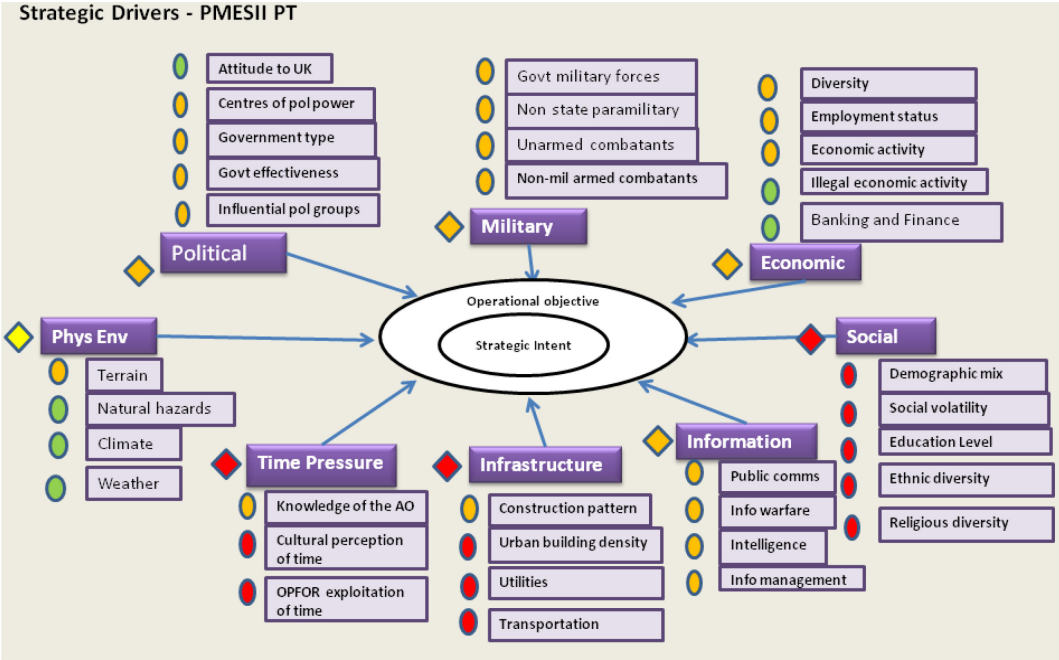


Figure 6-3: PMESII-PT Factors and Strategic Intent/Operational Objective to Help Define the Strategic Context (Developed from Ref. [12]).

Consider Figure 6-4, as an illustration of this step. The green colouring of the Direct and Collect boxes indicate that the tasks have been performed well. However, the processing of the collected intelligence, including the fusion of the other forms of intelligence as well as the appreciation of the cultural context and dissemination of the information to the required people is assessed as lacking, indicated by the amber colouring of the Process and Disseminate boxes on the dashboard. Taking that into account and the assessed failures within the command loop (indicated by the red boxes) in terms of understanding, planning, integration and control as well as the lack of preparation of personnel, there are likely to be serious concerns over the extent to which the operational objective can be achieved with existing capability.

6.4.1.4 Step 4: Understand the Strategic Consequences of Military Capability

The previous steps (steps 1-3) should have developed a shared understanding of the strategic context and an assessment of the effectiveness of existing capabilities. However, the attempt to make sense of this information would not be complete without an appreciation of the strategic outcomes that are likely to arise from the application of the current capability. This step considers the contribution of the military capability and examines the wider or broader implications.

As an example, the strategic objective might have been to maintain stability in the region and prevent escalation. However, the presence of military capability in-country including tanks and armoured vehicles, while maintaining a degree of stability, might affect the indigenous population's permissive attitude to a foreign military force. The population may feel occupied rather than supported, leading to hostility and negative strategic consequences. The appreciation that the force does not match the desired strategic outcome should lead to the conclusion that something needs to change – either the proposed military capability or the strategic intent. Figure 6-4 communicates an example of the capability, operational aim and strategic intent assessment where the limitations in capability, particularly in the Command, Inform, Operate and Prepare areas, is having a negative effect on achieving the operational aim. In order to achieve both the operational aim and better strategic outcomes, the current operation may need to be modified to be more effective tactically, operationally and strategically. This will be covered in a later section.

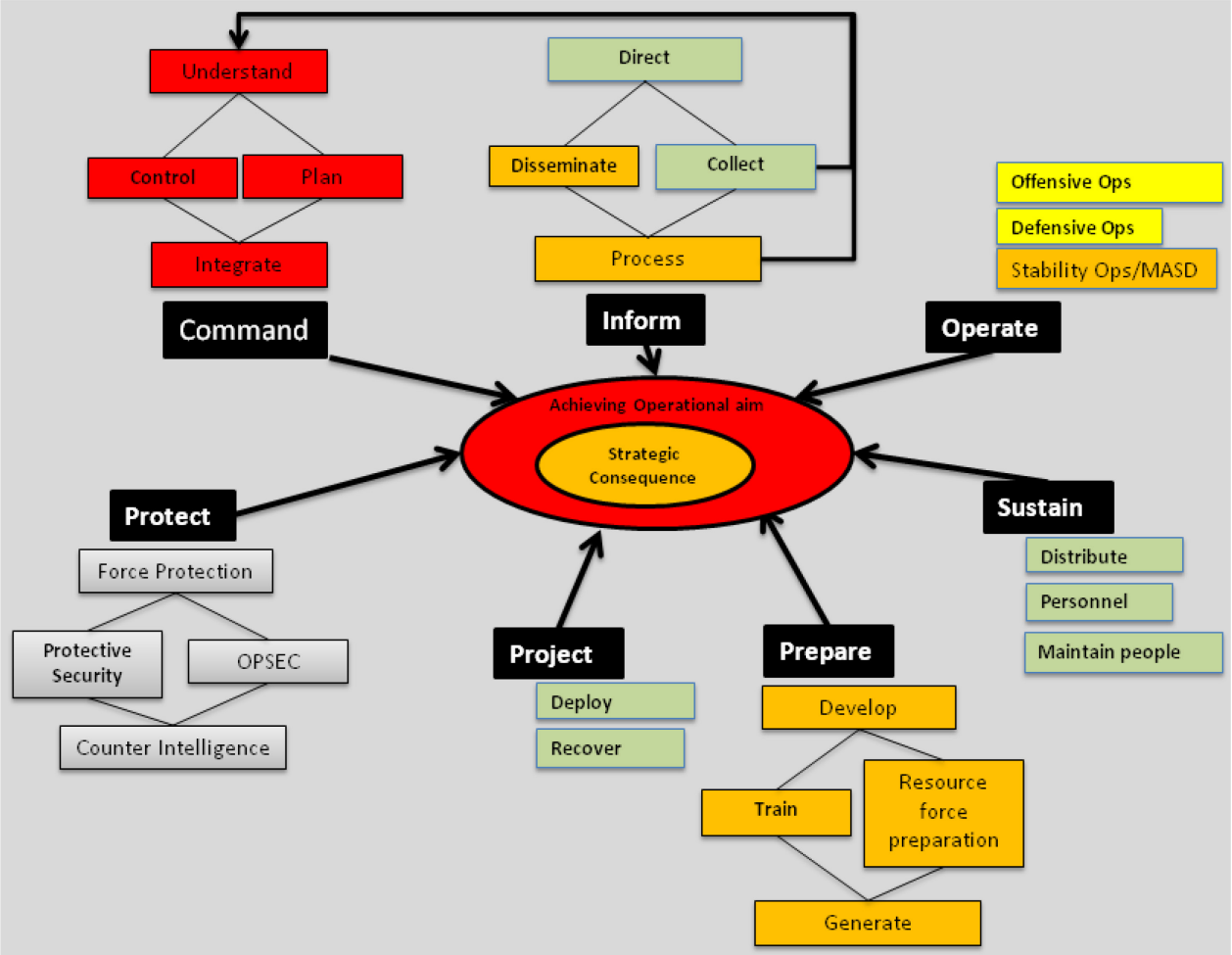


Figure 6-4: Assessment of Capability, Operational Aim, and Strategic Consequences [12].

6.5 LIVE WITH UNCERTAINTY

In establishing the baseline capability assessment, a number of assumptions have been made about each of the various factors that might be present, and the way in which these factors interact with each other. In Figure 6-4, there is only one line that connects the components of capability; the line between Inform: Process and Command: Understand. This is a loose connection but one that seems reasonable. However generally, the idea is to have a loosely coupled system and a loosely coupled representation so as not to hardwire any of the relationships. The strength of the relationships between capability components is not known and indeed some of the factors may not be relevant, but it is possible, even with these uncertainties, to look at scenarios where factors may interact with each other, for example increased armour (protection) has an impact on sustainment, and investigate the consequences of such interactions. Consideration of the potential ways in which the enemy's tactics might change will require some tolerance for uncertainty and therefore some resiliency and adaptability in our own forces.

6.6 BE INNOVATIVE

Having established a baseline, and perhaps having recognised some lack of effectiveness in achieving tasks, operational aims, and strategic objectives, there is an opportunity to alter operations to be more

effective. The improved approach would not necessarily adopt an equipment focus, it could be based on alternative ideas, such as improvements in social media and influence, or better messaging, rather than introducing troops.

It is important to recognise that innovation is key to realising successful outcomes. The innovative element is that which drives out particular nuances of an operational context. The *realist* community [13], [14] emphasise that consideration of an "intervention" or improvement in capability should focus on what works, for whom, in what circumstance. Each situation needs to be looked at in its own merit and the appropriate innovative alternative identified.

A continually-evolving appreciation of the strategic context, drawing from the PMESII-PT framework analysis, may help in modifying ongoing operations. Consideration of innovative modifications provides opportunity to think beyond current constraints to address gaps highlighted during re-analysis of the baseline capability and the assessment of operational objectives.

Adopting a pan-defence lines of development[4] similar to the US model approach to the identification of alternative military capability solutions, such as changes to ways of working, change of personnel, improved training and better information is proposed. Once these options for change have been identified, they should be expressed in relatively simple actions so that the potential impact of any one change to the complex system can be explored and understood. This is where "what if" questions, particularly in relation to changing capability, could be explored. For example, what if the dissemination was better – would it make any difference? If preparation was better, e.g., cultural training, and there was better dissemination to the appropriate level, would that make the military force more effective?

6.7 EVALUATE THE WIDER CONSEQUENCES

The next stage is to evaluate the wider consequences of the application of innovative alternatives identified to the complex problem. Both the innovation and evaluation stages use a "fast fail" approach looking to promote those things that have promise and discard those that do not. It is now possible to update the dashboard contained in Figure 6-4 with a colour-coded communication of the assessment of capabilities as shown in Figure 6-5.

The wider consequences of action could be reflected in Figure 6-5 and reflected alongside the characteristics of the operational environment (previously shown in Figure 6-3). The tolerance of the military capability to the evolving operational conditions should be considered as part of this evaluation.

6.8 ENRICHING THE EVALUATION

The complex system approach proposed is focused on analysis of the problem whilst acknowledging the complexity of the system. It does not discount other methods, such as controlled experimentation and simulation, however, the focus must be on maintaining a holistic approach to evaluating the complex system, rather than decomposing the problem, and adopting a reductionist approach. Investigation of the interaction between two factors within the complex system might be more appropriate using simulation or military judgement panels. The information gained from such investigation will enhance our understanding of the complex system thus enriching the assessment.

[4] Training, Equipment, Personnel, Information, Doctrine, Organisation, Infrastructure and Logistics. See Network Centric Operations (NCO) Case Study. The British Approach to Low-Intensity Operations: Part II for a useful explanation of the Lines of Development. Department of Defense, Office of Force Transformation, Washington, DC, 20303.

6.9 ITERATE FROM THE BASELINE

The final stage is to iterate, to go around the cycle again, building on the knowledge that has been gained through the development of the baseline, coupling understanding, uncertainty, and innovation. This stage enables steps to be taken towards achieving the operational objective and the strategic intent. This is highly dependent on the extent to which any alternative actions address the shortfalls identified from analysis of the baseline, but the evaluation activity should identify the extent to which particular alternatives or sequences of alternatives will address the gaps identified in order to be more effective.

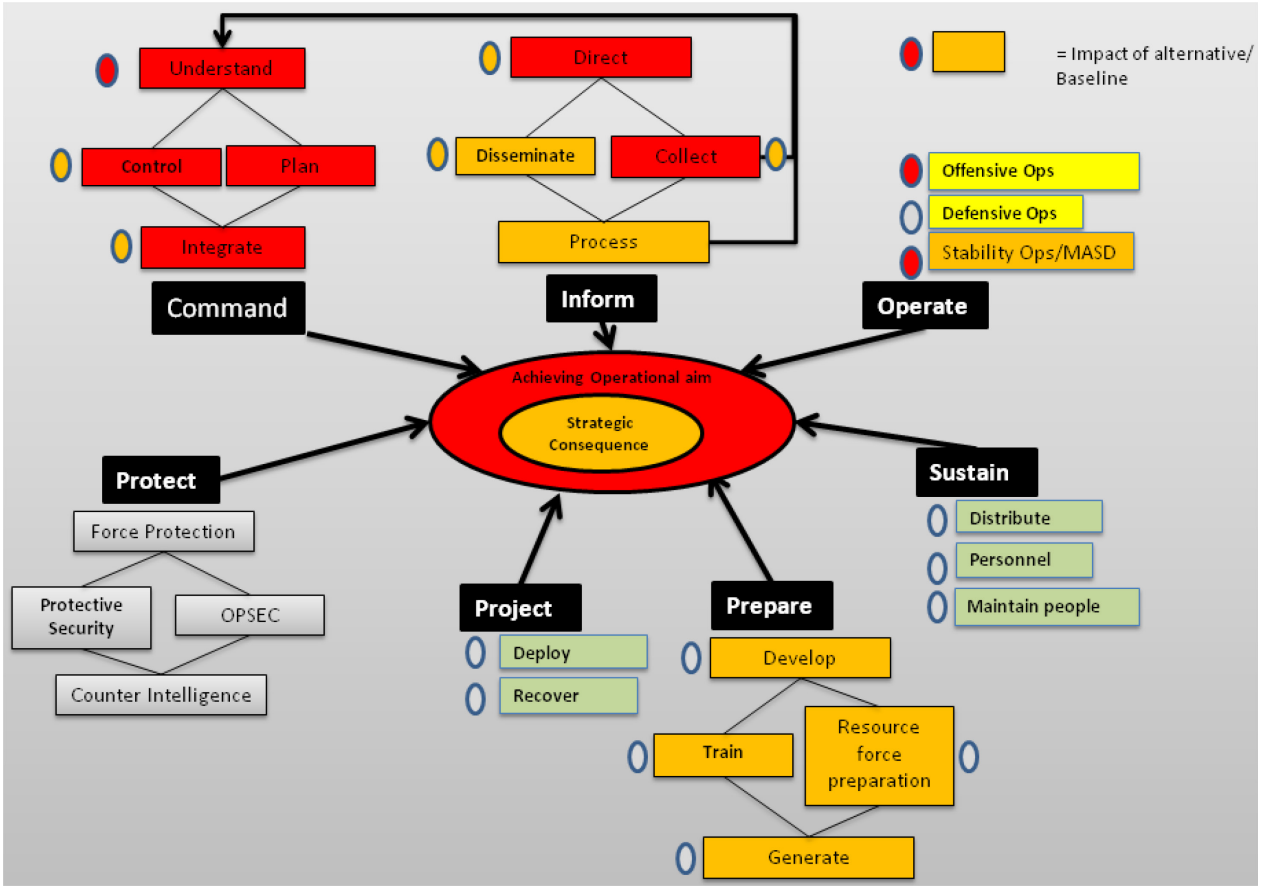


Figure 6-5: Example to Show How an Alternative Option Can Be Considered.

6.10 BENEFIT OF THE HOLISTIC COMPLEX SYSTEMS ANALYSIS APPROACH APPLIED TO OPERATIONS ASSESSMENT

The holistic complex systems analysis approach supports the exploration of an operational and strategic context, an understanding of the current capability and the consequences of employing that capability in context. It supports an understanding of how any change, e.g., through improved training, communications, protection, etc., might affect the complex system in a positive or negative way. The approach should encourage consideration of each of the factors, what they are sensitive to, and if affected, what impact each would have on any related factors. Such analysis is not as complex as an influence diagram or system dynamics model, but would benefit from application in a wargame setting or military judgement panel.

The ideas behind the approach encourage the user to think creatively about viable alternatives, and the overall approach permits examination of the complex system from a range of different perspectives: strategic; operational, and tactical, as well alternative perspectives from "red" teams or NGO participants, so that in future operations, assessment, and ultimately, operations are more effective.

6.11 AUTHOR'S BIOGRAPHY

Chris Jordan was the Chief Analyst for the Niteworks partnership. He has a degree in psychology from Queen's University Belfast and a Master of Philosophy in cognitive psychology from Reading University. He has broad experience of analysis, experimentation, test and evaluation, primarily in the Defence sector. From an early career in the Defence Research Agency, he became involved in the disciplines of Human Factors, Human Performance Modelling and Systems Engineering in the context of military command and control systems and subsequently wider military capability. Since then he has worked on a range of analysis and experimentation projects to support military capability development through decision support.

6.12 REFERENCES

[1] Jordan, C. and Wilkinson, M. 2015. Holistic Complex System Intervention Evaluation – Understanding the nature of defence capability. Niteworks NW/CR/0495/209.

[2] DCDC. 2011. Understanding and Intelligence Support to Operations Edition 3. UK Joint Doctrine Publication JDP 2-00.

[3] Snowden, D.J. and Boone, M.E. 2007. A leader's framework for decision making. Harvard Business Review. November, 2007.

[4] Connable, B. 2012. Embracing the Fog of War: Assessment and Metrics in Counter-Insurgency. RAND Corporation.

[5] Downes-Martin, S. 2011. Operations Assessment in Afghanistan is broken. Naval War College Review, 64(4), Autumn.

[6] Ackoff, R. 1981. Creating the corporate future: Plan or be planned for. New York: John Wiley and Sons.

[7] Hitchins, D. 2009. What are the General Principles Applicable to Systems? INCOSE Insight, 12(4):59-63, December.

[8] de Coning, C. 2013. Chapter 10 of Innovation in Operations Research: Recent developments in measuring results in conflict environments. Edited by A. Williams, J. Bexfield, F. F. Farina and J. de Nijs. NATO ACT.

[9] PA Consulting. 2009. Dynamic Planning for COIN in Afghanistan. PA Consulting Group.

[10] O'Neil, D. No date. Characteristics of Culture. California: Palomar College.

[11] Chambers II, J.W. 1999. The Oxford Companion to American Military History. Ed. John Whiteclay Chambers II. New York: Oxford UP.

[12] Wilkinson, M., Jordan, C. and Currie, F. 2016. Developing Strategic Systemic solutions to complex problems in the UK Defence Enterprise (Draft). 26th Annual INCOSE International Symposium (IS 2016). Edinburgh, Scotland.

[13] Bhaskar, R. 1978. A realist theory of science, 2nd Edition. Hassocks: Harvester Press.

[14] Pawson, R. 2013. The science of evaluation: A realist evaluation. Sage publications.

Chapter 7 – ASSESSING NATO'S COMMUNICATION EFFECTIVENESS WITH COMPUTER-BASED CONTENT ANALYSIS: AN EXAMPLE FROM AFGHANISTAN

Valentin Poponete
SHAPE Strategic Planning
BELGIUM

ABSTRACT

While the importance of communication and influence activities in complex operations has long been recognized, there is usually very limited assessment of the effectiveness of these activities. The few studies that have addressed this issue in the case of NATO and U.S. operations in Afghanistan concluded that, while media monitoring has been a constant activity, there was very limited analysis of effectiveness due to an absence of proper measures of effectiveness, assessment procedures, and methodologies.

These assessments were short-term oriented and based on measures of performance (e.g., counting of the number of communication products produced) and inadequate measures of effectiveness (e.g., counting of media article publishing a certain story, subjective weighing of published articles into positive, negative or neutral).

This paper presents a longer-term assessment and recommends a method that could represent a step forward. This method is a systematic and objective computer-based analysis of text content and could open a new world for NATO communication practitioners and support their understanding of how their messages impact any audiences, from the people in the remote villages of Afghanistan, to the internet-based communities throughout the world.

7.1 INTRODUCTION

The idea that warfare is changing is nothing new. But it is not only the nature of conflict itself that changes; it is also the relation of conflict with the media, which has become an integral element of wars. It seems that "contemporary wars and modern news media attract each other" [1]. The former NATO Deputy Supreme Allied Commander Europe General Sir Rupert Smith eloquently describes the current operating environment as being broader than the fields of the conflict zone: "We now come to the other manner in which we fight and operate amongst the people in a wider sense: through the media. Television and the Internet, in particular, have brought conflict into the homes of the world – the homes of both leaders and electorates" [2].

While this new dimension of the operations environment has different names such as the information environment, the cognitive dimension or the human environment, the reality remains: conflicts are also fought through acts of communication that could represent either effect multipliers or create undesirable effects. And, while NATO has improved its communication policies and practices over the last decade, there is one thing that requires more attention: assessing the effectiveness of their communication acts.

The aim of this paper is to point to the limited development that was achieved by NATO with regards to the way its communication is assessed, and to recommend a method that could represent a step forward. Through a systematic and objective analysis of texts, computer-based content analysis could open a new world for NATO communication practitioners and support their understanding of how their messages impact any audiences from the people in the remote villages of Afghanistan, to the internet-based communities throughout the world.

7.2 THE EVOLUTION OF NATO'S STRATCOM

In the context of the new operating environment, StratCom has become a catchphrase among anyone involved in conflict management, but not everyone understands its appropriate meaning. Strategic communication is considered a critical aspect in contemporary armed conflicts [3].

Although only institutionalised during the last century in parallel with the emergence and expansion of mass communication channels, the role of communication in influencing military conflicts has been considered for centuries. For example, Sun Tzu, stated that "supreme excellence consists in breaking the enemy's resistance without fighting" [4], and von Clausewitz recognized that one of the key strategic principles of war is gaining the support of public opinion [5]. While different organizations have developed their own definitions of strategic communication, in straightforward and simple terms, this represents "the purposeful use of communication by an organization to fulfil its mission" [6].

Mainly under the influence of the circumstances in different theatres of operations, especially Afghanistan [7], NATO implemented a series of changes to improve its communication and adapt it to new security challenges. It approved a policy, a military concept, and a capability implementation plan for StratCom. According to policy approved in 2009, NATO StratCom represents "the coordinated and appropriate use of NATO communications activities and capabilities – Public Diplomacy, Public Affairs, Military Public Affairs, Information Operations and Psychological Operations, as appropriate – in support of Alliance policies, operations and activities, and in order to advance NATO's aims."

While the importance of communication and influence activities in complex operations has long been recognized, there is usually very limited assessment of the effectiveness of these activities. Despite all the changes NATO adopted to improve its communication activities, NATO has not devoted sufficient attention to devising methods to assess the effectiveness of its messaging. As a case in point, for the entire duration of the NATO-led International Security Assistance Force mission in Afghanistan, despite increasing requirements, "assessments and analysis was weak – and by several accounts, a situation that endures at NATO HQs" [7].

In addition, NATO StratCom practitioners constantly discuss assessment of their efforts with the widespread agreement that this assessment is difficult [8]. As a consequence, there seems to be a tendency to focus on measuring performance (e.g., the number of hours of broadcast, the number of press releases issued, etc.) that provides no indication about the effectiveness of these communication activities on target audiences.

In the following pages, I will introduce a methodology known as 'content analysis' which we can use to assess the effectiveness of NATO's StratCom. I argue that although content analysis cannot ascertain the behavioural changes in the different target audiences, it can provide an assessment of the effectiveness with which NATO's StratCom influences the communication of these audiences. While analysing the communication of large audiences is very challenging, by examining the communication products of individuals or organizations who have the potential to influence larger audiences, such as village leaders in Afghanistan, internet community opinion leaders, and news media organizations, we can better assess the effectiveness of our communication.

Besides using the existing literature to support my arguments, I will also give an example of a computer-based content analysis of about 10,000 media articles related to NATO-led International Security Assistance Force (ISAF) mission in Afghanistan and a few hundred ISAF press releases and Taliban propaganda statements. Joint Forces Command (JFC) Brunssum Public Affairs (PA) performed this analysis in 2009 in an effort to improve the effectiveness of its communication efforts related to the ISAF mission in Afghanistan. While PA is just one of the StratCom capabilities, this case study exemplifies the way computer-based content analysis could be used in assessing effectiveness of any communication and influence activity.

7.3 ASSESSING STRATEGIC COMMUNICATION WITH NETWORK TEXT ANALYSIS

In Ref. [9], Paul compared about 20 conceptions of strategic communication and discovered that the related terminology is vague and often contradictory. He concluded that evaluation is key for effective strategic communication, and the fact that measuring influence is challenging does not suggest that it should not be done.

During its operations, especially in Afghanistan, NATO employed several methods to try to measure the effectiveness of its strategic communication campaigns. It used opinion polls, atmospherics, and a kind of content analysis of media reports that qualified, mainly based on the subjective judgement of the analyst, whether various media reports were supportive or not of NATO-led forces. Except the opinion surveys, which could give some indications about the effectiveness of the ISAF StratCom, the other assessments were rather short-term oriented, and changed virtually with every headquarters' personnel rotation.

In Ref. [10] p. 14, Holsti defines content analysis as "any technique for making inferences by objectively and systematically identifying specified characteristics of messages." By systematically evaluating texts, qualitative data can be transformed into quantitative data to which statistical analysis methods can be applied.

Over the last four decades, computer-based content analysis methodologies have been developed [11], [12] and are successfully being used to analyse large volumes of text in order to extract essential meanings. The advantage of such analytical aids lies in the assurances they can provide that text explorations are systematic, effectively countering the natural tendency of humans to read and recall selectively. One type of such computer-based content analysis methods is called Network Text Analysis (NTA). It enables the representation of written texts as concept maps and the determination of the meanings denoted within these texts [13].

NTA has been used for analysis in various contexts that represent relevant case studies and could provide valuable lessons for NATO's StratCom. For example, NTA was applied for such analysis as: the media coverage of the terrorist attacks of 11 September 2001 [14], the identification of deceptive news [15], the study of discourses of resilience across American and British online news and blog coverage of the 2005 London subway bombings [16], and research on Twitter communication during the 2011 Egyptian uprisings that resulted in the resignation of President Hosni Mubarak [17], [18].

In Ref. [19], Diesner identified seventeen NTA methods that, while differing along features such as automation, abstraction and generalization, consider that text content could be displayed as networks of words and their interrelations, and that the position of the words within the network provides an understanding of the text meaning and significant topics.

Centring Resonance Analysis (CRA) is an NTA method proposed by Corman and colleagues in a 2002 article. CRA is based on the linguistic ideas of centring theory, where "centres", mainly a noun or noun phrase, are the words that contribute the most to the meaning of the text and to which other words are connected to construct coherent messages.

CRA "identifies discursively important words and represents these as a network, then uses structural properties of the network to index word importance" [20] p. 157. In this way, it developed a metric called "influence" which measures the degree to which a centre generates coherence in a text [21]. A word influence value above 0.01, is considered significant, while an influence value above 0.05 is considered very significant [22].

Another metric called "focus" measures the coherence of a text, the degree to which the concepts are organized in the text [23]. The value of this metric can be between zero and one, with higher scores indicating more coherent, focused texts. Longer texts tend to be less focused as their authors need to be more skilled to organize more concepts. Besides, lower focus texts tend to be more casual or informal than higher focus texts [22].

In addition, CRA is the only approach among NTA methods that developed a metric called "resonance" that measures the similarity between two texts in terms of their respective centres [20]. The more texts that use the same words in influential positions, the higher their resonance is and the more similar they are. Resonance values can be between zero and one and, as resonance is a form of correlation, it can be interpreted the same way, with higher values representing a stronger resonance [22].

Figure 7-1 shows an example of the CRA analysis output of an ISAF press release [24] from 01 April 2008. It shows the different CRA statistics, the text network as well as the influence values of the most influential words and pairs of words in the text. The most influential nodes in the network are the words in coloured boxes that tend to be located near the top of the graph, while the less influential words are situated in black in the lower part of the graph. The ties between the nodes indicate that those words are connected with each other in meaningful ways in the text, and the darkness of the lines indicate the strength of ties between the various words.

7.4 ASSESSING ISAF'S PUBLIC AFFAIRS WITH CRA METHODOLOGY

Between 2008 and 2010, I was the chief of the ISAF Cell of the Public Affairs Office at Joint Forces Command Brunssum, the operational out-of-theatre headquarters in charge of the NATO-led mission in Afghanistan. In 2009, in an attempt to improve the planning of Public Affairs activities based on assessments, I decided to employ CRA to understand the potential of this method and make recommendations that, if implemented, could improve the effectiveness of ISAF's Public Affairs activities. At that time, there were no measurable definitions of ISAF Public Affairs objectives and effects and thus, the analysis started with two general questions:

- Which were the themes and messages conveyed in ISAF's Public Affairs activities and how were they reflected in the Afghan and International media?
- Which were the themes and messages conveyed in the enemy's (Taliban) information products and how were they reflected in the Afghan and International media?

These questions assume the themes and messages transmitted by both ISAF and the Taliban were a result of deliberate decisions based on their overall communication strategies.

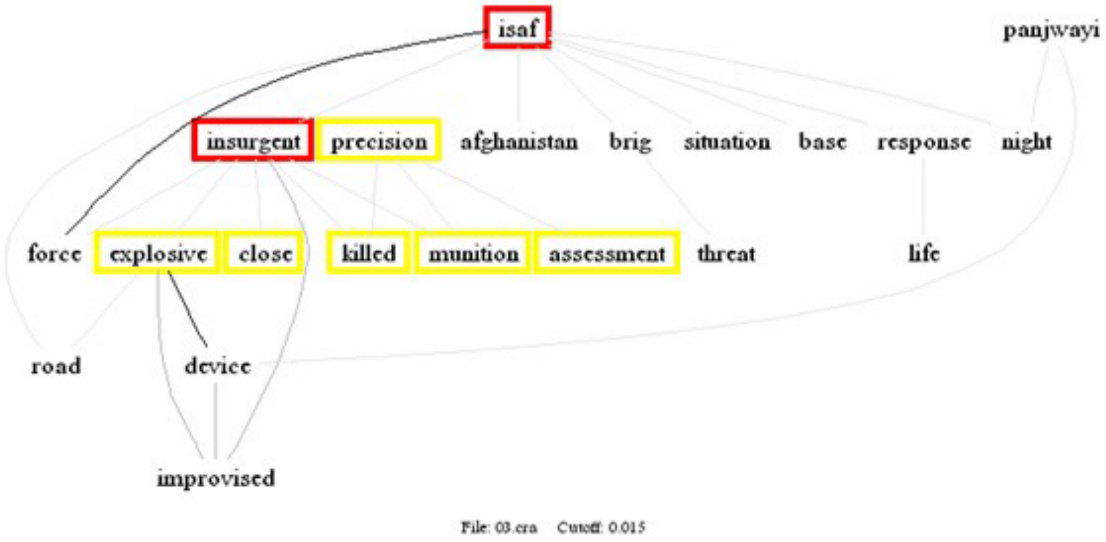
For this assessment, I analysed the following documents:

- A systematic random sample of the JFC Brunssum Daily Press Summaries from February 2007 until December 2008 that comprised about 5,500 international media articles addressing the ISAF mission.
- A systematic random sample of the ISAF PAO Daily Media Review from January to December 2008 that comprised about 4,500 Afghan media articles addressing the ISAF mission.
- A systematic random sample of about 350 ISAF press releases from January to December 2008.
- A systematic random sample of the ISAF PAO Daily Media Review from June to December 2008 that comprised about 200 Taliban propaganda messages published on their website and translated by ISAF in English.

CRA Network Statistics

Number of nodes: 56
Density: 0.059
Focus: 0.463

CRA Map



Influence Analysis

Words	Pairs
isaf 0.48675	isaf | insurgent 0.194
insurgent 0.3987	isaf | force 0.047
precision 0.09865	insurgent | munition 0.031
munition 0.0771	insurgent | killed 0.031
killed 0.0771	isaf | road 0.024
assessment 0.06195	isaf | brig 0.022
explosive 0.0551	insurgent | explosive 0.022

Figure 7-1: CRA Network Sample.

The reason for differences in the timelines for which the various documents were randomly sampled was their availability. For the JFC Brunssum Daily Press Summaries, I chose the 2007–2008 period because I had the 2007 Press Summaries already processed for a previous trial analysis. For the Taliban propaganda, June–December 2008 was the only period for which translations were available at the time of the study.

To do the analysis, I employed a software called Crawdad Text Analysis Software that was based on the CRA methodology. It required minimum processing of text before analysis, mainly the transformation of documents from .doc to .txt format. In addition, I used Statistical Package for Social Sciences (SPSS) to conduct further statistical analysis using the numerical data provided by Crawdad to aggregate words and identify communication themes for both the international and Afghan media.

7.4.1 Analysis

Using Crawdad, I identified the most influential concepts in the international media reporting. The upper part of the CRA network map was populated with words such as Taliban, United States, Pakistan, government, force, country, and NATO. Secondly, by performing factor analysis with SPSS on the influence value time series of the influential words in the CRA network, I identified the main themes of the international media and their trends over time. Each theme comprises several words highly correlated with one another over time. I also plotted influence values of the thematic time series to identify change points and reporting periods. This analysis is presented in Figure 7-2.

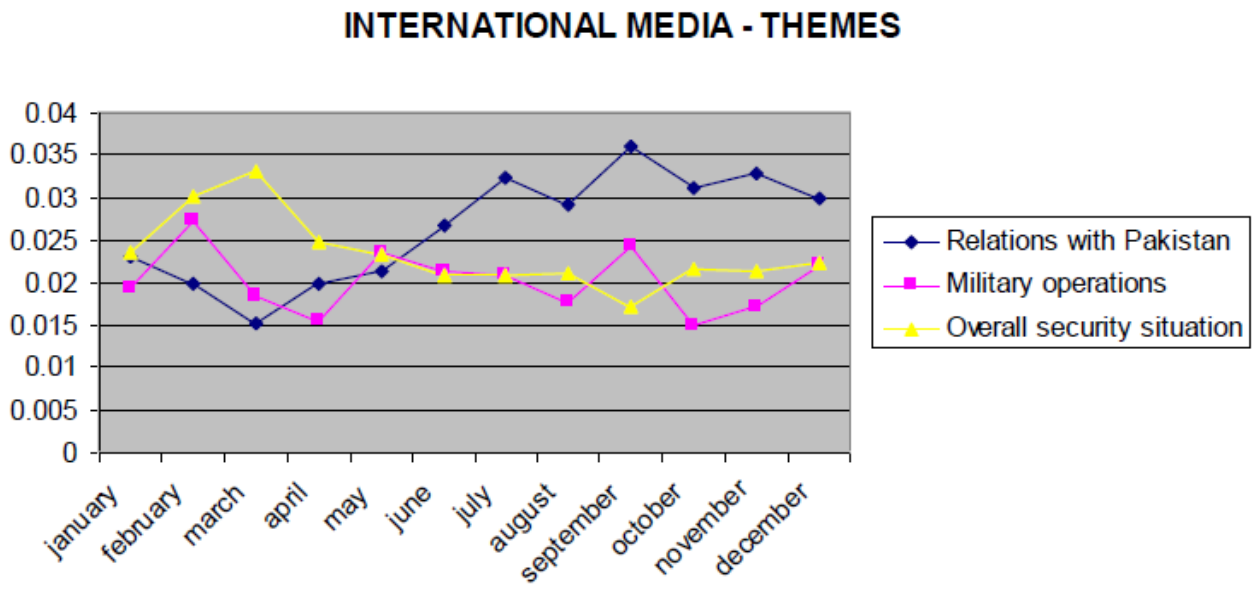


Figure 7-2: Themes in International Media (2008).

The analysis of most influential words in the Afghan media revealed that their reporting focused on concepts such as province, government, Taliban, people, district, country, force and security. The main themes of the Afghan media reporting were security, government institutions and relations with Pakistan (see Figure 7-3).

By simply comparing the themes and topics of international and Afghan media reporting, it is possible to identify that Pakistan and the security situation were similar themes of interest. At the same time, the Afghan media focused more on the country's internal affairs and the effectiveness of national and local government institutions while the international media was more interested in the external support and the conduct of military operations by the international forces. This second part suggested the need for the ISAF Public Affairs campaigns to tailor their messages to the different audiences. While theoretically the principle of tailoring messages to audiences is nothing new for the Public Affairs personnel, this has not necessarily been applied in practice, as further analysis will confirm.

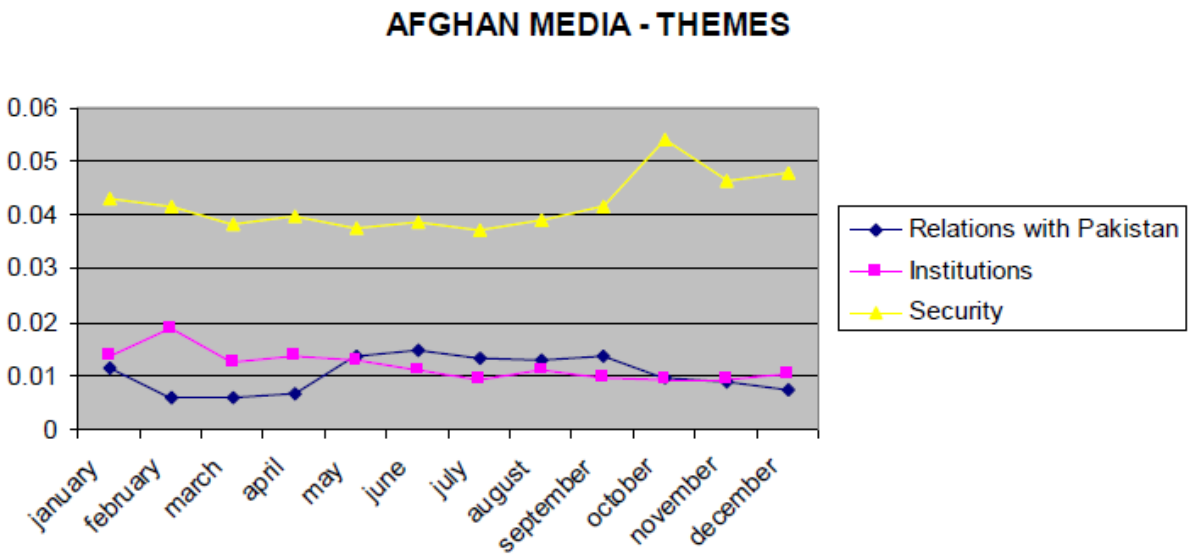


Figure 7-3: Themes in Afghan Media (2008).

The influence analysis of the ISAF press releases identified that they focused on topics that aggregated in three main themes: civilian casualties, Provincial Reconstruction Teams (PRTs), and military operations. This shows that there is a discrepancy between the ISAF messages in the press releases and the topics of interest of the media. It is especially the case of the ISAF press releases related to PRTs' activities that provided "good news stories" but were of very limited interest to the media. In addition, the analysis shows that communication related to the other two themes, civilian casualties and military operations, were reactive to events that happened in Afghanistan as all press releases related to these themes were published post factum. The analysis suggests that there was very limited initiative by the ISAF Public Affairs Office to engage with the media in a proactive manner on these topics to try to promote its own narrative.

To partially account for the time lag between the publishing of the press release and the media reporting, resonance analysis was conducted on daily, weekly and monthly press releases and media reporting with very small differences. When measuring the resonance of ISAF press release with the international and Afghan media respectively, the analysis showed that there was a very low similarity between the messages and themes communicated by ISAF and the topics and themes around which the media focused their articles. Thus, the overall resonance of the ISAF press releases with the international media reporting was 0.035 while the overall resonance with the Afghan media reporting was 0.05. These results suggest that most of the media reporting about ISAF was centred on topics that were influenced in a very limited manner by the ISAF communication campaigns.

During periods of lower intensity fighting, especially during the winter, the resonance of the ISAF press releases with the media reporting had the tendency to be even lower. This suggests that during these periods, other sources of information appear to have been even more relevant for the media. Moreover, ISAF seems to have missed the opportunities offered by these low-intensity fighting periods to engage the media and try to shape the message away from the short-term themes of casualties and operations.

The analysis of the Taliban statements published on their website suggests that their propaganda focused on military operations, political issues, and social/moral justice, themes which appear to be much more in line with the Afghan media themes of interest. The high visibility of civilian casualty incidents caused by the international forces gave Taliban propaganda additional opportunities to stress the social/moral justice theme. In addition, the resonance between the Taliban propaganda messages and both the Afghan and international media reporting was 0.05, which appears to be better or at least equal with ISAF's ability to influence the media.

In essence, if we assume that the messages in the ISAF press releases were a reflection of the overall NATO communication strategy for Afghanistan, we could easily draw some conclusions about the effectiveness of NATO's StratCom in Afghanistan: very little from what ISAF communicated was taken forward by the media, which means NATO's StratCom had limited effectiveness.

Further comparative analysis of the coherence of ISAF Press Releases and Taliban propaganda statements revealed the fact that they had similar values of the focus metric, both close to 0.5, which signifies that texts were relatively well organized, which was expected due to the fact that these were short in length. Moreover, I compared the coherence of the unclassified version of the ISAF Commander Counterinsurgency Guidance and the English translation of the Taliban 2009 Rules and Regulations Booklet. Both documents were issued in the same time period and were available on the Internet. The Taliban booklet focus score of 0.309 is lower than that of the ISAF Commander guidance of 0.366, but the second text is shorter in length. At the same time, the average focus of shorter texts, such as a news story published by Reuters's news service, is 0.358, and that of former American President Bush's speech on the Global War on Terror, delivered in April 2006, is about 0.250 [23]. As such, these comparisons suggested that the Taliban propaganda statements and booklet were written in a relatively professional manner when compared with those available in Western governmental and military organizations. This further confirmed the International Crisis Group (2008) report that concluded that the Taliban has developed a complex communications structure able to employ a wide range of media [25].

Civilian casualties has been a topic of high interest for media reporting, and one of the most controversial topics, not because of the moral and humanitarian aspects related to it, but because of the impact civilian casualties had on the situation on the ground, regarding the legitimacy of, and population support for, the Afghan government, international forces, and the insurgency. For these reasons, the analysis gave a special attention to the topic of civilian casualties.

In the overall reporting, the Afghan media tended to correlate civilian casualties with the presence of international (US, ISAF, foreign) forces, while the international media tended to correlate civilian casualties with both the international forces and the insurgents (not Taliban) almost at the same level, with a little more emphasis on the insurgents.

In addition to the analysis of the overall media reporting on civilian casualties, I conducted a special longitudinal analysis of the media reporting of an ISAF incident that occurred on 22 August 2008 in Shindand, northern Afghanistan, which resulted in a high number of Afghan civilian casualties. The analysis showed that, while the media lost initial interest in the subject after a few days following the incident, the ISAF investigation of the incident kept the media writing about it. Every time new information was released to the media, media attention increased, then decreased again in a few days. It increased again when new information became available. In essence, the analysis suggested that the longer the investigation continued until finalised, the longer the media continued to report about the incident and consequently influenced negatively the public perception of ISAF. This is an example of the lack of coherence and sustainability of the ISAF messaging related to the consequence management that followed the civilian casualty incidents resulting from ISAF operations.

7.5 CONCLUSIONS

Overall, at the time of its execution, the analysis presented above enabled the development of a set of conclusions and actionable recommendations to improve the effectiveness of NATO's Public Affairs in Afghanistan. As NATO's Public Affairs releases related to the ISAF mission were driven mainly by the headquarters in Kabul, with some influence by NATO HQ in Brussels, an analysis developed at the out-of-theatre headquarters did not generate any major change, but provided another perspective to all the communication practitioners and decision makers in JFC Brunssum, who understood that the situation was very complex and that more messaging did not provide a solution.

Nevertheless, the use of computer-based content analysis proved a useful approach to assessing the effectiveness of long-term communication campaigns. This analysis process should be considered as soon as possible to improve the conduct of NATO's StratCom activities. Computer-based content analysis has the potential to analyse Big Data in the information environment which could only improve our ability to communicate effectively in a world where more and more data are generated. With various NATO entities currently monitoring communication products of news media, internet communities, and opinion leaders, there are a lot of data that are collected, but are not sufficiently exploited to benefit the Alliance. The improvement of assessing the effectiveness of communication activities will benefit NATO in all areas of StratCom, Public Affairs, Information Operations and Psychological Operations.

7.6 BACKGROUND READING

Calha, J.M. 2015. Hybrid Warfare: NATO's New Strategic Challenge? Report to the NATO Parliamentary Assembly, Defence and Security Committee.

Department of Defence. 2015. Enhancing Security and Stability in Afghanistan, Report to Congress. https://www.defense.gov/Portals/1/Documents/pubs/1225_Report_Dec_2015_-_Final_20151210.pdf [Accessed Date: 24 Sept. 2016].

Lasconjarias, G. and Larsen, J. (editors). 2015. NATO's Response to Hybrid Threats. NATO Defense College. NDC Forum Papers Series, Forum Paper 24

NATO. 2008. Bucharest Summit Declaration. http://www.nato.int/cps/en/SID-D1453

C0E-DCF302EF/natolive/official_texts_8443.htm?selectedLocale=en [Accessed Date: 22 Sept. 2016].

NATO. 2009. NATO Strategic Communications Policy.

NATO. 2016. A "Comprehensive Approach" to Crises. http://www.nato.int/cps/en/natolive/topics_51633.htm [Accessed Date: 25 Jun 2016].

Riffe, D., Lacy, S. and Fico, F. 2005. Analyzing Media Messages, Using Quantitative Content Analysis in Research. Lawrence Erlbaum Associates Publishers.

7.7 AUTHOR'S BIOGRAPHY

Valentin Poponete is a civil-military planner/analyst at SHAPE Strategic Planning. He worked in Public Affairs for more than 12 years. Between 2006 and 2010, worked in the JFC Brunssum Public Affairs Office and was the head of the ISAF Cell in the same office from 2008 until 2010. He deployed two times to the Public Affairs Office of the ISAF HQ in Kabul. Valentin Poponete is a retired Romanian Army officer. He has a BA in Sociology, a Post-university Degree in Public Relations, and a MA in Political Science and is working on his Ph.D. that employs social network analysis to the study of human migration. To be effective, a comprehensive approach to crisis management must be complemented by sustained and coherent public messages. NATO's information campaigns are substantiated by systematic and updated information, documenting progress in relevant areas. Efforts are also being made to share communication strategies with international actors and to coordinate communications in theatre.

7.8 REFERENCES

[1] Balabanova., E. 2007. Media, Wars and Politics. Comparing the Incomparable in Western and Eastern Europe. Routledge.

[2] Smith, R. 2007. The Utility of Force: The Art of War in the Modern World. New York: Alfred A. Knopf.

[3] Santos, R.F. 2013. Strategic Communication in Modern Conflicts: Afghanistan. Revista del Instituto Español de Estudios Estratégicos, No. 2.

[4] Tzu, S. 2000. On the Art of War. Allandale Online Publishing.

[5] von Clausewitz, C. 1812. Principles of War. Translated by Hans W. Gatzke, 1942.

[6] Hallahan, K., Holtzhausen, D., van Ruler, B., Verčič, D. and Sriramesh, K. 2007. Defining Strategic Communication. International Journal of Strategic Communication, 1(1):3-35.

[7] Boudreau, B. 2016. We Have Met the Enemy and He Is Us. An Analysis of NATO Strategic Communications: The International Security Assistance Force (ISAF) in Afghanistan, 2003–2014. Riga, Latvia: NATO Strategic Communications Centre of Excellence.

[8] Tatham, S. and Le Page, R. 2014. NATO Strategic Communication: More to be Done? Policy Paper No. 01. National Defence Academy of Latvia, Center for Security and Strategic Research.

[9] Paul, C. 2011. Strategic Communication: Origins, Concepts, and Current Debates. Praeger.

[10] Holsti, O. 1969. Content Analysis for the Social Sciences and Humanities. Addison-Wesley Publishing Company, Inc.

[11] Newman, I. and DiSalvo, V. 1980. Use of a Computer-Based Content Analysis Technique. Journal of School Health, 50(4):214-217.

[12] Wiedemann, G. 2013. Opening up to Big Data: Computer-Assisted Analysis of Textual Data in Social Sciences. Forum: Qualitative Social Research, 1(42):23.

[13] Diesner, J. and Carley, K. 2005. Revealing Social Structure from Texts: Meta-Matrix Text Analysis as a Novel Method for Network Text Analysis. In Narayanan, V.K. and Armstrong, D., Casual Mapping for Research in Information Technology.

[14] Corman, S. and Dooley, K. 2002. Dynamic Patterns in Terrorism News: Reuters Coverage of 9/11. Presented at Computational Analysis of Social and Organizational Systems Conference (CASOS '02), Carnegie-Mellon University. http://www.casos.cs.cmu.edu/publications/papers/CASOSConf_2002_Day3.pdf [Accessed Date: 25 Aug 2016].

[15] Conroy, N., Rubin, V. and Chen, Y. 2015. Automatic Deception Detection: Methods for Finding Fake News. Proceedings of the Association for Information Science and Technology, 52(1).

[16] Bean, H., Keränen, L. and Durfy, M. 2011. This Is London: Cosmopolitan Nationalism and the Discourse of Resilience in the Case of the 7/7 Terrorist Attacks. Rhetoric and Public Affairs, 14(3): 427-464.

[17] Papacharissi, Z. and Oliveira, M. de F. 2012. The Rhythms of News Storytelling on #Egypt. Journal of Communication, 62:266-282.

[18] Meraz, S. and Papacharissi, Z. 2013. Networked Gatekeeping and Networked Framing on #Egypt. The International Journal of Press/Politics, 18(2):138-166.

[19] Diesner, J. 2012. Uncovering and managing the impact of methodological choices for the computational construction of socio-technical networks from texts. Technical report CMU-ISR-12-101.

[20] Corman, S., Kuhn, T., Mcphee, R. and Dooley, K. 2002. Studying Complex Discursive Systems. Centering Resonance Analysis of Communication. Human Communication Research, 28(2):157-206.

[21] Dooley, K., Corman, S. and Ballard, D. 2004. Centering Resonance Analysis: A Superior Data Mining Algorithm for Textual Data Streams. Report No. STTR 001-1. Arlington, VA: US Air Force Office of Scientific Research.

[22] Crawdadtech. 2009. Crawdad text analysis system 1.2. www.crawdadtech.com [Accessed Date: 2009].

[23] Goodall, L., Cady, L., Corman, S., McDonald, K., Woodward, M. and Forbes, C. 2008. The Iranian Letter to President Bush: Analysis and Recommendations. Corman, Steven, Trethewey, Angela and Goodall, Lloyd (eds.), Weapons of mass persuasion: strategic communication to combat violent extremism. Peter Lang Inc., International Academic Publishers.

[24] ISAF. 2008. Press Release 18-2008, Insurgents killed while placing improvised explosive devices in Panjwayi District. http://www.nato.int/isaf/docu/pressreleases/2008/04-april/pr080401-128.html [Accessed Date: 2009].

[25] International Crisis Group. 2008. Taliban Propaganda: Winning the War of Words? Asia Report, 158.

Chapter 8 – ALIGNING SURVEYS TO THE MISSION: THE ROLES OF PUBLIC OPINION POLLING IN COMPLEX OPERATIONS

Philip Eles
NATO Communications and Information Agency
CANADA

ABSTRACT

This chapter examines the relationship between military operations and the public opinion polling programs designed to support them. It does so through the lens of the NATO survey program in Afghanistan and how it evolved to meet a changing NATO mission which transitioned in 2015 from a focus on counterinsurgency to security force assistance.

This chapter defines five roles that surveys can play in informing operations. Surveys can: provide situation awareness of the strategic operating environment; provide a direct measure of progress against those strategic effects that are related to population perceptions; provide an indirect measure of effectiveness for assessing the impact of other operational tasks not specifically aimed at influencing perceptions; provide a direct measure of effectiveness for assessing those operational tasks which are conducted specifically to influence perceptions; and help establish population demographics to inform other actions. All of these aid the command in making its operations more effective.

Through this framework, the evolution of NATO-sponsored surveys in Afghanistan in 2015 is interpreted as a shift in survey role priority to better align with the new Resolute Support campaign priorities. More generally, this framework can be used to give insights into how best to find the appropriate balance between consistency for measuring important trends and responsiveness to the evolution of information requirements. Specifically, those items related to effectiveness of operational tasks should be responsive to evolving needs, while those items related to the strategic environment and strategic effects require long-term consistency.

8.1 BACKGROUND: SURVEYS IN COMPLEX OPERATIONS

In complex operations, the causal linkages between actions on the ground and intended effects may not be well understood, and actions, both military and non-military, can have additional and often unintended effects. A particularly important example of the complicated link between actions and effects is the impact of operations on attitudes of the citizens of the host nation.

Perceptions of the host nation population are important for a variety of mission types, not just those for which winning over public support is central to strategic objectives (e.g., counterinsurgency). Public support can translate directly into the local population providing logistics support or facilitating the cause of one side or the other, or being more or less receptive to various initiatives and messages. Consider, for example, a dismounted infantry section patrolling a village in which the population supports its mission, versus a village in which the population does not. Similarly, consider the support that insurgents receive from a community which favours the current government versus a community which opposes it. In both cases, public support can be an enabler, while lack of support can represent a threat. The end-game of military interventions often requires transition to a host nation lead, and ultimately the public's acceptance of the conflict outcome.

Public perceptions impact operations at all levels from the tactical to the strategic. Understanding those perceptions is therefore key to a commander's awareness of the battlespace as well as his ability to operate within it. Public perceptions are therefore inextricably tied to the effectiveness of operations.

This is especially true for counterinsurgency operations in which public perceptions are the centre of gravity of a campaign that seeks ultimately to win over support for a host nation government and to undermine and erode support for the insurgency [1]. However, understanding the complex interactions between counterinsurgent actions and effects in counterinsurgency has long challenged military planners and assessors [2]. Surveys give commanders a direct measure of public attitudes, and are therefore a means of gauging the downstream effects of counterinsurgency efforts.

8.2 BACKGROUND: NATO-SPONSORED SURVEYS IN AFGHANISTAN

In Afghanistan, surveys have long played a key role in supporting International Security Assistance Force (ISAF) assessments of the progress of counterinsurgency efforts. Over the course of the ISAF mission, survey programs supported commanders at all echelons, ranging from regional polling programs sponsored by troop contributing nations in support of national Task Forces and Regional Commands, to several nation-wide opinion polling programs which informed 3- and 4-star ISAF Headquarters in Kabul [3], [4], [5]. One of the latter survey programs, tailored to the information requirements of Commander ISAF, was the Afghan Nation-wide Quarterly Assessment Research (ANQAR) program.

Starting in 2008, ANQAR collected nation-wide public opinion data on a quarterly basis, covering topics on a range of issues related to Governance, Reconstruction and Development, and Security, the main ISAF lines of effort. The NATO-sponsored ANQAR program partnered with a private sector company for all aspects of data collection, while operations analysts within ISAF's Afghan Assessment Group (AAG) provided data analysis as well as project oversight and coordination.

At the start of 2015, when NATO transitioned from ISAF to Resolute Support (RS), the campaign shifted focus from counterinsurgency to security force assistance, representing a shift from tactical support of a nascent Afghan security force in fighting a counterinsurgency, to the provision of training, advice, and assistance to Afghan security institutions (e.g., ministries) and to the Afghan Army and Police at the corps and police zone levels and above. The RS OPLAN [6] outlined new operational objectives and with these came new information requirements. A smaller NATO mission also meant fewer in-theatre resources dedicated to, among other things, campaign assessment, which resulted in significant technical and contracting aspects of ANQAR being managed as a reach-back task from NATO bodies in Europe. At the same time, the security environment in Afghanistan deteriorated, the Afghan government changed in a disputed election, and ISAF retrograde negatively impacted the Afghan economy.

In response to a changing operation and a changing environment, the ANQAR program evolved. By examining how and why it changed, we can gain insights into the relationship between a survey program and the operation it supports.

8.3 THE ROLES OF SURVEYS IN SUPPORT OF A MISSION: A FRAMEWORK

Before we examine Afghan surveys and how they changed as the mission evolved, we outline a general framework that describes the roles that surveys play in informing operations. We use this framework as a lens through which to examine the Afghan survey program. This framework can inform planners for future survey efforts.

Public opinion surveys can inform operations in a number of ways depending on the mission type, the strategic and operational objectives, the commander's priorities, and the specific information requirements of the staff. Understanding the various roles of surveys in informing the campaign can help analysts to design a survey that is appropriately balanced for the mission or to modify a survey in response to an evolving mission.

The following five roles of surveys reflect how they are used to inform operations. They are listed here in order from strategic to tactical, and described in more details in the subsequent sections.

- Role A: Surveys Can Provide Situation Awareness of the Strategic Operating Environment.
- Role B: Surveys Are a Direct Measure of Progress Against Those Strategic Effects which are Related to Population Perceptions.
- Role C: Surveys Provide an Indirect Measure of the Effectiveness of Those Operational Tasks which are not Specifically Aimed at Influencing Perceptions.
- Role D: Surveys Provide a Direct Measure of the Effectiveness of Those Operational Tasks which are Conducted Specifically to Influence Perceptions.
- Role E: Surveys Can Be Used to Determine Population Demographics.

8.3.1 Role A: Surveys Can Provide Situation Awareness of the Strategic Operating Environment

Surveys can give insights into the operating environment in much the same way as other intelligence gathering techniques. The population's support for insurgent groups, satisfaction with local or national government, economic issues, and expectations for the future all shape the operating environment, even if they do not have a direct linkage to operational objectives. Surveys, therefore, give a commander an additional "sensor" to monitor the strategic space in which he is operating and inform his Commander's Critical Information Requirements.

In principle, the understanding of the strategic environment also informs the operational planning processes by informing mission assumptions and strategic risks. In practice, however, survey data are rarely available for a new mission (the exception being the case of a follow-on mission).

A good recent example from Afghanistan is the use of surveys to measure public perceptions of the Afghan National Unity Government (NUG), the political coalition assembled to resolve an electoral impasse following the 2014 presidential elections. The NUG's fragility or stability is a key risk to NATO's Resolute Support Mission, though it is not directly related to strategic effects or operational objectives, nor are there operational tasks aimed at affecting it. Public perceptions of the government influence the strategic space in which the commander operates.

8.3.2 Role B: Surveys are a Direct Measure of Progress Against Those Strategic Effects Which are Related to Population Perceptions

In some operations, strategic objectives and strategic effects explicitly involve influencing population perceptions. In this case, attitudes of the population are the relevant measure of effectiveness, and surveys provide the relevant metrics by which progress is marked.

The notable example is counterinsurgency operations, which explicitly seek to influence perceptions such that the population rejects the insurgency and accepts the legitimacy of the government [1]. By measuring public support for insurgents and the government, surveys provide a direct measure of progress against strategic effects.

Unfortunately, in complex environments, when cause and effect are not well understood, such high-level assessments of campaign progress are difficult to translate back into operational plans to improve operational effectiveness since the actions required to change perceptions are non-linear and often not fully understood. That is to say, because causality is difficult to prove, it is often impossible to establish whether those effects were a result of military actions, or whether they were a produce of other exogenous factors. In particular,

when certain actions don't produce the expected effects, it may be a result of external factors or other indirect mechanisms which were not considered in the original planning. This is a general feature of complex operations, where the connection between actions and effects is non-trivial, and especially when it comes to public perceptions, where individual opinions are shaped by many factors.

While counterinsurgency represents the archetypical example of an operation whose objectives involve influencing public perceptions, recent NATO activities in response to the crisis in the Ukraine also seek to "reassure the populations" of NATO nations regarding Alliance cohesion, readiness, and resolve [7].

8.3.3 Role C: Surveys Provide an Indirect Measure of the Effectiveness of Those Operational Tasks Which are not Specifically Aimed at Influencing Perceptions

Many operational tasks are not conducted with the primary intent of influencing population perceptions. However, they may have an impact on public perceptions. Therefore, surveys may provide an indirect, or proxy, measure of effectiveness for those actions, and can reflect on the success or failure of those tasks, or on the unintended consequences of friendly force actions in executing those tasks.

For example, in Afghanistan, operational tasks with unintended effects on public opinion include airstrikes, route clearance operations, host nation security force capacity building efforts, or infrastructure development projects. Collateral damage from airstrikes alienates the population – an unintended negative effect of an operational task which may undermine other efforts.

As with strategic effects, it is difficult to prove a causal relationship between tactical actions and their effects, though the focused and localized nature of many tactical actions means that it may be easier to reasonably infer that certain actions resulted in certain effects – for example, that airstrikes causing civilian casualties contributed to worsening in public opinion of international forces in the area. Unfortunately, for nation-wide surveys the sample sizes at the local level are often small, so that margins of error are high, making it difficult to infer tactical effects from survey results (as discussed in more detail below).

8.3.4 Role D: Surveys Provide a Direct Measure of the Effectiveness of Those Operational Tasks Which are Conducted Specifically to Influence Perceptions

Tasks directly related to influencing public perceptions are performed by, for example, Information Operations (IO), Strategic Communications (StratCom), and Public Affairs (PA) staff as well as elements of the Intelligence Branch (J2). These tasks include, for example, sponsoring public media campaigns. A survey is a tool which provides a direct measure of public opinion, and therefore can aid in assessing the overall impact of messaging efforts. Survey data can also be used to inform IO planners about the target audience's use of media (radio, television, print) in order to maximise message penetration.

For example, in Afghanistan, media campaigns were run to make Afghans aware of a new telephone hotline established to facilitate the reporting of insurgent activity such as the emplacement of IEDs. Surveys were able to measure public awareness of the hotline, awareness of the number to call, as well as the reasons why people might choose not to use it. Surveys were able to demonstrate an increased awareness in those areas where media campaigns were conducted, suggesting that the campaigns were effective. In another example, surveys improved the effectiveness of media campaigns aimed at improving recruitment of Afghan National Army and Police by examining public sentiment towards those organizations.

8.3.5 Role E: Surveys Can be Used to Determine Population Demographics

In addition to gauging public opinion, surveys also collect demographic data including age, ethnicity, education (secular and religious), household income, occupation, and employment status. While demographic data are typically used to examine differences in opinions between various demographic

groups, they can also be used to construct demographic profiles of geographic areas. Demographic data can be used to support target audience analysis to improve planning of IO media campaigns, or for example, to examine ethnic distributions to support planning.

For example, in advance of the 2014 Afghan presidential elections, 23 waves of ANQAR survey data were combined to produce an ethnic map of Afghanistan which included ethnic distributions down to the district level, the most detailed ethnic breakdown to date. The results were used to examine the average level of violence seen by ethnic groups, revealing that a pre-elections spike in violence affected some ethnic groups more than others. In another example, the reported level of education of young Afghan women were observed increasing over time, both within Kabul and across the rest of Afghanistan.

8.4 HOW SURVEYS INFORMED THE ISAF/RS MISSION

The ANQAR program attempted to fulfil *all* of the above roles to varying degrees. The survey became an instrument to answer a wide range of questions for a wide range of customers both within the headquarters and at higher headquarters. Certain survey questions addressed specific information requirements, and as a whole, the ANQAR questionnaire tried to satisfy a broad range of users within the survey stakeholder community.

Commanders and their planning staff were typically most interested in understanding the impacts of their immediate actions (Roles C-D), while higher headquarters and national political leadership were often primarily interested in the strategic environment and progress towards stated strategic objectives or an assessment of strategic risks (Roles A-B). Information Operations and StratCom staff typically were most interested in Roles D-E to better understand the immediate impacts of their specific messages, and to plan effective future messaging campaigns.

Assessment staff typically utilized the survey data across the full spectrum (Roles A-E) because they were often required to produce assessment products that satisfied both operators, by providing feedback to make operations more effective, and higher headquarters, by providing an assessment of long-term campaign progress. As a result, in Afghanistan, assessments cells were typically the main user and owner of Afghan surveys. Assessments cells typically also had operational analyst staff members with the technical skills and software tools required to conduct statistical analysis of survey data.

8.4.1 Managing Evolving Information Requirements in a Changing Mission

Military operations can evolve over time for a variety of reasons, such as when strategic and operational objectives are reprioritized or redefined, when resource levels change, when the strategic environment changes causing new issues to come to the fore, or when commanders and staff turnover. In order to remain relevant, surveys must remain aligned to evolving information requirements of survey stakeholders.[1] When ISAF transitioned to Resolute Support, the information requirements changed as well, which necessitated a re-evaluation of the survey questionnaire. So, on the one hand, surveys needed to change with the changing mission.

On the other hand, consistency is key to the success of a longitudinal surveys program. Maintaining consistency in how a survey is implemented (i.e., sampling method, sampling plan, etc.), and consistency in the questions that are asked allows for measurement of changes in perceptions over time. Measuring change over time is fundamentally more important than a single measurement obtained in a single survey. This is because results from a single survey contain inherent biases, including response biases (e.g., social

[1] Survey stakeholders are a community of both in-theatre and out-of-theatre actors whose information requirements are addressed by the survey. They contribute to the development of the survey instrument (questionnaire) and are users of the survey data. Typically, stakeholders include assessors and planners at various levels of the command structure (including subordinate and higher headquarters) as well as the information operations, strategic communications, and intelligence elements. Strong engagement of the survey stakeholders is key to the success of survey programs.

desirability bias, anchoring bias, etc.) [8] and sampling biases introduced by the sampling strategy employed (i.e., multi-stage clustered sampling) [9]. However, because these biases do not change significantly over time, results from consecutive surveys should be equally biased and therefore changes over time should reflect a real change in the environment, independent of biases. Survey information is therefore most informative for assessments when changes over time are examined, and consistency is key.

The pull between these two seemingly opposing pressures – to remain responsive to a changing mission while remaining consistent in order to track long-term trends – is an enduring challenge of survey programs [10], [11].

The key to finding the balance in this apparent contradiction is in understanding the roles of the survey as outlined in the framework described above. Consistency is most necessary in those questions which address long-term progress against strategic objectives (Role B) and the monitoring of strategic space (Role A), as well as in tracking demographic information (Role E). Responsiveness is most necessary for those areas where surveys are meant to inform on the impacts of specific actions or tasks (Roles C, D).

It is the responsibility of analysts to understand how survey questions are used, and which role they fulfil. Strong engagement by analysts with the full range of survey stakeholders is therefore critical for a survey program. A formal process is required for engaging with the survey stakeholder community for coordinating modifications to the survey questionnaire, in particular when the mission changes. Requested changes should come with either formal or informal justifications regarding why they are required. Based on the requests from all stakeholders, and as assessed against the spectrum of survey roles, survey analysts must present firm recommendations to senior leadership of the unit responsible for the survey regarding the acceptance or rejection of requested changes.

The utility of surveys at the strategic level should not be underestimated. Even if strategic commands do not explicitly provide input to questionnaire development, operational-level survey analysts must be aware of those questions which tend to be included in strategic reports to political leadership (e.g., to the North Atlantic Council or U.S. Congress). The challenge is that operational analysts often have limited visibility on how survey results are used by higher headquarters, particularly if those analysts are embedded within an in-theatre headquarters and are not privy to how survey results might be used by others (in or out of theatre) to inform decision making at other echelons of command.

8.4.2 The Evolution of Afghan Surveys from ISAF to Resolute Support

In the pre-2015 ISAF era the justification for NATO's sponsorship of surveys in Afghanistan was straightforward. In a counterinsurgency operation, population perceptions were a key downstream effect of NATO's operations: insurgent groups which could not be defeated by purely militarily means were to be defeated by eroding popular support for their cause and winning over support for the government through reconstruction and development efforts and the restoration of governance structures and basic services. Strategic effects were therefore directly tied to the population's support for the Afghan government and their rejection of the insurgency.

In 2015, ISAF transitioned to Resolute Support (RS), a security force assistance mission which focused on building capacity within Afghan security forces at the corps-level and above, as well as within Afghan security institutions (e.g., Ministries of Defence and Interior). As ISAF forces drew down across Afghanistan, the Taliban shifted away from asymmetric tactics (e.g., IED's) to more conventional tactics (e.g., direct engagements) against a near-peer adversary, and ANDSF operations evolved from heart-and-minds counterinsurgency towards more kinetics operations. Public perceptions were no longer at the centre of the mission as the overall focus shifted to developing the operational capability of Afghan National Defence and Security Forces (ANDSF).

The ANQAR surveys were no longer directly tied to a key strategic objective, and the importance of surveys for Role B was reduced significantly when ANQAR transitioned to supporting RS. Some questions that had previously directly addressed strategic effects (Roles B) became measures of the strategic environment (Role A).

In the transition to RS, the NATO footprint was reduced dramatically. Compared to ISAF where coalition forces were conducting counterinsurgency activities across Afghanistan, the RS posture was Kabul-centric with a limited regional presence. Beyond Kabul, regional Train, Advise, and Assist Commands (TAAC) provided advisors to ANDSF units at corps headquarters. Therefore, what had been ISAF's interest in tactical/operational effects on the battlefield (Role C) shifted with RS into an interest in monitoring second order effects of ANDSF operations (a more indirect form of Role C). The focus of the theatre commander shifted from local to nation-wide issues.

With a reduced coalition footprint, commanders had fewer "sensors" to provide an awareness of the operating environment; including fewer patrols, fewer engagements with local leaders, and fewer intelligence collection assets. The remaining sensors, including surveys, became that much more important in providing commanders with situational awareness (Role A). Thus, in the shift to RS, ANQAR surveys grew in relative importance as a "sensor", informing the strategic picture and supplementing other intelligence and information gathering techniques.

The role of surveys in informing Information Operations (Role D) remained largely unchanged, though the focus of those operations shifted. Resolute Support IO and support to Afghan public awareness campaigns shifted to focus on influencing public perceptions of the ANDSF and the National Unity Government. In tailoring their media campaigns, IO operators used surveys to better understand the perceptions related to the recruitment of ANDSF soldiers and police, and recruitment of women in particular. Questions related to Roles D and E therefore retained or increased their relative importance in the RS mission.

Figure 8-1 depicts conceptually the overall shift in the relative roles of ANQAR surveys from the ISAF era to the Resolute Support era, where the size of the bubbles reflects the relative importance of each role to the overall mission. As discussed above, overall the shift was a decrease in ANQAR's importance in assessing progress against strategic objectives (Role B) and operational objectives (Role C), and increase in its role in providing general information about the operating environment (Role A), with little change in its importance in supporting the assessment of Information Operations (Role D) and DEMOGRAPHICS (Role E).

As a result of the above changes in information requirements within RS, ANQAR stakeholders reviewed the questionnaire in early 2015. Many of the questions related to Role C were removed, particularly those that dealt with perceptions of tactical issues (e.g., perceptions of district governor, satisfaction with local schools and healthcare). New sections were added addressing emerging issues such as checkpoint extortion, migration awareness and intention, government accountability and corruption, and perceptions regarding ANDSF directly related to recruitment efforts.

8.4.3 Understanding the Challenges that Afghan Surveys Have Faced

Afghan surveys have by no means been perfect. The ANQAR program has faced challenges over the years, and has had its detractors. In Afghanistan, nation-wide perceptions have not changed as dramatically as one would hope after a decade and a half of counterinsurgency operations. This apparent lack of progress has led some to criticise survey methods (e.g., biases, regional coverage, etc.) rather than trying to understand why counterinsurgency efforts have not positively affected public perceptions, or why any local effects would not have been measured by a nation-wide survey.[2]

[2] In fact, the general lack of measurable progress with the NATO mission in Afghanistan has been a significant challenge for the assessment community. As a result, rather than (or at least in addition to) reflecting on the root causes of this lack of progress (e.g., unintended consequences of certain aspects of ISAF operations which alienate the population as much as they win it over), many detractors have attacked the assessment and survey methodologies. In part, this may be responsible for the high turnover in assessment frameworks.

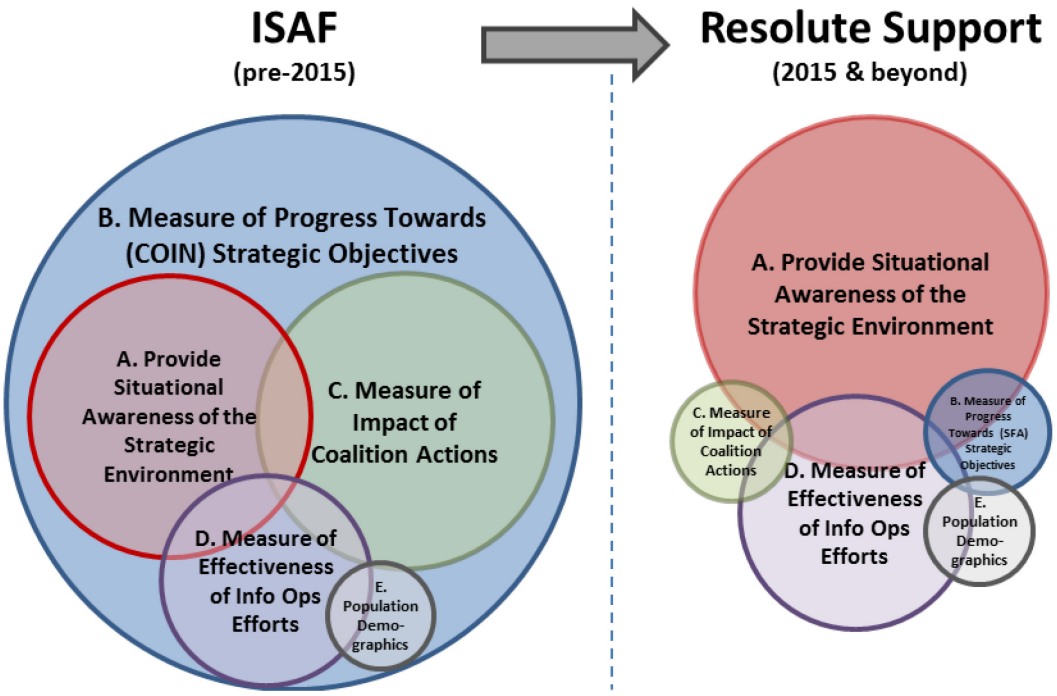


Figure 8-1: Changing Importance of Survey Roles from ISAF to RS.
Note: Size of the bubble reflects its importance to the mission.

To some extent, the lack of impact of ISAF's regional counterinsurgency efforts on nation-wide or province-wide perceptions may be unsurprising. Even at its peak, ISAF's counterinsurgency efforts were not nation-wide but rather focused in "Key Terrain Districts" and on "Village Stability Operations" (i.e., at district and village level in select areas). ANQAR, however, was designed as a general barometer of public perceptions at the nation-wide and provincial level and was never intended to inform tactical commanders. Small sample sizes at the sub-provincial level often precluded any meaningful analysis to support measurement of change.[3] Any local effects of ISAF operations would be washed out in the provincial or nation-wide average by the voices of those who had not experienced the insurgency and ISAF's counterinsurgency efforts first hand. Other survey programs sponsored by Regional or Task Force commanders and conducted by troop contributing nations had more success in linking local actions to effects on local perceptions. Even then, areas where the insurgency was most violently contested remain inaccessible to survey teams, and therefore the direct impact of counterinsurgency operations on those at the centre of the conflict remains elusive.

Another criticism which has been levelled against ANQAR surveys was its limited ability to provide actionable information that could inform planners. However, counterinsurgency operations have a complex linkage between actions and effects, particularly with respect to public perceptions. For example, see the so-called "Spaghetti Diagram" which attempted to articulate those linkages [12]. Second and third order effects of actions, as well as unintended consequences, or counter-coalition insurgent propaganda can all have as much of an impact on public perceptions as the counterinsurgency efforts themselves. It is therefore difficult to assess whether particular actions in the battlespace translated into a change in public perceptions (or lack of change), and therefore difficult to feed back to planners regarding which effects were and were not successful.

Furthermore, critics sometimes questioned the utility of certain survey questions, or the specific question formulation, as being either too abstract, too vague, or even too specific. Questions specifically tailored to

[3] That being said, in each survey ANQAR sample sizes were increased in those provinces and districts that were the focus of ISAF operations at the time. In each survey, a set of "booster" samples was allocated to key regions of interest.

address one of the roles discussed earlier, often do not appear to be good questions when consider from the perspective of a different role. Strategic questions (Roles A, B) may not make much sense to those trying to evaluate operational effects (Roles C, D), and vice versa. Furthermore, those who criticise the specific question formulation (i.e., wording) often overlook the fact that it is the Dari and Pashto language versions that are fielded, while the English translation may not exactly reflect question wording.

Understanding the various roles of surveys is key in helping extract information most relevant to each stakeholder. It is also important in helping to understand why certain questions are asked, when this may not be obvious from some perspectives.

8.5 WHERE NEXT – THE FUTURE OF NATO SURVEYS?

8.5.1 Institutionalising NATO Surveys

Some nations have successfully institutionalised the capability to conduct surveys in support of military operations. Notably, the U.S. Centre for Army Analysis (CAA) maintains a core survey capability and an ability to train and deploy operational analysts who receive basic training on survey analysis through CAA's Deployed Analysts Course. Valuable lessons can and have been gleaned from these national efforts, and as they support different mission types, these lessons are valuable and complementary to those observed from the NATO experience in Afghanistan. In contrast, NATO's survey capability is more fragile, as is its deployable operational analysis capability in general. Over the past eight years, NATO-sponsored surveys in Afghanistan have relied on a mixture of analysts scattered across the NATO Command Structure (e.g., JFC Brunssum, ACT), within NATO's technical agencies (e.g., NATO Communications and Information Agency), and through Crisis Establishment (CE) civilian positions in theatre. As such, NATO has treated surveys as an ad hoc capability that can be developed on an as required basis by tapping into existing operational analysis communities within NATO. Industry is seen as the primary source of expertise to fulfil any requirement for surveys in future missions. However, it is not well recognized that an analytic capability internal to NATO is required. Such internal expertise is indeed needed in order to articulate technical survey requirements, to independently evaluate survey methods, and to conduct detailed, focused, timely, and often sensitive or classified analysis of the survey data to feed directly into tailored products for commanders. Therefore, NATO requires an institutionalised capability to conduct survey analysis.

8.5.2 Afghan Surveys Beyond Resolute Support

NATO's mission in Afghanistan continues to be extended year after year, and to date so have ANQAR surveys. At some point in the future, it is likely that NATO's military contribution to Afghanistan will further decline, possibly replaced by a civilian-led presence. Regardless, NATO's interest in its legacy in Afghanistan will likely outlast its physical presence. Surveys can, in principle, continue to provide NATO with a means to gauge the long-term impact of its investment in Afghanistan beyond the current mission.

The considerable dataset collected through the ANQAR programme will remain an important and useful baseline against which future efforts are judged. Future surveys in Afghanistan can also be a bellwether to give strategic warning of imminent changes in the operating environment for a future mission.

Whether ANQAR ends with Resolute Support or it continues beyond, it is important that during the lifetime of the project, while all pieces are still in place, the lessons are adequately captured and documented. This chapter represents one output of this effort. The hope is that in years to come, such documentation will prove to be a valuable reference to analysts involved in providing support to perception surveys that would complement the body of analytical products on public perceptions produced for commanders.

Documentation of best practices and lessons learned represents an initial step towards institutionalisation of the survey capability within NATO. These best practices are already being incorporated into formal documents such as the NATO OA Handbook [10] and Deployed Analyst's Handbook [11].

8.5.3 NATO Surveys Beyond Afghanistan

Opportunities currently exist to leverage the expertise and partnerships created through the NATO's survey experience in Afghanistan, to expand survey efforts to other NATO activities. Surveys can inform planners and assessment staff for ongoing NATO operations, as well as aid in monitoring of emerging crises where surveys might be used as early warning indicators of conflict, or for pre-intervention baselining.

For example, surveys may be a potentially useful tool to help NATO assess the effectiveness of its responses to the crisis in Ukraine, NATO's Assurance Measures [7]. These efforts, among other things, seek to influence the perceptions of citizens of NATO nations regarding Alliance cohesion, readiness, and resolve. Interest in using surveys has been additionally fuelled by independent surveys conducted by the Pew Research Centre of public perceptions of the Ukraine crisis within NATO nations, within the Ukraine, and within Russia [13]. Proposals for assessment frameworks have included population perception metrics.

As another example, amidst the 2015 migrant crisis in Europe (to which Afghans fleeing conflict are the second highest contributor behind Syrians), ANQAR's recent results regarding Afghan perceptions of migration issues, including awareness of migration from communities, desire to leave, intended destinations, and drivers underlying desire to leave, have sparked a broader interest in surveys as a means to inform decision makers on topics beyond those directly related to the mission in Afghanistan.

It remains to be seen whether these opportunities to leverage the Afghan experience are realized, but they remain another avenue through which the NATO survey capability can be institutionalised, or at least retained while true institutionalisation happens through changes to doctrine.

8.6 CONCLUSION

Designing a survey program to support complex operations, and guiding that program such that it remains responsive to the needs of the mission is an enduring task. Analysts must understand the roles that the survey plays in informing operators, planners and assessors, as well as commanders and higher headquarters. This is an important first step to ensuring that the various information requirements of the headquarters are adequately met, and that the right balance is found between survey responsiveness and continuity. Ultimately, analysts must be intimately familiar with the mission, the decisions which surveys inform, and the limitations of the dataset itself.

8.7 AUTHOR'S BIOGRAPHY

Philip Eles, PhD, is an operational analyst at the NATO Communications and Information (NCI) Agency in The Hague, Netherlands, where he has been providing analytical support for strategic assessments of Afghanistan since 2013 through projects for SHAPE, JFC Brunssum, and ISAF/Resolute Support. He has deployed on five occasions to the Afghan Assessment Group (AAG) in Kabul, and is part of NCI Agency's Reach-Back Support Team for the AAG. Since 2014, he has been the Technical Lead on ANQAR, NATO's survey program in Afghanistan. Between 2006 and 2013, as a Defence Scientist at Defence Research and Development Canada, Dr. Eles worked on assessments to support to the Canadian combat mission in Kandahar. He was one of the lead analysts for Canada's Kandahar Public Opinion Polling (KPOP) program from 2007 to 2011. Dr. Eles holds a Ph.D. in Physics from the University of British Columbia in Vancouver.

8.8 REFERENCES

[1] NATO. 2011. Allied Joint Doctrine for Counterinsurgency (counterinsurgency), [AJP-3.4.4]. Brussels, Belgium: NATO Standardization Agency, February 2011. https://info.publicintelligence.net/NATO-Counterinsurgency.pdf.

[2] Connable, B. 2012. Embracing the Fog of War: Assessment and Metrics in Counterinsurgency. RAND Corporation. http://www.rand.org/content/dam/rand/pubs/monographs/2012/RAND_MG1086.pdf.

[3] Eles, P.T., Vincent, E., Vasiliev, B. and Banko, K. 2012. Opinion Polling in Support of the Canadian Mission in Kandahar: Final Report for the Kandahar Province Opinion Polling Program, Including Program Overview, Lessons, and Recommendations. Defence R&D Canada Centre for Operational Research and Analysis, Technical Report DRDC CORA TR 2012-160U. September 2012.

[4] Eggereide, B., Martinussen, S.E., Marthinussen, E. and Barstad, A. 2012. Faryab Survey – wave 6. Norwegian Defence Research Establishment (FFI), FFI-rapport 2012/02506, December 2012. https://www.ffi.no/no/Rapporter/12-02506.pdf.

[5] Purslow, N. 2013 Establishing a Public Perceptions Polling Capability in the Early Stages of Deployment. DSTL Policy and Capability Studies, DSTL/705716 1.0, March 2013.

[6] Resolute Support. 2014. Resolute Support (RS) Operations Plan (OPLAN) 38312. Headquarters International Security Assistance Force (ISAF).

[7] NATO. 2015b. NATO Assurance Measures Factsheet. Supreme Headquarters Allied Powers, Europe. http://shape.nato.int/land and http://shape.nato.int/nato-assurance-measures.

[8] Banko, K. 2013. Public Opinion Polling in Operations Assessment with Examples from Afghanistan. In A. Williams et al. (Eds.), Innovation in Operations Assessment: Recent Developments in Measuring Results in Conflict Environments.

[9] D3 Systems, Inc. 2016. Methods Report, ANQAR Wave 31. March 2016.

[10] NATO. 2015a. NATO Operations Assessment Handbook version 3.0, Annex E: Polling to Support Operations Assessment. Norfolk, VA: Supreme Allied Commander Transformation. July 2015. pp. E1-E10.

[11] United States Army. 2015. Deployed Analysts Handbook, Chapter V: Public Perception Surveys: Development and Analysis. U.S. Army, Centre for Army Analysis (CAA).

[12] PA Consulting Group. 2009. Afghanistan Stability / Counterinsurgency Dynamics (aka the Spaghetti diagram). http://msnbcmedia.msn.com/i/MSNBC/Components/Photo/_new/AfghanistanDynamic_Planning.pdf.

[13] Pew Research Center. 2015. NATO Publics Blame Russia for Ukrainian Crisis, but Reluctant to Provide Military Aid. http://www.pewglobal.org/files/2015/06/Pew-Research-Center-Russia-Ukraine-Report-FINAL-June-10-2015.pdf.

Chapter 9 – BEYOND DESCRIPTIVE STATISTICS IN SURVEY ANALYSIS: PRACTICAL EXAMPLES FROM NATO-SPONSORED SURVEYS IN AFGHANISTAN

Philip Eles
NATO Communications and Information Agency
CANADA

ABSTRACT

The conduct of surveys in conflict zones can be challenging, and may be constrained by the security environment, by cultural aspects of data collection, and by a limited information on population distributions and population displacement. Survey methods may stray from textbook examples. Analysis of the resulting survey data, and its use in informing military operations (including in operations assessments), is itself a challenge.

This chapter is meant to give operational analysts and assessment staff a practical perspective on some of the real challenges of implementing surveys in the field and of analysing the resulting data. It covers some recent developments in NATO-sponsored surveys in Afghanistan, such as dealing with a changing security environment, getting more out of survey data, and complementary data collection efforts. Topics covered herein are intended to complement and enhance the existing resources which have been developed for deploying operational analysts who may be tasked with oversight of a survey program in support of complex operations.

9.1 INTRODUCTION

In the context of military operations, public opinion polling data can provide commanders and their staff with situational awareness of the human terrain within the area of operation. Surveys also provide a measure of effectiveness for assessing progress against strategic effects and operational objectives, particularly those that are directly linked to public perceptions.[1] Because of this link between public perceptions and assessing progress, operational assessments cells within deployed headquarters are often the main consumers and sponsors of survey data collection efforts. Assessments cells typically also have operational analyst staff members with the technical skills and software tools required to conduct statistical analysis of survey data. However, operational analysts deployed to support assessments cells are rarely adequately prepared to take on the responsibility of managing a survey program as part of their responsibilities.

On the surface, the analysis of survey data is simply a calculation of the proportion of respondents expressing certain views. Through the appropriate application of population-proportional weights, inferences can be made about the population as a whole. In reality, the roles and responsibilities of analysts responsible for managing a survey program go far beyond the calculation and reporting of simple descriptive statistics. These responsibilities include both technical and administrative aspects of survey design, implementation, analysis, and reporting.

The conduct of surveys in conflict zones in support of military operations can be challenging and methods may stray far from textbook examples [1], [2]. Managing survey programs requires an understanding of

[1] For a detailed discussion on how surveys support complex operations, see the companion chapter by the same author in this volume on the Roles of Surveys in Support of Complex Operations.

complex factors ranging from dealing with statistics governing complex sampling methods; to managing data collection in a changing security environment which may impact accessibility to the population for field teams; to coordinating information requirements across a broad stakeholder community who use survey results in a myriad of ways (see the author's other chapter in this volume).

Deployed operational analysts may find themselves tasked with managing a survey program for the duration of their deployment, often taking over from predecessors who had themselves inherited the program from others. Many may feel underprepared for technical and administrative aspects of the task.

Existing pre-deployment courses and practical guides for deploying analysts [3], [4] have done well to address many of the basic training requirements in the area of surveys. Deploying analysts have also been able to rely on the knowledge and experience of those industry partners contracted to conduct surveys – industry partners who make themselves available at pre-deployment courses and/or through in-theatre base visits by their local or visiting staff. However, little has been written specifically for deploying analysts about the practical aspects of survey design and implementation.

This chapter aims to incrementally add to the available literature in order to better help prepare analysts to support survey work in complex operations. It describes several practical aspects of survey data collection and analysis from NATO's recent experiences in Afghanistan through the Afghanistan Nation-wide Quarterly Assessment Research (ANQAR) program. ANQAR has been collecting quarterly nation-wide survey data (n~12,000 interviews) every three months since mid-2008.

9.2 COLLECTING SURVEY DATA IN A CHANGING SECURITY ENVIRONMENT

Consistency in survey design and implementation is key to measuring changes in perceptions over time; consistency helps overcome many of the response biases and sampling biases inherent in surveys conducted in conflict zones.

However, in active conflict zones such as Afghanistan, one factor for which it is impossible to maintain consistency is the security situation. The ability of civilian (host nation) survey teams to access certain areas due to insecurity is not under the control of survey designers, and is likely to change over time.

When certain segments of the population are inaccessible to survey teams, they are in effect excluded from the sampling frame. In essence, inaccessibility means that survey results are not nationally representative – results cannot be expressed as "X% of Afghans say Y", but rather as "X% of Afghans living in accessible communities say Y". In principle, as long as the accessible population does not change significantly over time, survey results always represent a consistent sub-population of Afghanistan (albeit a potentially non-representative sub-population). However, if the size of the accessible population were to change significantly, it would be impossible to know whether changes in measured perceptions were due to real changes in public sentiment or just due to a change in who was surveyed.

The above description represents a real and practical limitation in survey design and implementation in a conflict zone. The degree to which it compromises the data quality is situation dependent, and there are ways to overcome the bias through the use of intercept interviews as was done for ANQAR as we describe now.

9.2.1 Dealing with Changing District Accessibility Through Intercept Interviews

In ANQAR, the proportion of districts inaccessible to survey teams gradually increased over time as the security situation deteriorated starting in mid-2015 during the first summer fighting season after the withdrawal of ISAF forces from across Afghanistan. Figure 9-1 shows the districts inaccessible to survey teams, and how that has changed over time.

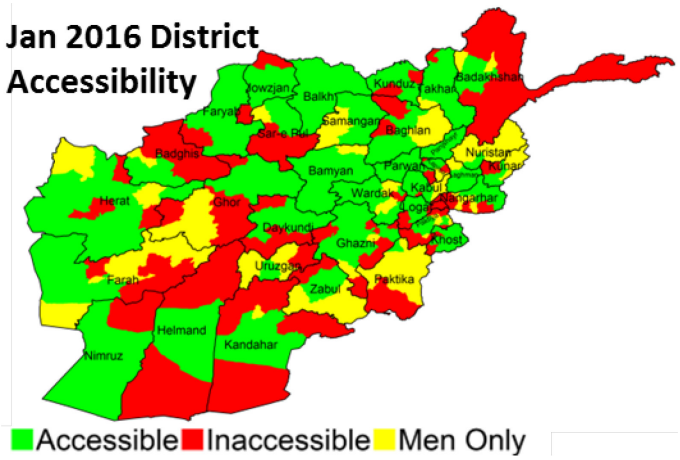


Figure 9-1: Districts Accessible to ANQAR Survey Field Teams in Jan 2016.

Up until 2014, individuals living in inaccessible districts were not included in the ANQAR sample. As a result, those areas where the insurgency was most highly contested did not have a representative voice in the nation-wide poll.

While on the surface, this may seem unsettling, it should be kept in mind that the inaccessible districts also tend to be very sparsely populated. Densely populated areas of Afghanistan, such as major towns and population centres, have consistently remained open to survey teams. Figure 9-2 shows the percentage of the nation-wide population not directly accessible to survey teams (green line), compared to the number of districts not accessible (brown line). While the number of inaccessible districts has increased considerably since 2013, the newly inaccessible districts have been in areas that are not heavily populated, so that the size of the inaccessible population has not increased as significantly, at least until mid-2015. The exclusion of those sparsely populated insecure districts which are not accessible to survey teams does not greatly impact survey results at the national aggregate level. The impact is greater when regional results are reported (e.g., provincial results), though due to larger margins of error for the smaller sample sizes, sub-national results are less frequently reported and are used with caution.

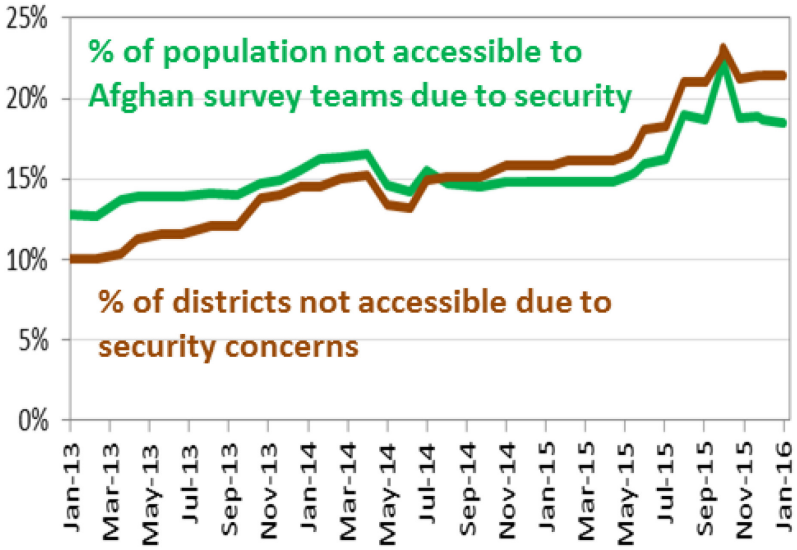


Figure 9-2: Change in Districts' Accessibility and Population Accessibility Over Time.

Recognizing these issues, the ANQAR sampling strategy was modified in 2014 in order to better reflect nation-wide sentiments. "Intercept interviews" were incorporated more broadly into the survey data collection.[2] Intercept interviews are conducted by speaking with residents of inaccessible districts while they are travelling on routes into and out of their districts [5]. Intercept interviews are conducted by male interview teams only and performed any time a sampling point is chosen inside a district that is not accessible. As of Dec 2015, intercepts made up 16% of ANQAR interviews.

Intercept interviews allowed ANQAR surveys to capture the perceptions of Afghans from across all districts of Afghanistan. However, as one might expect, Afghans travelling outside their districts who are the likely candidates for intercept interviews may represent a demographically skewed sample. Analysis of the results from intercept interviews compared to data from the rest of the nation shows unsurprising trends: intercept interview respondents tend to be employed, older, and male as one would expect of someone travelling between districts, and these individuals tend to have poorer perceptions of security, and of the police, and they are more likely than the rest of the population to say that insurgents have grown stronger [6].

Intercept interviews therefore represent a compromise: the demographic representativeness of the survey sample is sacrificed in favour of better geographic representativeness. However, when folded into the remaining nation-wide data collected by the usual means, intercept interviews do not significantly affect nation-wide results (a shift of between one and three percent overall even when perceptions from inaccessible districts are very different from the rest of Afghanistan).

District accessibility for surveys is determined solely by the company contracted to conduct the surveys, i.e., D3 Systems, Inc. in the case of ANQAR. These company assessments are therefore independent measures of security which can be used as a proxy metric for security, to complement other security assessments performed by NATO.

9.3 WHAT THE SURVEY DATA SAYS ABOUT NATION-WIDE ATTITUDES: DESCRIPTIVE STATISTICS

The vast majority of useful survey analysis still comes from generating descriptive statistics from the data. Calculations of population proportions (i.e., "X% of the population says Y") and reporting on how they change over time and how they vary across the battlespace satisfies most headquarters information requirements. This is true for household surveys in general, not just for surveys in support of complex operations [7].

Descriptive statistics of data collected by means of a random sample of the population are themselves are useful and important. They are not merely a summary of what certain respondents told survey teams. Because those respondents were randomly selected from the nation-wide population, and because population-proportional weighting are applied in the calculation, the results reflect national attitudes.[3] This allows one to make the leap from saying "X% of respondents said Y" to "X% of Afghans would say Y". It also allows one to attribute margins of error to the results; we know exactly how many respondents gave a particular answer, but we can only estimate what percentage of the entire population feels this way. This leap – from describing the collected data to inferring about population proportions – is a fundamental aspect of population surveys, and one that falls directly out of "simple" descriptive statistics.

[2] Previous to that, intercept interviews were used on a limited basis in select districts where ISAF insisted on survey penetration even though districts were inaccessible. This was done very sparingly pre-2014.

[3] Just like drawing a sample of colored marbles from a jar allows one to make inferences regarding the distribution of marble colors within of the jar.

9.4 GETTING MORE OUT OF SURVEY DATA: BEYOND DESCRIPTIVE STATISTICS

Analysts conducting survey analysis in support of military operations often do not get a chance to go beyond calculating and reporting on descriptive statistics. Given the time constraints related to reporting survey results as soon as fieldwork is complete (often 7–10 days from when the data are received by analysts to when initial reports are due), the frequency of the surveys (new data every three months), and the limited resources available (one or two operations analysts dedicated to survey analysis), more detailed inferential statistics are typically done in the margins and/or as part of longer-term studies. Detailed results of statistical inference (e.g., p-values from hypothesis tests) are often not explicitly included in the main survey products provided to commanders.

This section gives examples of techniques employed within ANQAR which go beyond simple descriptive statistics. Some of these are standard analytical techniques which can be found elsewhere in the literature (e.g., Ref. [8]) while others are novel and developed specifically to address Afghan survey data.

9.4.1 Pair-Wise (Bivariate) Correlations

In addition to simple descriptive statistics, analysis of the correlation between pairs of variables is routinely done when reporting on survey results to commanders. Simple correlations between opinions and demographics are most often used. For example, examining the differences in perceptions of the Taliban between Pashtun, Tajik, and Hazara ethnicities.

In Afghanistan, there are several demographic variables along which public opinions are most often divided. Perceptions tend to differ most between ethnic groups; and between residents of Kabul City, other urban areas, and rural areas. Opinions tend to also be split by age, gender, and geographic regions. Other factors such as literacy, household income, employment status, and occupation can have a high degree of correlation with certain questions. Pair-wise correlations are typically done using simple weighted contingency tables or correlation coefficients.

An understanding of the operational context should guide analysts in anticipating which variables might be correlated. (e.g., Pashtuns should favour a Pashtun president, Kabul residents may be more cynical of central government corruption, etc.).

The risk in examining correlations in a pair-wise fashion is that there may be other variables that explain the correlation. For example, the reason that Pashtuns may have a worse perception of security is a result of the fact that Pashtuns tend to live in the South and Southeast of Afghanistan along the border with Pakistan (the so-called “Pashtun belt”) which is also where the insurgency is most fiercely contested. Thus, the correlation between ethnicity and perceptions of security is explained by a third variable: geography. To disentangle multi-variate correlations, we can (though infrequently do) turn to logistic regression models.

9.4.2 Logistic Regression Modelling

In order to disentangle the effects of more than one explanatory variable on a variable of interest, logistic regression and generalized linear modelling techniques are appropriate. For categorical or ordinal items (e.g., questions with two or more response options), logistic regression models are appropriate. Most statistical packages have the functionality to build such models, and ANQAR analysts have more recently tended to use R statistics software for this purpose [9].[4]

[4] R is a free software environment for statistical computing and graphics.

For example, a logistic regression can give us more insights about Afghans' willingness to join the Afghan National Army (ANA). Figure 9-3 shows the explanatory variables used in such an analysis as well as the regression results. The Log Odds Ratio can be interpreted as, for example, Kabul residents are ($10^{0.27}$ =) 1.9 times more likely than the average Afghan to be willing to enlist in the ANA (when all other factors are taken into account), whereas those living in the 215th Corps Area of Responsibility (i.e., Helmand and Nimroz provinces in Afghanistan's southwest) are 3.9 times less likely than the average Afghan to be willing to enlist (again, with all else equal). Error bars allow us to discern variables which have no bearing on recruitment intention (even if bivariate correlations suggest they might). For example, when all other factors are included, ethnic Hazaras are neither more nor less likely to say they would be willing to enlist.

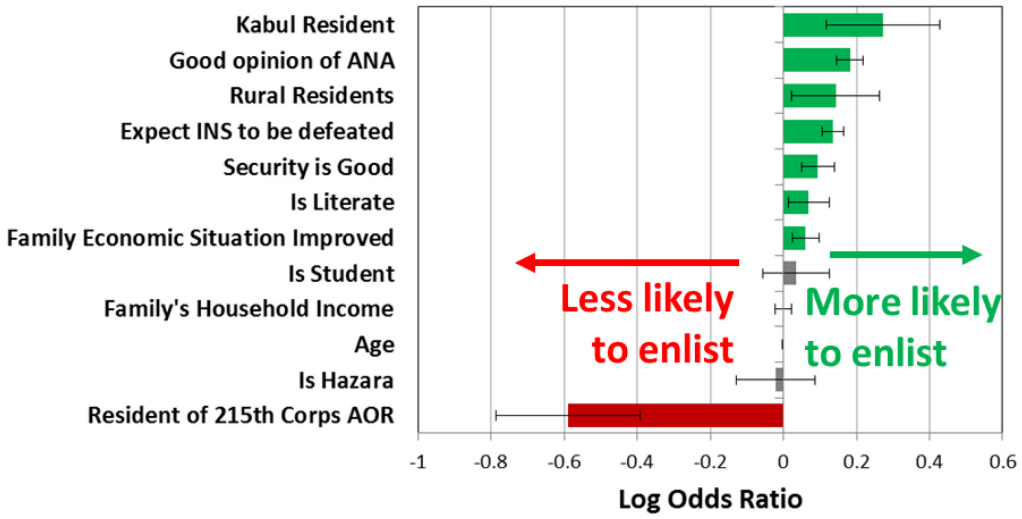


Figure 9-3: Logistic Regression Model for Willingness to Consider Enlisting with the Afghan National Army.

Care should be taken when trying to include too many variables in a model. The more the data are split across variables, the smaller the sample size for each possible combination of variables, and therefore the higher the margins of error in the resulting correlation coefficients/regression parameters. If done properly, this should be reflected in the size of the error bars such as those shown in Figure 9-3.

Furthermore, care should be taken in using automated techniques for determining which dependent variables should be included in the model (e.g., automated step-wise regression). Such techniques tend to identify the many obvious and uninteresting correlations within the data (e.g., Pashtuns live in the South), and operationally interesting correlations can be overwhelmed by the trivial ones. Knowing what to look for before building a model is often key. This requires analysts to have an intimate understanding of the operating environment, or a close link to operators who can guide such analysis.

9.4.3 Margins of Error and Hypothesis Testing

Analysts must have a firm understanding of survey margins of error in order to properly interpret results. Margins of error communicate the likely range in which the true population average falls, based on the value measured from a sample of the population. While simple margins of error are often reported, they do not reflect the true error which is higher due to the survey design. Specifically, the stratified multi-stage clustered sampling strategy often employed in complex environments results in additional contributions to the margin of error (for more detailed discussion, see Ref. [10]).

"Complex Margins of Error" (CMoEs) are more reflective of the effects on statistical error of a complex survey design and should be used when reporting survey errors. These effects are rolled up into a single parameter called the "design effect" (deff). This reflects the increase in margin of error resulting from a sampling method that deviates from a Simple Random Sample (SRS). A conservative complex margin of error at the 95% confidence level is given by:

$$\text{CMOE} = \pm 0.98\sqrt{\frac{\text{deff}}{\text{N}}} \tag{9-1}$$

where N is the sample size. The design effect can be calculated using the R Statistics Survey Package [11].

Design effect estimates are currently provided by the survey contract company, D3 Systems, Inc., as part of their deliverables for each survey. An average over a standard set of questions is used to determine an overall survey design effect which is then used for all questions. For ANQAR, the value of the design effect is typically around two for the nation-wide sample (i.e., deff = 2). Thus the CMoE is about 40% higher than the simple margin of error for a simple random sample with the same sample size.

Incorporation of design effect into hypothesis tests is done automatically within environments such as R Statistics when used with the survey package. Exact tests have also been developed take into account design effect, and which are valid for small sample sizes, e.g., for comparing district results [12].

9.4.4 Novel Ways to Amalgamate Results Along Non-Traditional Geographic Boundaries: Natural Opinion Boundaries in Afghanistan

Even with a large nation-wide sample size (e.g., N = 12,500 for ANQAR), at the lowest level of granularity (e.g., districts) the number of interviews can be small – e.g., as few as eight to 64 interviews per district in ANQAR – and therefore the margins of error at the district level can be quite large (±15% to ±50% after taking into account design effects). As a result, analysts tend not to report district-level survey results, instead opting to calculate province-wide or regional aggregates when examining sub-national trends.

However, provincial or regional administrative boundaries tend not to reflect natural population boundaries, for example ethnic, cultural, or economic boundaries, nor do they reflect operational boundaries for counterinsurgency operations or for insurgent groups. Province or regional aggregates therefore tend to "wash out" any local changes in perceptions resulting from operational effects, or due to ethic/demographic divisions.

To overcome this problem, NATO Communications and Information Agency analysts recently developed and implemented a technique to aggregate districts in a way that is not reliant on provincial boundaries [13]. The technique uses a hierarchical clustering algorithm and a Bayesian statistical test to identify clusters of districts where public opinions are the same, but which are statistically different from adjacent areas. The resulting district clusters define regions in which opinions are not statistically distinguishable. Distinct regions are separated by natural "opinion boundaries" (see Figure 9-4). Some of these boundaries align nicely with existing ethnic and operational boundaries in Afghanistan, lending confidence to the approach.

The fact that there are large geographic areas where public perceptions are statistically indistinguishable suggests that the current survey sampling method is not overly sensitive to exactly which primary sampling points (villages) are included in the random sample. This gives increased confidence in survey methods in general.

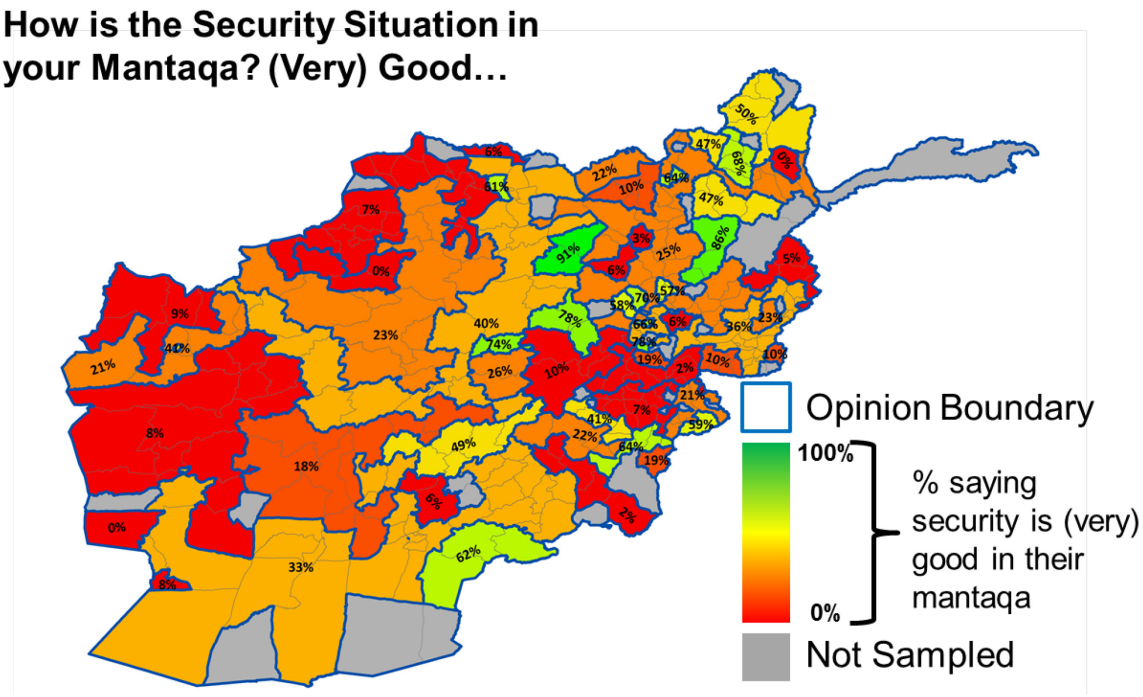


Figure 9-4: Natural Opinion Boundaries for Perceptions of Security.

9.4.5 Combining Survey Datasets for Demographic Mapping

Since the start of ANQAR in September 2008, through the projected end of the programme at the end of 2016, a total of approximately 384,000 interviews will have been conducted in households across Afghanistan.[5]

A by-product of the consistent collection of such large quantities of survey data is that the cumulative dataset paints a detailed demographic picture of the country. The one caveat to this approach is that combining multiple surveys collected over the span of eight years washes out any changes in demographics over that time, resulting in a time-average picture. However, certain demographic information, such as ethnicity, is unlikely to vary significantly over time, so that the combined data should give a relatively accurate picture.

Figure 9-5 and Figure 9-6 depict the results of such an effort which gave ISAF the most detailed picture of ethnic distribution to date (including ethnic percentages per district). Cross-validation against other sources of ethnic distributions at the national level (CIA World Factbook 2004–2010, The Asia Foundation Survey 2012, ABC/BBC/ARD Surveys 2004–2009) aligned remarkably well, giving additional confidence in the approach.

This ethnic mapping was completed just before the 2014 Afghan presidential elections. Additionally, an ethnic violence metric was developed which gave an ethnic breakdown for the average levels of violence seen by individuals in their district. The results showed that before the elections violence increased more rapidly among some groups than among others. These results had implications on the freeness, fairness, and transparency of an election in which candidates and voters were split along ethnic lines.

[5] Given a population of around 30 million and given that the number of people living in each household is 9.5 (based on amalgamated ANQAR results from Waves 1–30 [14]), this suggests that nearly one in eight Afghan households have been surveyed at one time or another in the past eight years. This possibility ignores the likelihood of some households being sampled more than once in the past eight years though even if the same household was selected, there is still a good chance that a different individual was selected given the large average size of households.

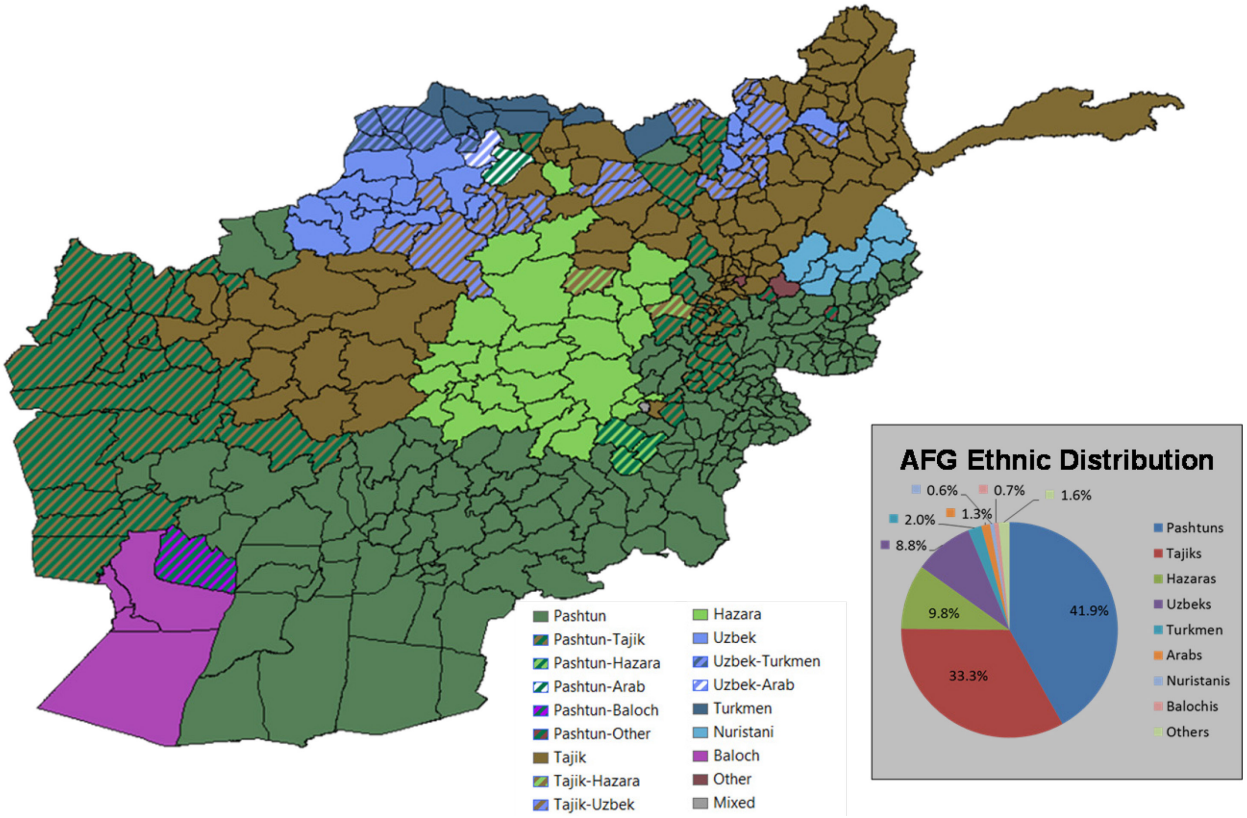


Figure 9-5: Survey-Based Ethnic Map.

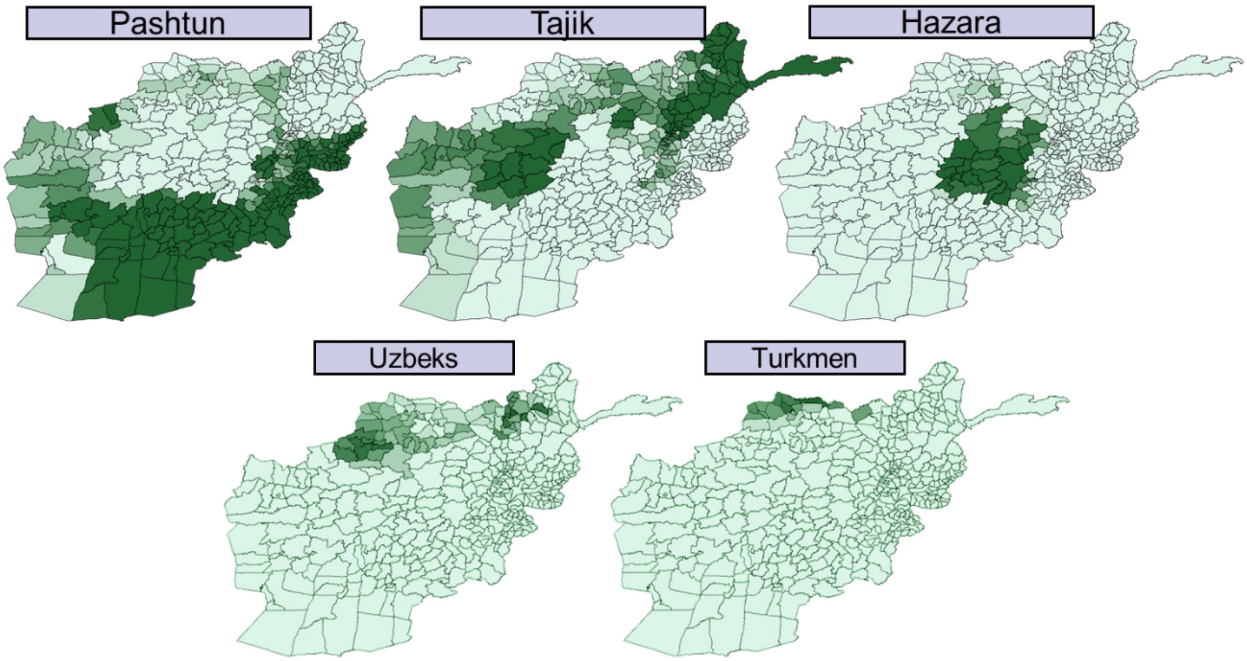


Figure 9-6: Ethnic Proportions by District.

9.5 COMPLEMENTARY DATA COLLECTION EFFORTS – BEYOND SURVEYS

Analysts should consider other parallel data collection efforts to complement survey data. In some cases, this can be done within the envelope of the existing survey contract with the data provider (e.g., sacrificing sample size for the nation-wide poll such that the additional effort remains cost-neutral). In other cases, separate contracts and funding sources may be required.

This section briefly touches on four complementary data collection efforts which have run in parallel with ANQAR.

9.5.1 Focus Groups and In-Depth Interviews

In September 2014, in parallel with the Wave 29 data collection effort, and under the ANQAR contract umbrella, D3 Systems, Inc. ran a limited scope trial of focus groups and in-depth interviews.

The intent of this qualitative data collection effort was to help explain recent trends in the quantitative survey data. For example, while ANQAR was reporting a deterioration in satisfaction across a range of issues related to national government, the economy and security, few details were available regarding the root cause of this growing dissatisfaction.

Interviews were conducted in the form of focus groups and in-depth interviews in five provinces. Respondents were asked questions on various topics including security, crime, reintegration, and the presidential election. A total of ten focus groups were conducted: two focus groups in each province with eight participants each, and a total of 40 in-depth interviews divided into eight interviews in each province. Half of the respondents were male and half were female between the ages of 18 and 36 [14].

The results provided an interesting additional layer of context to the quantitative ANQAR survey data, but in the end, fell short in the attempt to explain the sources of public perceptions. Several reasons for the shortfall were identified:

- **Opinions lacked in-depth insights**: The responses to the questions and the discussion during focus group lacked depth and insight. Whether because of the demographics of the participants (i.e., average working class Afghans, not necessarily literate or informed on national issues) or because individuals had not developed well-formed opinions on complex topics, the responses recorded from the interview did not provide very many additional insights into issues of interest. See for example, an extract of results in Figure 9-7.

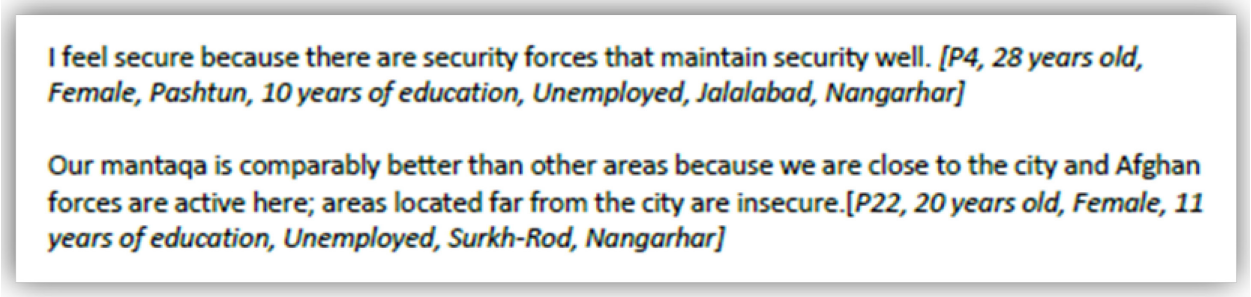


Figure 9-7: Sample of Results from Ref. [14].

- **Respondent deference to interviewers**: The results suggested that respondents often deferred to interviewers. For example, when asked "is there anything else you would like to add?", some

respondents made remarks to the effect that surely the interviewers had thought of everything, and that they had nothing more which they could add. Some respondents may have felt intimidated by the literate, educated, "big city" interviewers, or at minimum, they may not have been accustomed to expressing their opinions openly in such a forum. One 29 year-old male shopkeeper in Helmand responded, "we are illiterate, you know better" when asked if he had anything else to contribute.

- **Pre-scripted interview questions did not offer the ability to drill down**: Because questionnaires were pre-scripted and delivered by a trained Afghan interviewer, there was limited opportunity to delve deeper into subjects based on a particular response. Such deep dives would have to be anticipated in advance and scripted into the questionnaire in the form of probing or follow-up questions. In response to some follow-up questions such as "why", respondents often reiterated the question: "security deteriorated because there are more attacks in our area".
- **Too many topics were covered**: As a corollary to the point above, the initial trial attempted to cover too many topics, and failed to focus on a few key issues. This limited the planning put into potential responses and follow-up questions which may (or may not) have yielded more substantive answers.
- **The volume of qualitative data produced from transcripts was too large to be efficiently analysed by in-theatre staff**: The sheer number of pages of transcripts of all interviews and focus groups (747 pages, over 200,000 words) was a challenge to consume and integrate with the resources and time allocated. To read all the transcripts was impractical, let alone to pull together themes and key issues (e.g., broken down by areas, demographics, etc.).
- **Not enough resources (and possibly data) to effectively implement standard qualitative data analysis techniques**: There are specific software tools (e.g., NVivo) which allow researchers to code and analyse unstructured text. However, such techniques are time consuming and under the current contract, such analysis was not delivered by the survey company. It was also suggested that the size of the dataset was not quite large enough to infer anything useful from this approach. By the time researchers broke down comments by demographics and regions, they had a very small sample in each group. That being said, the D3 Systems' analysis report did provide high-level results which were of interest.
- **Qualitative Opinion Research requires a different skill set than typically resides in theatre Assessment Cells**: Deployed quantitative analysts (e.g., NATO Operational Analysts or U.S. ORSAs) often do not have the training and background in qualitative research to properly conduct data analysis. In the case of ANQAR, D3 Systems, Inc. was able to provide a wealth of knowledge and guidance to Operational Analysts to support the questionnaire design and analysis of the qualitative data. However, the ability to do any additional in-house analysis, including the availability of qualitative research software tools, was lacking within the deployed HQ. Neither the funding nor the contracting vehicles were in place to properly enlist the help of experts in the field[6] (see also Minkov's article on "Red Teaming" in this volume).
- **Challenges in extrapolating results to make nation-wide inferences**: The fact that one (or a few) Afghans offered an anecdote that may have given insight into certain issues (e.g., their experience with police checkpoints), was difficult to interpret and include in a strategic report. It was impossible to know whether these comments were representative, or merely reflected the experience of one individual. It was, however, useful to include interesting quotes when presenting the quantitative survey data. This gave ANQAR survey results a more human feel that pie charts and trend lines lacked on their own. However, analysts should be aware that additional content accompanying the survey results (e.g., quotes, pictures, etc.) can cause cognitive biases which affect how a reader interprets the information. This is known as a "framing effect" [15].

[6] The insights expressed here reflect those gained from extended discussions with D3 Systems regarding qualitative survey (staff deployed to Kabul) who agreed that similar issues affect other qualitative research in conflict zones.

Some of the issues mentioned above may be remedied with a better implementation. Future analysts are encouraged to pursue similar efforts while keeping in mind those lessons identified above. In the case of ANQAR, it was decided by the survey sponsor at ISAF that while this trial had been successful in demonstrating the ability to conduct qualitative research, any further work would be put on hold.

9.5.2 Collection of Complementary Atmospherics Data by Survey Teams

In February 2015, D3 Systems, Inc. ran a pilot atmospherics study in parallel with ANQAR data collection, leveraging the nation-wide field force which they already employed for conducting interviews.[7] For the pilot study, in addition to collecting public perception data, field teams and regional supervisors completed questionnaires regarding atmospherics within district centres. After visiting district centres, field teams completed a questionnaire on issues such as: presence of insurgents, availability of goods and services, infrastructure, telecommunications, and travel precautions required.

These data were intended to complement public opinion data by giving an in-depth look at what the reality is on the ground for the people that live in these areas.

Most importantly, atmospherics data were also collected in districts otherwise inaccessible to survey teams due to insecurity. These districts are in fact not completely inaccessible to Afghan interviewers, but rather security prevents them from overtly asking residents about their opinions on potentially sensitive topics. However, local Afghan survey company employees are still able to travel to the district centre to discretely and unobtrusively observe conditions in the district.

Complementary data collection, such as this, requires additional resources from the survey company, and while the information obtained was deemed to be highly valuable, particularly in complementing public opinion data, additional funding for ongoing data collection was not available at the time and the atmospherics data collection was not continued beyond the initial trial.

In the future, analysts may consider this type of data collection when developing a data collection plan in support of assessments. While in principle, such data collection could be done independently of other survey data collection, the fact is that survey companies already have on staff teams of trained local people who travel across the country to collect opinion polling data. Harnessing that capability to collect additional data represents an incremental effort and cost, and is therefore likely to be easier and cheaper than setting up a separate data collection effort.

9.5.3 District Accessibility as a Proxy Measure of Security

One by-product of managing teams of interviewers who travel the country to collect public perception data is that survey companies must maintain their own independent assessments of local security in order to determine whether it is safe to send male and/or female interview teams to certain areas. This dataset is in itself potentially valuable to assessment staff. It represents an impartial and unbiased assessment of local security. In a way, this is a more direct measure than, for example, other measures such as number of security incidents in a district.[8]

One challenge is whether the survey company retains a consistent risk averseness – i.e., that its willingness to send teams into marginally secure areas remains roughly unchanged. There may be, for example, a

[7] An "atmospherics" program refers to the systematic collection and reporting of information about the environment by individuals sent into the field. This, includes collection on indicators related to economics, infrastructure, and governance (e.g., is the bazaar open? are there insurgent checkpoints?). This information is complementary to other efforts to understand the environment including Human Terrain Teams and population surveys [16].

[8] Counting security incidents as a proxy for security has certain flaws, particularly when it comes to areas under uncontested insurgent control where usually security incidents are low.

tendency for districts which were previously deemed inaccessible to remain closed even if conditions change for the better. It is easier and cheaper for a company to minimise the number of locations to which they have to send their teams. However, survey companies also recognize that maximum coverage is critically important to the sponsors, and that "nation-wide" results excluding the opinions of the residents in the most-contested areas will not continue to have the support of the sponsors.

In fact, reliance on intercept interviews for inaccessible districts (discussed above) means that interview teams must still travel to the vicinity of the insecure district, with an additional burden of "intercepting" and interviewing individuals travelling to/from the inaccessible district. Therefore, there is in fact limited savings seen by the survey company in keeping certain districts closed once security has improved.

9.5.4 Collecting Opinions from Target Groups – Surveying Afghan National Army Soldiers

While public perception data continued to be important after the NATO mission transition from ISAF to Resolute Support (RS), it did not inform many of the tactical issues related to a security force assistance mission. In particular, Resolute Support's focus on capacity building within Afghan security institutions also included supporting recruitment and professionalization of Afghan soldiers and police forces. Issues related to recruitment, retention, and morale within the Afghan National Army (ANA), and more generally, the experiences of Afghan soldiers and recruits were much more directly related to the objectives of Resolute Support. In late 2015, the Afghan Assessment Group (AAG) at RS Headquarters, working in collaboration with other RS units and with technical support through a new contract with D3 Systems, Inc., implemented a series of surveys of Afghan soldiers and recruits.

The Afghan National Army Nation-wide Atmospherics Study (ANANAS – or Pineapple as it became known)[9] represented a series of small-scale data collection efforts for which the target population was Afghan Army soldiers and recruits. The data collection effort revolved around two key issues: recruitment and retention of ANA personnel.

In a pilot study conducted in late 2015 and early 2016, a paper-based survey was administered on three separate occasions to recruits at the ANA Recruiting Command in Kabul. A total of nearly 400 responses were collected. Recruits were asked questions about their experience with recruitment centres, their expectations, and experience upon intake. Initial results provided some valuable insights into the experience of new recruits and informed recommendations on how to improve the recruitment process [17].

Following the pilot study which demonstrated the utility of surveying soldiers, a subsequent survey used electronic data collection with tablet computers. The tablets ran software which read aloud each question and the possible choices of answers in the local language (Pashto or Dari).[10]

Tablet-based data collection mitigated the difficulties of surveying a largely illiterate population by allowing respondents to follow audio and visual cues. The tablets were also intended to give respondents an increased confidence of anonymity compared to in-person interviews. The issue of anonymity was particularly important when asking about possibly sensitive topics related to soldiers' perceptions of their professional experiences. The intent was to alleviate respondents' concerns about the career implications of a negative response. The degree to which soldiers trusted the anonymity of their inputs into an electronic device is an area for further study.

Since the tablets upload their findings to a secure server, they provide the additional benefit of data being instantly available for analysis. This eliminates the need to have paper surveys translated and decreases the

[9] Ananas is the word for pineapple in a number of languages other than English.

[10] Technical support to the implementation of this data collection effort, including the tablets, software, translation and recording of questions/answers, and data collection/consolidation/reporting, were all conducted under a contract with D3 Systems, Inc.

turnaround time for developing analytical products. As of March 2016, 200 ANA recruits were being surveyed monthly with the tablets.

The "Pineapple" study also evolved to surveying field units about the adequacy of training and equipment, quality of life, morale, satisfaction with career progression opportunities, leave policies, and experience with corruption. As of March 2016, a paper-based pilot survey has been fielded with 207th ANA Corps and tablet-based surveys are being prepared [17].

It is still too early to assess the long-term success of the program, although initial indications from the recruit surveys are promising. One potential issue may be the difficulty to conduct a random sample from field units, both in terms of ways to randomize respondent selection, and in terms of understanding the target population demographics to be able to make inferences about the whole population from the responses of a few (i.e., to stratify and weight results by rank, gender, age, years of service, etc.).

Nevertheless, the targeted data collection effort is noteworthy in its novel medium (using tablets) approach, as well as in its focus on a population group of interest to the mission. In the future, analysts will be able to learn from this initial experience, and at minimum may be encouraged to look beyond traditional collection of perception data from the public at large, towards a more targeted approach.

9.6 CONCLUSION

The subjects discussed in this chapter represent practical issues faced in overseeing survey programs in support of complex operations in conflict environments. While these are not an exhaustive set, they are representative of the challenges that deployed operational analysts can expect to face.

As with other data collected in an operational theatre, surveys can inform commanders and their staffs; they can provide situational awareness, and inform plans and decisions. However, as with other operational data, survey data are collected in an environment that is inherently insecure, often in underdeveloped regions of the world. In turning operational data into products to inform decision makers, analysts must be aware of the limitations associated with these data. Similarly, with survey data, analysts must have a firm grasp of both technical and administrative aspects of surveys as well as an intimate understanding of the mission, the operational environment, and the needs of the survey sponsors. These will help analysts to better respond to the needs of the commander.

As the international engagement in Afghanistan continues to evolve, and as the cumulative experience on surveys continues to build, it becomes increasingly important to capture lessons learned from survey programs supporting such ongoing operations and to further develop and deepen the corpus of training material which will assist analysts who will be preparing for the challenges of overseeing surveys in support of future missions.

9.7 AUTHOR'S BIOGRAPHY

Philip Eles, PhD, is an operational analyst at the NATO Communications and Information (NCI) Agency in The Hague, Netherlands, where he has been providing analytical support for strategic assessments of Afghanistan since 2013 through projects for SHAPE, JFC Brunssum, and ISAF/Resolute Support. He has deployed on five occasions to the Afghan Assessment Group (AAG) in Kabul, and is part of NCI Agency's Reach-Back Support Team for the AAG. Since 2014, he has been the Technical Lead on ANQAR, NATO's survey program in Afghanistan. Between 2006 and 2013, as a Defence Scientist at Defence Research and Development Canada, Dr. Eles worked on assessments to support to the Canadian combat mission in Kandahar. He was one of the lead analysts for Canada's Kandahar Public Opinion Polling (KPOP) program from 2007 to 2011. Dr. Eles holds a Ph.D. in Physics from the University of British Columbia in Vancouver.

9.8 REFERENCES

[1] Warshaw, M., Kakar, R., Habibzei, T. and Mohsini, Z. 2006. Starting from Scratch – Building Social Research Capacity in Afghanistan. ESOMAR Congress 2006, London, September.

[2] Vincent, E., Eles, P. and Vasiliev, B. 2009. Opinion Polling in Support of Counterinsurgency. The Cornwallis Group XIV: Analysis of Societal Conflict and Counter Insurgency, 14, April.

[3] NATO. 2015. NATO Operations Assessment Handbook version 3.0. Annex E: Polling to Support Operations Assessment. Norfolk, VA: Supreme Allied Commander Transformation. July 2015. pp. E1-E10.

[4] United States Army. 2015. Deployed Analysts Handbook, Chapter V: Public Perception Surveys: Development and Analysis. U.S. Army, Centre for Army Analysis (CAA).

[5] D3 Systems, Inc. 2016. Afghanistan Nationwide Quarterly Assessment Research (ANQAR) Methods Report. Wave 31. March 2016.

[6] D3 Systems, Inc. 2015. Afghanistan Nationwide Quarterly Assessment Research (ANQAR) Final Analytical Report. Wave 27.

[7] Chromy, J.R. and Abeyasekera, S. 2005. Statistical Analysis of Survey Data. Household Surveys in Developing Countries: Design, Implementation and Analysis. ST/ESA/STAT/SER.F/96. UN Department of Economic and Social Affairs, Statistics Division, 2005. http://unstats.un.org/unsd/hhsurveys/pdf/Household_surveys.pdf.

[8] Fowler, F.J. 2013. Survey Research Methods. Sage Publications, 2013.

[9] R-Project. 2015. The R Project for Statistical Computing. https://www.r-project.org/.

[10] Eles, P.T., Vincent, E., Vasiliev, B. and Banko, K. 2012. Opinion Polling in Support of the Canadian Mission in Kandahar: Final Report for the Kandahar Province Opinion Polling Program, Including Program Overview, Lessons, and Recommendations. Technical Report DRDC CORA TR 2012-160U. Defence R&D Canada Centre for Operational Research and Analysis, September 2012.

[11] R-Survey, 2015. Package 'Survey'. https://cran.r-roject.org/web/packages/survey/survey.pdf [Accessed Date: February 20, 2015].

[12] Eles, P.T., Vincent, E. and Kalantzis, E. 2009. A Bayesian Approach to Hypothesis Testing for the Kandahar Province Opinion Polls. Technical Memorandum DRDC CORA TM 2009-066. Defence R&D Canada Centre for Operational Research and Analysis, December 2009.

[13] Eles, P.T. 2016. Natural Opinion Boundaries: A Novel Technique for Geographic Clustering of Public Opinion Polling Data. Client Report TR2016BRS010364. NATO Communications and Information Agency.

[14] D3 Systems, Inc. 2014. Afghanistan Nationwide Quarterly Assessment Research (ANQAR) Qualitative Summary. Wave 25. Oct 2014.

[15] Powell, T.E., Boomgaarden, H.G., De Swert, K. and de Vreese, C.H. 2015. A Clearer Picture: The Contribution of Visuals and Text to Framing Effects. Journal of Communication, 65(6):997-1017.

[16] National Research Council. 2011. Sociocultural Data to Accomplish Department of Defense Missions: Toward a Unified Social Framework, Workshop Summary. Robert Pool, Rapporteur. Planning Committee on Unifying Social Frameworks. Board on Human-Systems Integration, Division of Behavioral and Social Sciences and Education. Washington, D.C.: The National Academies Press.

[17] Hill, E. 2016. Resolute Support Headquarters, Afghan Assessment Group. Personal communication with author.

CHAPTER 10 – DESIGNING AND ASSESSING COMMAND AND CONTROL TO DEAL WITH COMPLEX AND ILL-STRUCTURED OPERATIONAL ENVIRONMENTS

Mark E. Tillman and Kathleen M. Conley
Institute for Defense Analyses (IDA)
UNITED STATES

ABSTRACT

An operation is, in its basic sense, a coordinated, collective response to a situation that an organization chooses to alter. Fundamental to the success of this response is the ability to control the organization's own forces while also influencing important external actors in order to shape the operation, either according to a plan or as the situation develops over time. While many factors can influence success, complex operations cannot succeed without effective Command and Control (C2). It is C2 that determines what the organization will do and develops operational plans to influence the situation. Operations assessments, a critical component of C2, enable measurement of progress toward objectives and thereby contribute to decisions that adjust the plan as operations unfold. Since C2 is critical to success, it should also be assessed as the operation progresses. Decision makers should know how to match their C2 Approach and methods to the circumstances at hand. This chapter considers what operators can do, not only to assess their C2 Approach and method as originally designed, but also to assess and modify the C2 Approach throughout an operation. This is called "C2 by Design" [1]. This pragmatic interpretation of C2 Agility theory as explored in SAS 085[1] extends the theory's key ideas to operations assessment. The goal is to enable staffs and commanders to focus and act appropriately to alter their approach to C2 "in the moment" in order to preserve, recover, or improve organizational effectiveness. Successfully altering situations in complex and ill-structured operational environments will require a C2 design that is attuned to circumstances and that can both address current demands and shift to meet new challenges.

10.1 INTRODUCTION

A growing consensus among NATO militaries recognizes that operational problems are becoming more complex (see Ref. [2] p. ii). The planning community has begun to approach such problems using a method known as *design*. Design evolved to address a class of problems that are much more difficult to address than complex problems, ill-structured problems. With this class of problem, the first challenge is developing sufficient understanding of the problem to begin charting an appropriate response (an operational approach), given the dynamic environment (see Ref. [3] p. v). For ill-structured problems in particular, the design method guides the formulation of an operational approach that includes:

1) Probing to learn more about the context of the problem and helping planners better describe the system as the operation unfolds; while also

[1] Agility is the capability to successfully effect, cope with, and/or exploit changes in circumstances. While other factors will also influence outcomes, C2 Agility enables entities to effectively and efficiently employ the resources they have in a timely manner in a variety of missions and circumstances. SAS-085 was formed to improve the understanding of C2 Agility and assess its importance to NATO. SAS-085 accomplished these objectives by articulating the principles of C2 Agility, in the form of a C2 Agility Conceptual Model, substantially validating this model and establishing the importance of improving C2 Agility with empirical evidence obtained from a set of retrospective case studies and simulation-based experiments. Further, it identified next steps toward practical implementation in NATO operations and priorities for increasing the rigor and practicality of methods for measuring and improving C2 Agility (see Ref. [2]).

2) Learning whether that system is responding favourably or otherwise to operational injects into the system.

Operations assessments help commanders and staffs understand the changing state of such systems and whether the operational approach is contributing effectively to the desired next state – an improved situation – and ultimately to success.

While many factors contribute to success, no operation can succeed without effective Command and Control (C2). Every operation has an approach to C2, either 1) defaulting to the pre-existing C2 system, along with ad hoc arrangements that surface as needed, or 2) as purposely planned, possibly through design [4]. A C2 Approach is understood in the context of myriad relationships among the organizational actors – also called *entities* – involved in the operation, some of whom are external to the organization of interest. Each C2 Approach has three components:

1) Patterns of interaction or collaboration;
2) Distribution of information between the entities; and
3) Distribution of decision rights for action.

Changing a C2 Approach involves modifying the relationships (adding or dropping entities) and tuning each relationship as needed for the situation, either current or anticipated. Therefore, theoretically, an infinite number of potential C2 Approaches is available – some are better than others, given the operational circumstances.

The effectiveness and appropriateness of a given C2 Approach to either the current or anticipated circumstances should be assessed as an operation unfolds. A C2 Approach should complement the operational approach, matched to the circumstances at hand; if not, the C2 Approach should adapt as circumstances change just as one would expect the operational approach to adapt. Being able to recognize the need to change C2 and then change the C2 Approach to better align with the circumstances is known as C2 Agility.[2] While agility can be a highly desired attribute of C2, exercising C2 Agility in the moment requires a way to assess the current or anticipated C2 Approach as operations are designed or executed. This is a C2 Assessment.

10.2 WHY ASSESS THE C2 APPROACH?

Given the undeniable importance of C2, the C2 Approach should be examined not only as part of the design process but continually thereafter. If, however, a new C2 Approach was not established during an initial design process, an important issue to consider early on is whether the C2 Approach currently being practised is adequate for implementing the operational approach.[3] This may be the case when circumstances nearly matched those envisioned when the operation was originally designed. If an operation is quite routine, the existing C2 Approach may suffice. Even if existing C2 arrangements were assessed as more or less adequate, small adjustments to the C2 Approach may serve to better align it with the operational approach and better enable the Lines Of Effort (LOEs) and supporting functions envisioned for the operation. However, any

[2] C2 Agility is the capability to successfully effect, cope with, and/or exploit changes in circumstances. While other factors will also influence outcomes, C2 Agility enables entities to employ effectively and efficiently the resources they have in a timely manner in a variety of missions and circumstances. The SAS-085 panel was formed to improve the understanding of C2 Agility and assess its importance to NATO. SAS-085 accomplished these objectives by articulating the principles of C2 Agility in the form of a C2 Agility Conceptual Model, substantially validating this model and establishing the importance of improving C2 Agility with empirical evidence obtained from a set of retrospective case studies and simulation-based experiments. Further, SAS-085 identified the next steps toward practical implementation in NATO operations and set the priorities for increasing the rigor and practicality of methods for measuring and improving C2 Agility [2].

[3] If the initial design process included design of a C2 Approach, then the C2 design should already be matched to the initial operational approach and circumstances.

significant change in circumstances, such as a different set of coalition partners, rules of engagement, or command relationships, will likely drive a need to modify the initial C2 Approach.

Before adopting a new C2 Approach, the newly-selected approach should be deemed feasible from both technical and process perspectives.[4] Finding, implementing, monitoring, and adjusting a new C2 Approach is an iterative process that should continue throughout planning and execution of the operation. Other important considerations in the selection of a C2 Approach are the new linkages[5] and modifications to existing linkages as well as adjustments to staff C2 activities[6] required to implement it. C2 Approach assessments begin with asking three basic questions:

1) Are all existing linkages and staff C2 activities (those associated with the old approach) still needed? Can some be eliminated?
2) Should some activities be modified?
3) Should any of the staff's C2 activities be included?

Examples of new linkages include new subordinate organizations, new partners not subordinate but working laterally, or external organizations and actors crucial to success. Staff activities include all the battle rhythm events such as boards, cells, working groups, and command briefings.[7] Such activities should be periodically re-examined to ensure their continued relevance so that staff time is directed toward producing desired outcomes rather than being consumed with low value-added work.

The Joint Operation Planning Process (JOPP), as described in *Joint Operations Planning* [5], includes a two-page overview[8] of the assessment process and its importance to success. The overview states that "assessment and learning enable incremental improvements to the commander's operational approach and the campaign or contingency plan", and that "assessments by joint force commanders allow them to maintain accurate situational understanding and revise their visualization or operational approach appropriately". Finally, "assessment precedes and guides every activity within the JOPP and concludes each operation or phase of an operation" [5] pp. III-44-45. Operational assessment of C2 empowers commanders to fine tune the alignment of their operational approach and their C2 Approach over time, contributing to sound decision making as an operation unfolds.

Accepting the critical importance of assessment, we describe how to assess the C2 Approach that is actually being practised (which may turn out to be different from the one described in commander's guidance). Normally, the C2 Assessment is closely tied to, and dependent upon, assessment of the operational approach, which, in turn, depends upon assessment of the operational design from which it was derived. These three assessments are conceptually separate activities; however, assessing the current C2 Approach may also reveal flaws in the operational approach that may require consequential revisions.

---

[4] For example, sharing information with partners, especially classified information, is nearly always a problem.

[5] Ref. [1] states that C2 linkages are all the human connections among organizational actors in a specific operation. Many may have existed prior to the operation, but normally new relationships or linkages will need to be established, both internal (subordinate) and external (lateral or upward, including supporting and supported organizations, host nation, allies and coalition partners, and international organizations). These can be military or non-military but in each case the modalities that govern the linkage must be confirmed, adjusted, or created to match the existing circumstances and enable the operational approach. Once these linkages have been established, attention needs to be paid to the necessary information exchanges.

[6] Ref. [1] p. 14 states that "C2 activities" include what Ref. [6] calls *activities*, which include planning, coordination, and analysis; what Ref. [7] refers to as *tasks*; and what other sources refer to as *C2 functions*. C2 activities include the full range of processes, tasks, and actions that may be taken to carry out the C2 function. The C2 activities become the means by which the C2 Approach is executed, just as operational activities are the means by which an operational approach is carried out.

[7] Ref. [4] pp. 20-21 contains a more exhaustive list of staff C2 activities.

[8] A ten-page appendix on assessment is also included in Ref. [7].

10.3 ASSESSING THE DESIGN

Design assessment should answer the question, "Is the design working?" Design assessment is essentially an assessment of the validity and relevance of the planning assumptions and the resulting understanding of current conditions, propensities, and expected changes in circumstances in response to stimuli (injects or actions). Design assessment also provides the information needed to adjust the situation in the desired direction through the actors' collective action along the Lines Of Operation (LOEs) included in the operational approach. In some cases, the assessing organization has direct control over other actors, and in other cases (e.g., when working with non-governmental organizations or coalition partners), it has more limited influence. The design effort aims to define the problem to be solved by developing a postulated understanding of the operational environment based on assumptions regarding missing yet critical information.

If the design is not working as initially forecast – which is frequently the case in a complex operation – the assessment should lead to changes in the design to reflect a better understanding of the interests and forces at work among the internal and external entities in the environment, including the enemy. The following sections presume the existence of an ongoing design assessment that informs and improves assessments of the operational approach and the C2 Approach.

10.4 ASSESSING THE OPERATIONAL APPROACH

The operational approach, derived from the design, is a visualization of actions along multiple lines of similarly purposed effort. It too is based on assumptions that must be validated by outcomes as the operation unfolds. The operational approach becomes, in this sense, a set of ideas that comprise the path to attainment of objectives, leading toward the desired next or end state. The ideas are translated into combinations of concrete actions sequenced over time that are grouped into LOEs. Assessing the operational approach involves answering two basic questions: "What are we doing? And is what we are doing working effectively?" The answers to these two questions will help inform the question, "What should we be doing differently to be more effective?"

First, what are we doing? Are events unfolding as envisioned? Or has something caused a deviation from the original design and operational approach?[9] If the operational approach has been clearly articulated and the LOE events are observable and monitored, this question can be answered. If events are not unfolding as envisioned, the result of the assessment could be an adjustment to a particular LOE or a complete 'reboot' of the operational approach to get back on track.

If the operational approach is unfolding as envisioned, the second question becomes "Is what we are doing working effectively?" For each LOE, and for the operation as a whole, each action or set of actions is expected to yield positive outcomes that are observable. If the approach is not working, it may be that more patience or greater effort will get it going. In all likelihood, however, examining the underlying assumptions may indicate that adjustments to the LOE or to the overall operational approach are required.[10]

As with design assessment, this chapter assumes an ongoing operational approach assessment that both informs and improves assessment of the C2 Approach. The C2 Approach can be assessed at two levels: a macro level that is more appropriate at an aggregated, holistic level where we find an operational approach; and a subsystem level that is more appropriate when disaggregating the operational approach or lower levels, perhaps specific LOEs or battlefield functions.

[9] For example, the Iraq War underwent a change in circumstances, from major combat to counterinsurgency operations, in 2003–2004.

[10] For example, Iraq in 2007. See Section 10.9 for background reading materials.

10.5 ASSESSING THE C2 APPROACH AT THE MACRO LEVEL

The C2 Approach should be tailored to the overall operational approach and its individual LOEs. A C2 Approach can be multi-dimensional and multi-level. Thus, the assessment needs to consider multiple levels: one of the operational approach itself (the macro assessment), another at the level of individual LOE, and another at the level of cross-cutting battlefield functions (such as fire support or intelligence) and activities affecting more than one LOE (subsystems).

The first step in assessing the C2 Approach at the macro level is to understand what it is supposed to be. This includes the desired relationships (i.e., the linkages) with other internal and external actors, the desired information flows among the actors, and the expected collaboration and decision rights that enable actions and activities overall and along each LOE.

The next step is to ascertain what is actually happening: "Is the C2 Approach, as implemented, what was intended?" Assuming the intended C2 Approach was clearly articulated, the elements of its instantiation should be observable. For example, if the C2 Approach requires a new linkage with an actor outside the NATO or coalition military structure, has the relationship with that actor been established? Is the desired information exchange occurring? If the approach required changes to linkages within the military structure (for example, because of a change in a supporting-to-supported relationship), have they been completed?

Macro assessment should also address the question "Is the C2 Approach working effectively?" Even if the C2 Approach has been implemented as initially envisioned in the commander's guidance, it may be incomplete, the expected outcomes may not have occurred, or the environment itself may change as the result of friendly actions, enemy actions, or the actions of neutral actors over whom the military commander has no control. Fixing problems with the C2 Approach requires re-examining the operational design, the operational approach, and the C2 Approach itself to determine the underlying rationale. This step, coupled with learning from ongoing operations, may require adjustments to the C2 Approach.

Assessment at the macro level is not simply a series of stoplight charts reporting the status of communications links. Rather, it is an assessment of whether the C2 Approach is aligned with, and supportive of, the operational approach. This requires a deeper look into what is happening in important established linkages, including those that do not exist electronically (such as periodic meetings or an exchange of information on paper) and whether what is happening makes sense with respect to the operational approach. The macro C2 Assessment then should be designed to answer both questions ("What are we doing?" and "Is what we are doing working effectively?") in terms relevant to the operational approach and not just traditional C2 metrics.

In addition to the formal assessment process, having a separate Red Teaming effort would be prudent of the commander. The Red Team would focus on the operational environment with an eye toward whether the operational approach and C2 Approach remain aligned with it. Table 10-1 illustrates key elements of both macro C2 Assessment and macro C2 Red Teaming.

10.6 ASSESSING THE C2 APPROACH AT THE SUBSYSTEM LEVEL

C2 at the subsystem level may include any part of what comprises the macro level. This includes any subcomponent or subnetwork (in a relational sense, as opposed to an electronic sense) with two or more actors who comprise the subsystem. Each subsystem may be defined by a common function or the need to exchange information and collaborate to accomplish a common mission or task. Examples include a brigade combat team, a ship, a fire support net, an airborne warning and control system, a sustainment group, a humanitarian relief task force, or a group of key actors collaborating on actions along one LOE.

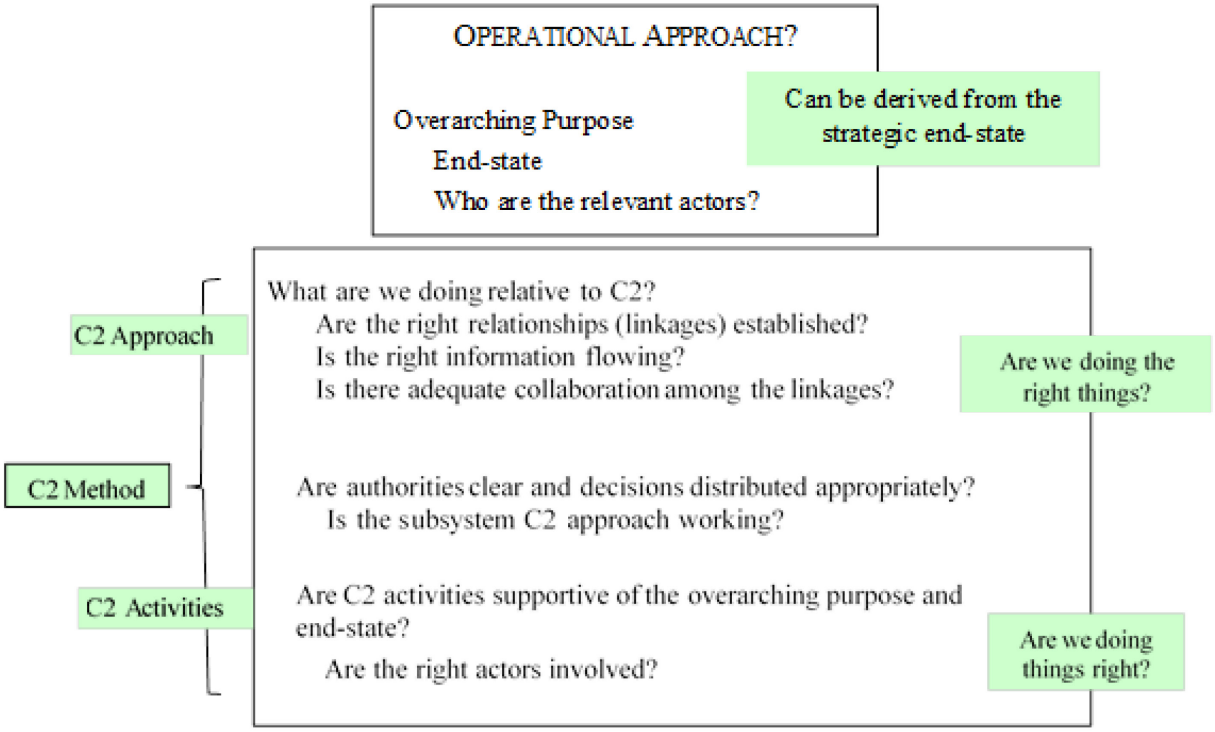
Here, the assessment of the C2 Approach in more detail is important as it can both 1) improve the operation of the subsystem and 2) contribute to the maintenance or improvement of the macro C2 Approach. Assessment at the subsystem level is similar to the macro assessment but it is generally simpler because it involves fewer entities and it is a more focused need for information and collaboration. Because it is simpler, problems are easier to identify and resolve. The process parallels the macro assessment but with a sharper focus. Moreover, it is even possible to make a useful C2 Assessment absent the details of the design, the broad operational approach, and possibly the C2 Approach.

Table 10-1: Key Elements of Macro-Level Assessment and Red Teaming.

Macro Assessment	Macro Red Teaming
What is the intended C2 approach? • Metric: The C2 plan has observable elements Is the C2 approach as implemented what was intended? • Metric: Actual C2 structures and activities are observable Is the C2 approach working? Is it enabling both the operational approach as a whole and its individual lines of effort? • Metric: Bottom-up reporting, not just on linkages but, more importantly, on whether the information flows, collaborations, and decision authorities are healthy and enabling both timely decisions and action. Reporting would be on friendly C2 information requirements	What has changed or could change in the operational environment that will impact the C2 approach? Example categories: • Mission change or mission creep • Organization (own or external) • Actors (more or fewer) • LOE (progress or lack of progress) • Changes in the enemy situation (positive or negative) or in factors beyond the commander's control that work for against mission accomplishment (such as weather and terrain) • Communications security compromises What are the most important changes to address first? • Consider risk and urgency? How will the most important changes impact the C2 approach? • What adjustment would be required? What indicators would illuminate change in the operational environment and how can they be monitored? • How can this be implemented? What are the commander's C2 information requirements?

The subsystem assessment requires at least a general understanding of the operational approach (upper box in Figure 10-1). Specifically, it requires knowing generally what the ongoing operation is trying to achieve. This includes the overarching purpose and the end state. Additionally, the assessment should be based on a clear understanding of how the subsystem fits into the operation (e.g., are coalition forces aiming to destroy the enemy or to compel surrender with minimal damage to infrastructure, and why?). Next, who are the other

entities, both habitual and new, that must interact to ensure the subsystem activities are contributing to a successful outcome? The assessment then requires an understanding of the subsystem C2 Approach and whether that C2 Approach is working effectively (lower box in Figure 10-1).



Note: Figure prepared by IDA authors.

Figure 10-1: Subsystem-Level Assessment.

Viewing the subsystem as a whole, the operator should evaluate the C2 Approach from the perspective of either the operational approach or the overarching purpose and end state. Three questions help focus this assessment:

- First, what is the operational approach, if relevant and known to the subsystem?[11]
- Second, whether the operational approach is known or not, what is the overarching purpose and end state?
- Finally, who are the relevant actors necessary to execute the operational approach or the overarching purpose and end state?

With respect to the subsystem itself, another set of questions guides the assessment. First, "Is what the subsystem actually doing or attempting to do relative to C2 making sense?" For example, a fires subsystem re-tasked to take on an additional function (e.g., civil affairs) in a counterinsurgency operation would require significant C2 adjustments. This question checks on what is happening with respect to C2 within the subsystem. Key questions in this area are as follows:

[11] The operational approach may not have been shared or known among all actors that are operating in a rapidly changing, common operational environment – for example, there may be little opportunity to communicate among emergency responders in the wake of a natural disaster.

1) Are all the right relationships (necessary linkages) established?
2) Are appropriate information flows occurring within each linkage?
3) Is there adequate collaboration between and among actors and linkages to achieve the operation's overarching purpose and end state?
4) Do the right actors in the subsystem have the right decision authorities to achieve the overarching purpose and end state?

If the above assessment concludes that the C2 activities make sense, the next question becomes: "Is it working effectively?" Answering this question necessitates an examination of C2 activities to determine:

1) Whether they are supportive of the overarching purpose and end state (here it is important to understand whether the C2 activities contribute to building the shared understanding that enables necessary decision making); and
2) Whether the right actors or entities are involved.

The answers to these questions will determine whether the subsystem C2 is healthy and, if it is not, where to make changes to improve the subsystem.[12]

10.7 ASSESSMENT DETECTS CHANGE; CHANGE DEMANDS AGILITY

Given that design and the resulting operational approach are based on a preliminary understanding of the operational environment, both are subject to change and very likely will do so. Both require assessment to determine when change is needed either because assumptions prove invalid or because of operational dynamics. Similarly, the C2 Approach can be expected to change or require modification as events unfold and learning takes place. Thus rigorous and continuous C2 Assessment is needed at both macro and subsystem levels to ensure the C2 Approach is aligned with, and supportive of, the operational approach and its LOEs. Figure 10-2 illustrates an example of how a process that designs C2 can complement the operational design process.

In Figure 10-2, C2 Assessment and Design are outlined in the dashed box, envisioned as a parallel and enabling supplement to the example design methodology. As shown in Figure 10-2, assessing and designing C2 involves feedback assessments and reframing loops that are needed to synchronize operations with more effective C2 methods. However, unless C2 is assessed in this way, it will not be clear whether C2 is contributing to or inhibiting operational success.[13] Thus, C2 Assessments are integral to getting the C2 method right.

[12] Managing and directing changes to the C2 Approach within a joint task force or similar staffing can be achieved through an adaptation of the '7-Minute Drill' (see Ref. [1] p. B-1). The adapted drill is known as the '10-Minute Drill' (see Ref. [1], B-2). Incorporating the '10-Minute Drill' entails C2 Activity owners, on their own, conducting a daily '10-Minute Drill'. The daily frequency would allow the incorporation of new command guidance and factoring in potential changes in circumstances. The chief of staff (or other battle rhythm event owner, e.g., Battle Update Briefing) could then consider tasking at least one C2 Activity owner (e.g., the Targeting Work Group) each day to present the results of its '10-Minute Drill' for review during the battle rhythm event. All C2 Activity owners in attendance would listen, participate, and learn from the presentation and resulting command guidance, thereby informing appropriate adjustments to their own '10-Minute Drill'. In short order, this could have the effect of shaping the C2 Approach across the headquarters, potentially affecting all battle rhythm events, as circumstances change. The '10-Minute Drill' includes brief inquiries into the problem, the context, linkages, information distribution, decision rights, and assessing changes to the C2 Approach.

[13] The reader is encouraged to review historical examples of C2 Approaches that contributed to or inhibited operational success found in Ref. [1][1] Appendix A.

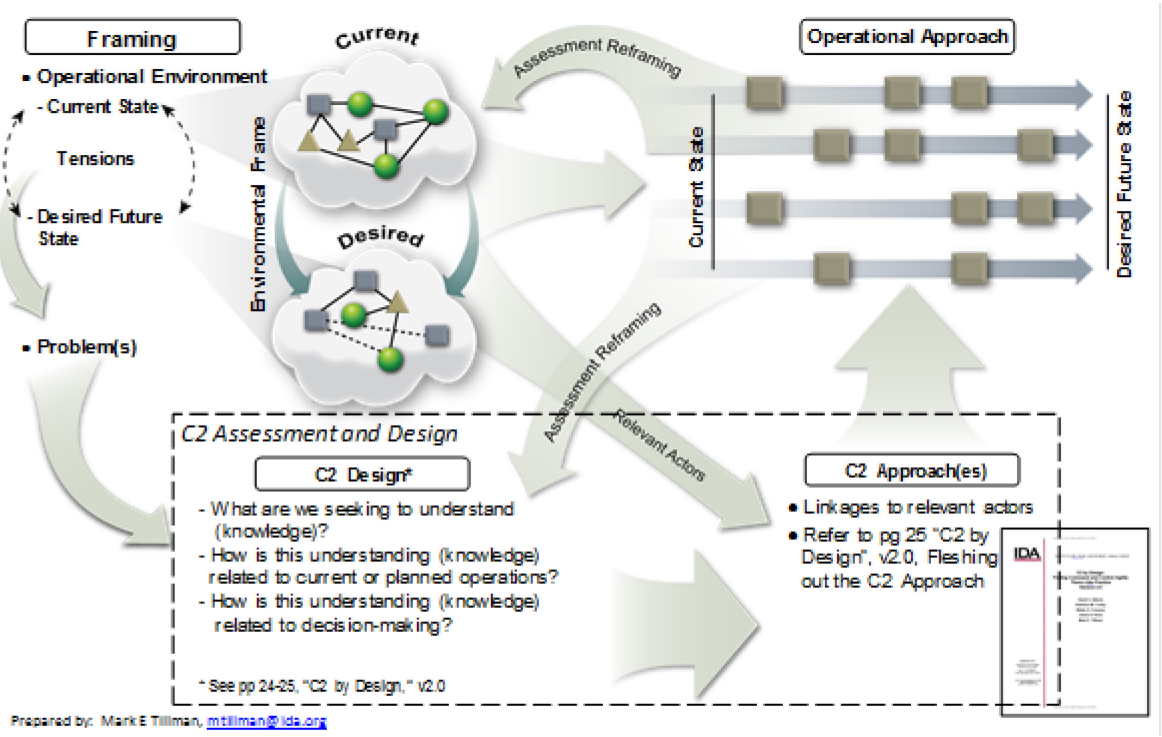


Figure 10-2: The Army Design Methodology Enhanced with the "C2 by Design" Assessment Approach. Note: The upper half of this figure was adapted from Ref. [3].

10.8 CONCLUSION

Assessing C2 in complex operations is achievable. The assessment is important, even vital, because it leads to effective C2 and can enable a well-designed operational approach to succeed. Indeed, all operations assessments are a part of the C2 function and, if properly focused throughout the design process, inform the leadership about the effectiveness of the actions underway. Operations assessment raises red flags when plans get off track. Similarly, assessing C2 can inform the commander as to whether the C2 Approach and activities are supportive of the overall operational approach and, ultimately, of achieving the desired end state.

10.9 BACKGROUND READING

Ricks, Thomas R. 2006. Fiasco: The American Military Adventure in Iraq. New York: Penguin.

West, Bing. 2008. The Strongest Tribe. New York: Random House.

10.10 AUTHORS' BIOGRAPHIES

Mark E Tillman, Colonel, US Army (retired) is currently a Research Staff Member at the Institute for Defense Analyses (IDA). Before joining IDA he served as a professor of National Security Studies at the National War College in Washington, D.C. and served on the USCENTCOM staff as the war planning chief for US military operations in Iraq and Afghanistan (Combat Operations ENDURING and IRAQI FREEDOM). Colonel Tillman has served in various operational assignments with infantry, armour, and field artillery units worldwide, including the storied 82d Airborne Division in combat. He has commanded tactical

field artillery units and served as an assistant professor of Systems Engineering at the United States Military Academy (West Point). Mr. Tillman is a graduate of the US Army Command and General Staff Officers' Course; holds Masters of Science Degrees in Operations Research/Applied Mathematics from the Colorado School of Mines and in National Security and Resource Strategy from ICAF (Eisenhower School); and as a Seminar XXI fellow holds a diploma in National Security Studies from the Massachusetts Institute of Technology (MIT).

Kathleen Conley, Colonel, US Air Force (retired) is currently a Research Staff Member at the Institute for Defense Analyses (IDA). Prior to joining IDA, she served as the Director of Land Forces Division, Cost Assessment and Program Evaluation (CAPE) in the Office of the Secretary of Defense. She provided direction and analysis for high-level ground forces studies, articulated guidance for future service programs, assessed service program proposals, and assisted in acquisition milestone reviews. She was also Director of CAPE's Projection Forces Division. She previously served in the US Air Force as a C-141 and C-17 pilot and in a variety of operational and staff assignments. Her education includes a Bachelor of Science Degree in Operations Research, Management, and Humanities from the United States Air Force Academy, a Master of Science Degree in Operations Research from Cornell University, a Master of Arts Degree in National Security and Strategic Studies from the Naval War College, and a National Defense Fellowship at Harvard University's Olin Institute.

10.11 REFERENCES

[1] Alberts, D.S., Conley, K.M., Freeman, W.D., Kurtz, J.H. and Tillman, M.E. 2015. C2 by Design: Putting Command and Control Agility Theory into Practice, Version 2.0. Alexandria, VA: Institute for Defense Analyses, September. https://www.ida.org/~/media/Corporate/Files/Publications/IDA_Documents/SFRD/2016/D-5614.ashx.

[2] NATO Research and Technology Organization. 2014. Command and Control (C2) Agility. Technical Report. STO-TR-SAS-085. http://ftp.rta.nato.int/public//PubFullText/RTO/TR/STO-TR-SAS-085///$$TR-SAS-085-ALL.pdf.

[3] Headquarters, Department of the Army. 2015. Army Design Methodology. US Army Techniques Publication (ATP) 5-0.1. Washington, D.C.: HQDA, July. http://armypubs.army.mil/doctrine/DR_pubs/dr_a/pdf/atp5_0x1.pdf.

[4] Alberts, D.S., Huber, R.K. and Moffat, J. 2010. NATO NEC C2 Maturity Model. Ft. Belvoir, VA: Defense Technical Information Center, February. http://www.dtic.mil/dtic/tr/fulltext/u2/a555717.pdf.

[5] Joint Chiefs of Staff. 2011. Joint Operations, Joint Publication 3-0. Signed by M.G. Mullen, ADM, U.S. Navy. Washington, D.C.: 11 August. http://www.dtic.mil/doctrine/new_pubs/jp3_0.pdf.

[6] U.S. Marine Corps. 1996. Marine Corps Doctrinal Publication 6: Command and Control. Signed by General Charles C. Krulak, Commandant of the Marne Corps. Washington, DC: HQ, USMC, 4 October. http://www.marines.mil/Portals/59/Publications/MCDP6CommandandControl.pdf. (Revised version, 20 June 1997. http://www.marines.mil/Portals/59/Publications/MCDP1Warfighting.pdf).

[7] Joint Chiefs of Staff. 2011. Joint Operation Planning, Joint Pub 5-0. Signed by M. G. Mullen, ADM, U.S. Navy. Washington, D.C.: Joint Staff, 11 August. http://www.dtic.mil/doctrine/new_pubs/jp5_0.pdf.

Chapter 11 – SPEAKING TRUTH UNTO POWER (STRATEGY, ASSESSMENT, AND DECISION MAKERS)

Bruce Pennell
NATO Communications and Information Agency
UNITED KINGDOM

Sergio Miller
Defence Consultant
UNITED KINGDOM

ABSTRACT

Recent conflicts in Afghanistan, Iraq, Libya and Syria demonstrate the challenge faced by NATO and allies in developing strategy, and intervening in crises to pursue strategic goals. Analysis and assessment of whether, and to what extent, interventions are making effective progress towards achieving these goals, are therefore key components of this wide-ranging issue.

Critics have claimed that the problem resides in poor strategy, overambitious or unachievable goals, or unworkable bureaucracy. Other critics suggest that the main issue is inadequate, inappropriate or simply over-optimistic assessments.

This chapter seeks to examine the difficult relationship between strategic goal setting, analysis and operations assessment. This requires some consideration of how strategic goals are developed, a look at the difficulties in honestly assessing progress in any mission, and a reconsideration of the decision-making model which underpins current analysis and assessment practices. The chapter draws on examples from recent (and less recent) history to make some recommendations for the future.

11.1 BAD NEWS

Strategy making is difficult. How conflict plays out is unpredictable, setbacks inevitably occur, and how to bring bad news to leaders when things go badly is a perennial challenge. So much is unremarkable, and well understood by the earliest commentators.

Due to the uncertainty inherent in conflict, wisdom lies in understanding when bad news represents just a 'bump in the road,' a more grave development demanding a change in plan, or – *in extremis* – a revision of strategy. The role of advisors and supporting staffs must therefore be to offer the analysis and insight which enables strategic leaders to achieve that understanding and to decide what to do in response.

The challenge is not new to strategists. Around 500 BC the original Athenian 'Strategos' were elected leaders who also acted as military chiefs and collectively formed a decision-making council to select policies and courses of action for the state in peacetime and in war [1]. Plutarch, the Greek historian, wrote later in his *Life of Lucullu*s, of the dangers which can befall advisors to such strategists. Lucullus was pursuing a military campaign against Tigranes the Great, King of modern-day Armenia, and mounted a surprise attack. Tigranes was alerted to this by a messenger who, instead of being rewarded for the information, was:

> *...so far from pleasing Tigranes that he had his head cut off for his pains; and no man daring to bring further information, without any intelligence at all, Tigranes sat while war was already blazing around him, giving ear only to those who flattered him* [2].

NATO's recent conflicts and crises have been accompanied by a vocal chorus of critics, claiming that NATO has been unable to think or act strategically, hobbled by over-optimistic assessments which promote poor

decision making. NATO has been unable to convince these critics and sceptical audiences that effective progress is being made. NATO Operations Assessment doctrine claims precisely to:

> *...enable the measurement of progress and results of operations in a military context, and the subsequent development of conclusions and recommendations that support decision making* [3].

If this is not, in practice, what happens, this should be sufficient cause for critical self-reflection.

This chapter therefore sets out to explain how and why the relationship between strategy, decision makers, and assessments staff does not function as intended. The chapter is illustrated by examples drawn mainly from NATO operations in Afghanistan, but uncomfortable parallels are also drawn with other campaigns. Although the views expressed in this chapter might be seen as unwelcome by some readers, Plutarch reminds that dismissing them and "giving ear only to those that flatter" is likely to be more disastrous than considering them in a spirit of critical reflection. In this same spirit, the chapter represents the authors' personal views and not those of any national or NATO body.

11.2 KEY ISSUES

Complicated problems rarely resolve themselves simply. For brevity in this chapter, we restrict ourselves to three areas we consider to be key:

- Goals – The challenge of strategic goal setting.
- Spin – Assessing Progress, Wishful thinking and Self-Censorship.
- Decision Making – The 'mysterious calculus' of decision makers and how they make use of analysis and assessment.

The first relates to how NATO deliberates before and at the start of an intervention; the second is concerned with ongoing operations and their assessment; and the third examines the predominant underpinning decision-making model from the perspective of current NATO doctrine.

11.2.1 Goals

In 2003 outgoing NATO Secretary General Lord Robertson struck a cautionary note in his farewell speech to the North Atlantic Council (NAC), warning that the political will expressed by allies in taking on an enhanced role in Afghanistan needed to be matched with real commitment:

> *Nations will have to waken up to what they have taken on. Expansion must be credible, and be seen to be credible. Money, troops and long-term commitment are the only ingredients of success – and the only way Afghanistan's problems will not come west to haunt us. Failure would be a crushing blow, not just for NATO but also for every NATO country...* [4].

Robertson's prescient words suggest he had already formed a view that the implications of the expanding NATO mission in Afghanistan (The International Security Assistance Force – ISAF) might not have been grasped by all participants. The evolution and expansion of the ISAF mission, from one narrowly focused on protection and security in the Kabul area in 2002, to a nation-wide mandate by 2005, was much more complicated than the basic description of NATO crisis response planning processes suggests [5]. NATO policy was influenced, or framed, by a wide range of factors which had little to do with the nature of the situation in Afghanistan, such as U.S. pressure on European NATO allies to show the relevance of NATO in a post 9/11 context, and the desire of some of those allies to show support for the U.S. in Afghanistan, in contrast to Iraq in 2003 [6]. By whatever means, by March 2006 a NATO Plan had been adopted which supported three strategic objectives. These were agreed consensually at the highest level by the member nations and the precise wording remains classified. However we can take from a 2010 unclassified U.S.

summary their approximate scope and for the purposes of this chapter we use these as placeholders for the actual, classified objectives [7]:

- To reduce the capability and will of the insurgency;
- To support the growth in capacity and capability of the Afghan National Security Forces (ANSF); and
- To facilitate improvements in governance and socio-economic development.

One possible key to understanding this expansion of the mission is offered by Schreer who argues that the NATO doctrinal description of strategy formulation as a top-down, politically driven process, failed to adequately capture what happened in Afghanistan [8]. This, he assessed, was much more dictated by events on the ground; from the bottom-up. NATO found itself forced by circumstance to take on additional responsibilities in the field of Security Sector Reform, Governance and Development Support, despite a light footprint and the constant challenge of encouraging NATO allies and partners to provide adequate forces and resources – as Lord Robertson had forewarned. At the heart of this bottom-up approach were Provincial Reconstruction Teams (PRTs). The PRT concept, which had been initially developed by the U.S. and other nations in their national interventions in Afghanistan, purportedly linked improved security with reconstruction and economic development throughout the country.

> *Focusing on quick impact projects, i.e., building critical infrastructure like water wells, schools and the like, PRTs link security with concrete development progress* [9].

PRTs, clearly derived from U.S. Counterinsurgency (COIN) doctrine [10] and U.S. experience in campaigns since Vietnam, continued under different lead-nations even while NATO ISAF assumed an overarching coordination role. However, despite the alluring simplicity of the *Clear-Hold-Build* mantra, and the logic of the PRT role within that, they did not prove the hoped-for panacea. While the role of PRTs was generally accepted, their composition, resourcing and engagement modes with other local and international actors all varied considerably between lead-nations. Overall their accountability and effectiveness was criticised by influential development practitioners and academics who pointed out, inter alia, that they:

> *developed into an incoherent network of lead-nation-driven units… counter-productive to Afghan ownership… bringing the risk that projects will be developed for short-term security rather than long-term development* [11].

Although NATO may have tacitly acknowledged some of this critique, it was unwilling or unable to refashion the PRT approach top-down.

By 2009, the salience of "bottom-up" strategy was made starkly clear when General Stanley McChrystal assumed command of ISAF. When his initial strategic assessment of Afghanistan became known to the media it opened a window into the challenges facing the mission. The specific content of the assessment is less relevant to this chapter than the fact that, according to eminent historian Hew Strachan, it reflected that:

> *McChrystal's strategy […] was shaped from the bottom-up; without a clear articulation by NATO or the United States of their political objectives and hence of their strategies, it could not be anything else. […] COMISAF's initial assessment was the best and fullest statement of what ISAF was seeking to do in Afghanistan in 2009. However, its attention was on the how rather than on the why; its focus was on the means, as the ends with which it was concerned were, in the standard hierarchy of military plans, essentially operational, not strategic, even if they posed as strategy* [12].

Strachan's implication is that if strategy is not clearly developed according to strategic aims, its chances of success are impaired. Strachan and several others have developed these ideas into broader critique of contemporary strategy formulation [13], [14], [15]. Their concerns are much broader than the scope of this chapter but are of interest if poor strategy formulation (as they see it) somehow intrinsically inhibits the understanding of whether operations can be effective.

To be clear, it is worth addressing this point more specifically by considering some generic crisis in which NATO develops plans to intervene. Let us assume it has 3 strategic aspects or domains {X, Y, Z}, each with an associated desired goal which NATO allies have agreed must be achieved to resolve the crisis. These may be purely military, or – more generally – be spread across a range of political, security and socio-economic areas. Let us also assume that these aspects and goals are not so esoteric or technical that they cannot be communicated in plain language. It should be clear enough that measuring progress towards these goals therefore simply requires an assessment of {X, Y, Z} and a comparison with similar assessment made at some earlier point in time, in order to judge whether the current status is closer to the agreed goals or not.

Even from this simple pseudo-model, it should also be evident that:

- Given that the {X, Y, Z} can represent any real world characteristics, they clearly do not need to conform to any specific rule such as independence (orthogonality) for an assessment to be made. A corollary is that change in each domain will only in the most unusual cases be at the same rate, measurable on the same scale, or be able to be aggregated meaningfully into a single dimension.
- Assessing change in {X, Y, Z} cannot in itself determine whether (and how) NATO actions are actually impacting on the situation. Similarly, forecasting or predicting the future status of {X, Y, Z} may equally be affected by other events and factors outside NATO control.
- If {X, Y, Z} are poorly selected, and do not in fact represent the most important aspects of the crisis to be resolved, then assessing change in these areas will likely be peripheral (if not irrelevant) to determining real progress. Similarly, if the methods adopted to assess {X, Y, Z} do not in fact give insight into their actual state, then the assessment will be unconvincing even when aspects (dimensions) and goals are well chosen.

The point of this somewhat laboured analysis is simply to emphasise some fundamentals, irrespective of the quality of strategy, or the number and type of goals chosen.

Firstly, all strategic goals in the real world should be broadly understandable in plain language, and can be assessed, irrespective of whether they are the best strategic goals which could have been chosen.

Except in the simplest cases, each aspect (domain or dimension) is likely to require a different method, approach and measurement 'scale' to assess it. Deciding whether an observed change in any domain is genuine and significant, rather than random variation or anomalous, may therefore require specialist expertise in that domain. (Of course, this does not resolve the issues arising when experts disagree!)

Finally, assessing change is not the same as assessing effectiveness; change may be attributable to many causes other than NATO intervention. However, if change cannot be assessed, or only negative change (movement away from strategic goals) is perceptible, then it is hard to argue the intervention is being successful. Assessing change is a necessary, but not sufficient, condition for operations assessment as defined by NATO.

Returning to the goals adopted by NATO in Afghanistan (as set out in 2006) we can see that, whether or not we agree they reflect 'good' strategy, they could be expressed in natural language and they should, therefore be amenable to being assessed. Hence it is illustrative to examine briefly how this assessment was done.

11.2.2 Spin

Anthony Cordesman, Chair in Strategy at the Centre for Security and International Studies (CSIS), has described the Afghanistan war as a 'Liar's Contest'.[1] He saved particular scorn for the way in

[1] A phrase attributed to Mark Twain. Liar's Contests are still held in his name in the USA, where opponents attempt to outdo each other's improbable stories.

which progress assessments produced by ISAF, the U.S. Central Command HQ (CENTCOM) and NATO were – in his view – consistently inadequate:

> *Corruption, nepotism, ethnic and sectarian splits, and the role of power brokers is grossly understated in all public reporting or simply ignored. The ability to create a credible, sustainable ANA [Afghan National Army] for 2014 and beyond is not explained, and the role of a highly corrupt mix of police forces in dealing with the justice system and insurgency has virtually no transparency or public credibility. Public estimates of future outside aid, trainers, and partners also lack transparency and credibility* [16].

Cordesman took particular issue with the kind of public statements made by senior NATO officers, of which the following are deliberately anonymised examples from the 2007-11 period:[2]

> "*So in summary, I'm pretty optimistic what's going on here. … I'm seeing some significant progress. You look at the institutional development; it's been phenomenal… overall, I'm incredibly optimistic…*"
>
> "*…from my point of view, we are making progress. Construction is challenging there… but we are very steadily making progress in a wide variety of areas…*"
>
> "*First, we are steadily making deliberate progress across Afghanistan, and we are on an upward trajectory.*"
>
> "*So we're definitely making huge progress...We continue to provide feedback... how to continue their success and provide security to the people, and messaging with them on how to connect with the population and having tremendous effect there.*"
>
> "*Now, we have made great progress with our civilian counterparts, both in the Afghan government and the international community … Evidence of our progress is clear.*"

Plainly, NATO generals and other senior leaders do not set out to be dishonest, and it would be unusual if the public statements of senior leaders were always comprehensive assessments of detailed aspects of the mission. Yet anyone involved in preparing such assessments being passed up the chain of command during this period, will likely recognise the same upbeat tone, which was maintained despite the fact that none of the resulting assessments (public or classified) convinced Cordesman, or other sceptical critics. How can this mismatch be explained?

One possibility is that the speakers sincerely believed what they were saying at the time. If so, this invites the question of how they became so convinced, whilst at the same time being so unable to convince critics such as Cordesman. A more plausible argument is that the speakers knew the situation was not as they were describing it in public but they were being 'politic' or 'spinning' their comments by focusing on the positive and diminishing or ignoring the negative. However, this also leads down an uncomfortable road. It may make sense as part of a Strategic Communications strategy to emphasise the positive, but it should be clear that such an emphasis can only have a limited role; a war cannot be won by spin alone. To spin is to lie by omission, and a lie once told can only be deepened. Moreover, spinning undermines the honest endeavours of subordinates, who are precisely those most acutely aware of the disparity between what they are experiencing, and what is being claimed by their seniors.

A third possibility is that some combination of self-delusion and spinning, or "active filtering" of bad news, was exacerbated by a more subconscious preferencing of positive information. This can be termed wishful thinking or, more specifically, confirmation bias. Such a bias is just one of a number of observable cognitive deficiencies which can be shown to impact on the quality of assessment in a range of situations [17]. Former UK Ambassador in Kabul, Sherard Cowper-Coles recalled of his role in these assessment processes:

[2] All of these statements were made in HQ ISAF or HQ NATO Press Briefings. Many are still available at http://www.nato.int/cps/en/natohq/.

> *We all wanted to believe that it [the plan] was working; we wanted to please Ministers, the armed forces and the Americans. There is nothing new about this; the same thing happened in the early years of the Vietnam war, when the best and brightest round John F. Kennedy [….] used the phrase that we used ourselves in Afghanistan: 'Progress is being made, but challenges remain.' It was wishful thinking, rather than some massive conspiracy* [18].

Lest such wishful thinking is regarded as an anomaly from Afghanistan (or Vietnam), it should be recalled U.S. Secretary of Defense Ash Carter felt compelled in September 2015 to re-affirm his commitment to the need for frank and honest intelligence reporting, following claims that pessimistic assessments of the ongoing air campaign in Syria against so-called Islamic State (IS) were being downplayed [19]. Senior analysts complained that their reports were being manipulated and a formal investigation was opened to establish whether senior leaders, including the director of intelligence at CENTCOM, changed analysis and assessments to conform more positively to a 'narrative' of progress [20].

The complex phenomenon of spin, over-optimism, wishful thinking or self-delusion appears by no means limited to the military/intelligence domain. Since the first draft of this chapter was submitted, a number of books have commented on the seemingly inexorable rise in almost all political and business spheres [21], [22]. Notwithstanding these commentaries, it seems still pertinent to examine the role of Operations Assessment staff in sustaining or challenging the particular military variant.

The aim of any commander and his supporting staff should be the development and implementation of winning plane. In order to achieve this, most military cultures recognise the need for the staff, and in particular Operations Assessment teams, to "*provide advice that contradicts and challenges, as well as confirms, even the most advanced plans*…" [23]. NATO doctrine echoes the same ethos, but is perhaps more muted, putting the onus on the Commander to generate with his staff "*a climate of mutual loyalty and respect rather than one that is sycophantic and unquestioning, the ability to tolerate 'loyal opposition"*" [24]. In theory at least, therefore, assessment staffs should be able to bring evidence and analysis to bear in order to counter cognitive bias, wishful thinking or other shortcomings.

It should therefore have been a crucial safeguard to the NATO mission in Afghanistan when a specialist team of operational analysts from the UK Defence Science and Technology Laboratory (Dstl) deployed to the fledgling ISAF HQ in Kabul in winter 2002 specifically to establish a formalised Operations Assessment staff. This team adopted a process of collating and analysing evidence to inform around 20 so-called Measures of Success designed to inform the overall assessment of whether and how campaign goals were being achieved. Looking back at contemporary accounts of these early operations assessment efforts there is an air of enthusiastic, pioneering optimism [25]. Data of every type was collated – ranging from the precise location and type of attacks on NATO forces, through to early efforts to garner the – views and attitudes of Afghans towards NATO and to the new Government through opinion polling. In subsequent years these NATO/ISAF efforts were supplemented by often overlapping and conflicting non-NATO efforts – National Assessments, assessments from the United Nations Assistance Mission in Afghanistan (UNAMA), and periodic reports from influential International Organisations (IOs) and Non-Governmental Organisations (NGOs) ranging from the World Bank Group, The Asia Foundation and Crisis International to name just a few.

Despite all this effort two things are apparent. Firstly the (broadly optimistic) in-house assessments from HQ ISAF and other HQs, however developed, did not appear able to reconcile the conflicting evidence and analysis from external (broadly less optimistic) assessments. Secondly, and perhaps more surprisingly, none of it seemed to matter. That is to say, it is hard to point at any strategic decision which apparently depended on the assessments made within the HQs. It was as if the entire assessment effort was instead serving another, less well-defined purpose, in order to be coherent with some sort of 'top-down' picture, notwithstanding contrary evidence.

The pernicious effects of these circumstances should be self-evident. No amount of sound assessment and analysis will count if 'wishful thinking' prevails. In such a climate, honest assessment is not just suppressed, but selectively cherry-picked or neglected, so that, ultimately, the assessment system adapts to meet these new requirements and incentives, rather than contradicting and challenging. In Afghanistan, assessment methods were endlessly revised where they inconveniently failed to support the desired arc of progress, and assessment and analysis effort were concentrated instead into reporting tactical-level actions which looked or felt positive, even though they were largely irrelevant to strategic progress.

Again, any reader thinking this overstates the case might pause to consider another ISAF example. We have already noted above that supporting the development of ANSF capacity and capability was an agreed strategic goal for NATO in Afghanistan. Yet the basic method and criteria for assessing whether ANSF capability was developing was changed repeatedly, to the extent that it became almost impossible to determine whether any progress had been made at all. In 2010, according to the refreshingly blunt reports made by the U.S. Special Inspector General for Afghanistan Reconstruction (SIGAR), the assessment system in place since 2005 was criticised for being inconsistent, over reliant on quantitative metrics and inadvertently creating disincentives to progress [26]. In 2014 a similar independent evaluation concluded the replacement assessment method was itself being replaced because ISAF leadership found it "*difficult to read, inconsistently applied, and not useful*" [27]. Despite all the efforts made to refine, revise and recalibrate the measure of ANSF development, after 10 years it was still apparently unclear whether progress could be determined at all.

11.2.3 Decision-Making Models

Arch-rationalist Robert McNamara counter-intuitively observed a 2003 retrospective documentary film: "*…rationality will not save us*" [28]. The 'Pentagon Papers' study now housed in the U.S. National Archives make a similar point in respect of U.S. policy and decision making in the Vietnam-era (authors emphasis added in bold):

> *Only upon the basis of interpretations (judgements) of the importance, meaning and relevance of things could policy be made. And that* ***judgement or interpretation was seldom or never inescapably inherent in the measurable, sharply definable, completely unarguable concrete detail****. It might be derived from or directly reflect such data, but* ***its form would be determined equally, or even more, from the perspective in which it was viewed.*** *And this perspective was comprised of the whole context of incompletely described, not fully identified values, and imperfectly described priorities, that determined the weight and place given to the factual detail in the mysterious calculus of the decision maker* [29].

If it were just a matter of getting all the facts – and there is hardly a lack of fact-getting in NATO assessments and intelligence staffs – then strategic decision making should be straightforward. Strategic decision making may be more complex given the challenge of both NATO (consensus) and national (more or less democratic) interactions. Influential policy analysts argue that it is complexity in decision-making mechanisms, as much if not more than the intrinsic complexity of the problem, which explains best why rationality is not enough [30], [31].

Majone has argued for many years that scientific-rational models – what he terms the 'decisionist' paradigm – neglect the essential characteristics of democratic bureaucracies, and hence the process by which decisions and policies are made by them [32]. Dialogue, persuasion and argument instead dominate between competing political actors: the executive, legislature, judicial and other branches of formal government, along with informal influencers. To be effective, therefore, analysis and assessment in support of decision making must reflect the needs of these actors and influencers, rather than focusing on a narrow, technical formulation of the problem in terms of (so-called) facts. All constraints are relevant to considering whether a policy option is feasible, not just the technical aspects, even if they are most amenable to quantitative techniques and therefore most appealing to analysts.

Majone also strongly argues that particular attention should be placed by analysts on evidence and its use in the decision-making process (the 'analytic argument'):

> *The nature of the evidence plays a crucial role … evidence is not the same as data or information. Rather, it is information selected from the available stock and introduced at a specific point of the argument in order to persuade a particular audience of the truth or falsity of a statement* (Ref. [30], p. 10).

Majone reminds readers that the disciplines of law and history routinely deal with problems concerning both complex arguments and voluminous information, where the reliability and relevance of this information itself cannot automatically be trusted. By contrast, in the 'decisionist' paradigm, quantitative evidence is considered as having intrinsic legitimacy, hence it is often preferred or weighted more heavily than it deserves. Majone further contends that, whilst the legal profession in particular acknowledges the crucial role of persuasion in creating a winning argument, in the rational-scientific disciplines the role of persuasion is downplayed:

> *Rational arguments are only partial determinants of attitudes: an element of persuasion is involved in any attempt to suggest one rather than another view, judgement, or course of action. Thus in order to be effective, an analyst must also be an advocate. But he is also a firm believer in the virtues of the scientific method, and this belief is generally associated with a distaste for problems of communication and persuasion* (Ref. [32], p. 19).

Another celebrated policy analyst, Deborah Stone, is also critical of rational decision-making models for policy formulation in democratic organisations (or alliances like NATO). Stone likens the rational policy model to efficient market theories of economics – more useful for the application of mathematical techniques than a helpful representation of the world as it is. Stone offers by contrast a 'polis' model – where inducements, formal and informal rules, rights, and powers are the driving forces. In the polis model, symbolism is more important than rationality. Successful policymakers harness symbols in the form of narrative stories, metaphors and ambiguity to mobilize support for their desired policies – to persuade, in Majone's terms. Some features of successful policy formulation identified by Stone – such as stating goals ambiguously and keeping undesirable alternatives off the agenda by not mentioning them – may appear unethical, but they nevertheless chime with some of the issues identified earlier in this chapter.

What does this mean for analysts? Firstly, analysts should be aware that the problem space is almost always broader than the policy objectives. Secondly, the so-called 'facts' of the matter will rarely be conclusive, but instead the symbols deployed by advocates of one policy position or another. In the Afghanistan context, NATO's strategic goals may have focused on Afghanistan but the problem space extended geographically across Pakistan, Iran and the Central Asian states, and also conceptually included the political will and cohesion of the Alliance and its partners. Additionally, this problem space will range across topics which cannot be represented by simple linear measures. This does not mean quantitative techniques have no role to play, but that they will require augmenting with a wider range of tools and approaches if genuine options and constraints are to be adequately portrayed.

Majone and Stone are just two (albeit significant) policy analysts who propose a more sophisticated understanding of the 'mysterious calculus' of strategic decision makers. This requires analysts and advisors at the strategic level to acknowledge the roles of advocacy and persuasion in exploiting evidence to support assessment and decision making. This does not mean that analysts should discard rational approaches. But neither can they ignore the real world political context in which rationality is set.

11.3 RECOMMENDATIONS

It is a cliché that the main lesson from history is that we don't learn from history. When British Prime Minister Harold Macmillan called his successor Alec Douglas-Home to offer advice, he is alleged to have

reassured him: "My dear boy, as long as you don't invade Afghanistan you'll be absolutely fine" [33]. The authors of this chapter cannot match Macmillan's distilled wisdom. Instead a few observations are offered by way of recommendation.

11.3.1 The Wisdom of Experts…and of Crowds

Since 2010 NATO has set itself the challenging task of being more comprehensive in its planning; to recognise that non-military aspects, actors and goals must be considered as part of any complex intervention [34]. Arguably, however, it has lagged behind in developing how to develop comprehensive understanding of these aspects.

NATO does not routinely employ a wide range of experts with deep understanding in non-military domains. In theory, it maintains institutional links with relevant experts and think tanks that offer this understanding, but in practice these links could be stronger and more systematic. Many authoritative experts work further afield than in NATO establishments or national administrations – in academia, in smaller NGOs and in industry. Outside experts are unlikely to hold necessary security clearances. Some may be antipathetic to NATO. Yet their views are important and NATO requires the will and means to access them.

Innovative ways can be found to find and exploit such specialist expertise. One route might be more routine use of the civilian committees and networks which NATO maintains with allies and partners. Another might be experimentation with 'crowdsourced' or other 'superforecasting' techniques, some of which claim to routinely outperform intelligence analysts [35]. None of these approaches are expensive, or technologically challenging, but they do require imagination, humility and boldness of vision.

11.3.2 Challenge Assumptions and Doctrine

Most people agree that a 'Devil's Advocate' is useful to uncover weak logic and explore contingency plans, who should not be regarded as an irritant, but essential. The problem, recognisable to anyone who has worked in large organisations, is that nobody wants to be the Devil's Advocate. Conformism is a powerful force. Questioning can be viewed as disloyalty and even as career – threatening. Leaders can be intolerant of challenge or dissent. Only after the event does it become permitted to announce that flaws were obvious all along. Critics might say this chapter is a good example.

NATO already practises 'Red Teaming' and has embraced 'Alternative Analysis' techniques as a stated part of its decision-making process. But unless the environment and culture can be created within which 'devil's advocacy' is firmly institutionalised then overall decision making will not benefit, particular when the decisions have strategic impact.

11.3.3 Offer Legitimate Avenues for Redress

The 2015 controversy over whether U.S. intelligence staff assessments were being 'sanitised' came to light through the actions of whistleblowers, analysts acting outside the chain of command by writing openly or leaking material to the media. Whistleblowing can be seen as a noble and ethical imperative, but inevitably is harmful to the organisation which finds itself the object of such action.

To avoid this embarrassment there should be legitimate, institutionalised redress mechanisms for staff who claim they are being put under pressure to conform to a political 'narrative', or who feel that their products are being unfairly misrepresented. Such a mechanism may act as a 'safety valve' for analysts and also as a restraint on those tempted to manipulate or misrepresent them.

11.3.4 Recognise Bias and Methodological Limitations

The ascendancy of the rational-technical model in the NATO decision-making and assessment doctrine has strengths and weaknesses. It is simple and methodical, and straightforward to teach, but it can lack agility and gives an impression of objectivity, and bias-elimination, which it simply does not merit. All processes of categorisation and identification of data for analysis are subjective, and are influenced by the background (and biases) of analysts. The weakness of some indicators and methods should be acknowledged, so that those areas where assessment can be more confidently made are distinguishable from more speculative conclusions. Margins of error and uncertainty in estimates should be reflected in the final assessment products.

Equally, analysts need to grow their current toolset beyond the well-developed statistical charting and quantitative data analysis methods. How to visualise diverse information – whether quantitative data, text or images – how it evolves over time, and how it can be communicated are equally important in building convincing analytic arguments, particularly at the strategic level.

11.3.5 Prepare to Recognise Failure

Rory Stewart, the British parliamentarian and former diplomat with extensive Afghanistan expertise, argued that "*the ability to recognise failure, and then to reform, is a defining mark of a serious country*" [36]. The analogy extends to alliances such as NATO, which has an extensive education and training philosophy in order to prepare for crisis response, yet does little – if the experience of the authors is at all representative – to practice how to recognise failure and to reform in response. By contrast, every NATO training event is always declared a success, sometimes even before completion. Such exercises do not prepare intelligence and assessment staff to identify the warning signs of progress not being made, or how to assimilate and reconcile uncomfortable information.

It may well be unrealistic to expect that the ritual of a major showpiece NATO exercise can ever be the venue for such failure-training, but suitable training events and methods have to be developed in which leaders at all level can experience the disconcerting sensation of a plan going slowly but inexorably awry, whether by inadequate understanding, shoddy planning, poor execution or sheer bad luck.

11.3.6 Recognise the True Scope of the Problem

Assessing strategic-level progress in a comprehensive intervention may require an extended scope well beyond the geographical limits of the crisis theatre or even the region. For example, in the ISAF mission a range of non-theatre factors were pertinent to the strategic situation including U.S. and Alliance domestic politics, relations with and between Iran, Saudi Arabia and Pakistan. These factors were doubly difficult to consider within strategic assessments – firstly because they were not explicitly described within the scope of the NATO mandate and secondly because they required specialist expert appraisal using methods which were not always easy to reconcile or weigh against other areas. Intelligence communities are comfortable with the notion of a broader 'Area of Interest' which lies beyond the 'Area of Responsibility' and such a conceptualisation would benefit the assessments community too.

11.4 SUMMARY

For half a century NATO honed its crisis management and warfighting skills through exercises and training designed to deter its main adversary and reassure allies. The absence of failure in these elaborate ritualistic and heavily scripted exercises had been interpreted as success. Processes of analysis and assessment were assumed to be valid and validated by such training, when in fact the skills required to recognise, brief and assimilate 'bad news' or lack of progress were simply not being developed. Wars in the 21st century have proved a shock; adaptable adversaries have failed to conform to NATO wishful thinking and NATO assessment products have failed to convince sceptical audiences.

This chapter has argued that NATO's understanding of strategic goal setting as a linear, top-down approach did not reflect the reality of the Afghanistan challenge. Nevertheless, although the selected strategic goals may therefore have been strategic only in name, they should nevertheless have been amenable to honest assessment. However, it is clear that progress assessments were inhibited by weaknesses in assessment methods, particularly in addressing essentially political aspects of strategic goals. An atmosphere of self-censorship in assessments prospered, smothering any inconvenient truths which might have challenged wishful thinking. Current NATO understanding of how assessments can or should inform decision making is rooted in an over-simplistic 'decisionist' paradigm. These factors combine in a tendency to focus inappropriately on lower level quantitative information which may be seen as useful, but which fail to adequately address strategic aspects.

The good news for NATO in addressing these issues is not technical but attitudinal; and it would indeed be the 'defining mark' of a serious organisation if it were now to accept and tackle them.

11.5 AUTHORS' BIOGRAPHIES

Bruce Pennell is a Senior Scientist with the NATO Communications and Information Agency and has worked primarily on Operations Analysis and Assessment issues since 2009. Previously a British Army Officer with operational experience in Northern Ireland, the Balkans, the Gulf and in South Asia, he has a technical background in Defence Simulation and Modelling, a Masters in Defence Studies and International Relations and a Postgraduate Certificate in Adult Education. He was co-chair of an award winning NATO Study into Operations Assessment and Planning for Transition in Afghanistan (SAS-091) and contributed to the 2013 NATO 'Innovation in Operations Assessment' report, along with many other internal papers and studies which remain largely unread.

Sergio Miller is a former British Army Intelligence Corps officer, with operational experience in Northern Ireland, the Gulf and with Special Forces. Since leaving the Regular Army he has worked in Defence as a self-employed consultant on a wide range of studies and projects. He is a regular contributor to the British Army's house journal, the British Army Review, and he continues to serve in the Reserves.

11.6 REFERENCES

[1] Cummings, S. 2012. Brief Case: The First Strategists. Long Range Planning, 26(3):133-5.

[2] Perrin, B. (translator) 1914. Plutarch: Life of Lucullus. Chicago. p. 551. http://penelope.uchicago.edu/Thayer/E/Roman/Texts/Plutarch/Lives/Lucullus*.html.

[3] NATO. 2011. NATO Operations Assessment Handbook; Version 1.0. p. 1-1.

[4] Robertson, G. 2003. Farewell Speech to NAC. http://www.nato.int/docu/speech/2003/s031217a.htm [Accessed Date: 17 Dec 2003].

[5] NATO. 2015. A comprehensive Approach to Crises. http://www.nato.int/cps/en/natolive/topics_51633.htm [Accessed Date: 1 Sept 2015].

[6] Steele, J. 2011. Ghosts of Afghanistan: Hard Truths and Foreign Myths. London: Portobello Books

[7] US GAO. 2010. Strategic Framework for U.S. Efforts in Afghanistan. US GAO-10-655R. Washington, D.C. http://www.gao.gov/assets/100/96809.pdf.

[8] Schreer, E. 2012. The Evolution of NATO Strategy in Afghanistan. Edstrom, and Gyllenspore, Pursuing Strategy: NATO Operations from the Gulf War to Gaddafi. London: Palgrave Macmillan. pp. 139-156.

[9] Koenders, A. 2006. Afghanistan and the Future of the Alliance. 174 PC 06 E. Brussels: NATO Parliamentary Assembly.

[10] US Army. 2004. Counterinsurgency Operations (ISBN 7-116-69200-2). Washington, D.C: U.S. Department of the Army.

[11] Fishstein, P. Wilder, A. 2012. Winning Hearts and Minds? Examining the Relationship between Aid and Security in Afghanistan. Medford, MA: Tufts University.

[12] Strachan, H. 2010. Strategy or Alibi? Obama, McChrystal and the Operational Level of War. Survival, 52(5):157-182.

[13] Strachan, H. 2013. The Direction of War. Cambridge UK: Cambridge University Press.

[14] Freedman, L. 2013. Strategy; A History. Oxford UK: Oxford University Press.

[15] Posen, B. 2014. Restraint, A New Foundation for U.S. Grand Strategy. Cornell University Press.

[16] Cordesman, A. 2014. Avoiding Creeping Defeat in Afghanistan: The Need for Realistic Assumptions, Strategy, and Plans. Washington D.C.: CSIS.

[17] Evans, J. 1989. Bias in Human Reasoning: Causes and Consequences. Hillsdale USA: Lawrence Erlbaum Associates.

[18] UK Parliament Defence Committee. 2015. Decision Making in Defence Policy. London: UK Parliament.

[19] Ackerman, S. 2015. U.S. spy chief's 'highly unusual' reported contact with military official raises concerns. http://www.theguardian.com/us-news/2015/sep/10/james-clapper-pentagon-military-official [Accessed Date: 12 Sept].

[20] Apuzzo, M. 2015. Pentagon Expands Inquiry into Intelligence on ISIS Surge. http://www.nytimes.com/2015/11/22/us/politics/military-reviews-us-response-to-isis-rise.html?_r=0 [Accessed Date: 22 Nov].

[21] D'Ancona, M. 2017. Post-Truth. London: Ebury Press.

[22] Davis, E. 2017. Post-Truth: Why we have reached peak bullshit. London: Brown, Little.

[23] UK MOD. 2010. Army Doctrine Publication: ADP Operations. 71632. UK Ministry of Defence. p. 0530.

[24] NATO. 2010. Allied Joint Publication (AJP) 01. p. 0637.

[25] Evans, D. 2003. Operational Analysis in Support of HQ ISAF. http://www.ismor.com/cornwallis/cornwallis_2003/2003_12Evans-Sept12.pdf [Accessed Date: 22 Nov 2016].

[26] SIGAR. 2010. SIGAR Audit 10-11 Actions Needed to Improve the Reliability of Afghan Security Force Assessments. Washington D.C.: Special Inspector General for Afghanistan Reconstruction, June.

[27] SIGAR. 2014. SIGAR Audit 14-33 Afghan National Security Forces: Actions Needed to Improve Plans for Sustaining Capability Assessment Efforts. Washington D.C.: Special Inspector general for Afghanistan Reconstruction, February.

[28] Morris, E. 2003. The Fog of War: Eleven Lessons from the Life of Robert S. McNamara. http://www.errolmorris.com/film/fow_transcript.html.

[29] US National Archive. 2011. Pentagon Papers, Part IV-C-1. p. 54. https://www.archives.gov/research/pentagon-paper.

[30] Majone, G. 1992. Evidence, Argument, and Persuasion in the Policy Process. New Haven, Connecticut: Yale University Press.

[31] Stone, D. 2002. Policy Paradox: The Art of Political Decision Making. New York: Norton.

[32] Quade, E. and Majone, G. 1980. Pitfalls of Analysis. Wiley and Sons.

[33] Dalrymple, W. 2013. In Afghanistan, we never seem to learn our lessons. Evening Standard, London, 10 May. http://www.standard.co.uk/comment/comment/william-dalrymple-in-afghanistan-we-never-seem-to-learn-our-lessons-8610992.html.

[34] NATO. 2015. Crisis Response – Who Decides? http://www.nato.int/cps/en/natolive/topics_49192.htm. [Accessed Date: 29 January].

[35] Tetlock, A. and Gardner, D. 2015. Superforecasting: The Art and Science of Prediction Crown Publishing Group: New York.

[36] Stewart, R. 2014. Afghanistan: Britain got almost everything wrong and should admit its failure. Daily Telegraph, 4 April. http://www.telegraph.co.uk/news/worldnews/asia/afghanistan/10745977/Afghanistan-Britain-got-almost-everything-wrong-and-should-admit-it.

Chapter 12 – DEALING WITH COMPLEXITY AND CHAOS – THE MILITARY EXPERIENCE

Jan Frelin
Swedish Defence Research Agency (FOI)
SWEDEN

ABSTRACT

Military doctrine stresses that defence forces have to handle complexity and chaos. In order to do that, it is necessary to have a working mechanism for feedback from the environment. Military theory stresses the requirement for an integrated feedback mechanism to enable command and control.

In this chapter, I investigate how the formal methods for creating effective feedback have assisted military forces over the last quarter century, using unclassified sources. I found that these methods have generally fallen short of reasonable requirements on effective feedback, creating an important gap between military practice and doctrine.

In the second part of the chapter, I look for possible explanations of the failure, using theoretical ideas from cognitive research, organisational learning, and the research on complexity and chaos. Finally, I suggest possible ways forward that shows promise in addressing this challenge.

12.1 INTRODUCTION

Military doctrine stresses that defence forces have to handle complexity and chaos. In order to do that it is necessary to have a working mechanism for collecting feedback from the environment. In this chapter, the aim is to investigate how the challenge of creating such an effective feedback mechanism has been met in practice.

The formal activity of analysing feedback from the environment is known as "operations assessment" [1], or "operation assessment" in U.S. doctrine [2]. Intelligence is the main source of feedback information, but operations assessment also requires analysis of the relationship between the situation in the field and the commander's own intentions and designs.

In the following, I investigate how operations assessment has worked in practice in four different major operations conducted by either U.S. led coalitions or NATO. The operations are Desert Storm 1991 (U.S. led), Allied Force 1999 (NATO-led), Iraqi Freedom 2003 (U.S. led) and ISAF 2003–2014 (NATO-led). The choice of cases was decided by whether the operations included significant combat operations, but also by the availability of sufficiently rich unclassified documentation in English. Thus a potential limitation of this study is the dependence on unclassified sources.

12.2 CASES

12.2.1 Operation Desert Storm

12.2.1.1 Background

Iraqi forces invade Kuwait on 2 August 1990, meeting only token resistance. Kuwait City is in Iraqi hands within 24 hours. In direct response, forward elements of the U.S. 82nd Airborne Division lands in Saudi Arabia on 9 August on Operation Desert Shield, in order ensure the territorial integrity of Saudi Arabia. U.S. and coalition forces are gradually built up in Saudi Arabia during the autumn of 1990, while the decision to liberate Kuwait is taken in Washington with the support of the UN [3].

On 17 January 1991, Operation Desert Storm is initiated with offensive air operations. Ground operations commence on 24 February. Kuwait City falls on 27 February, and on 28 February, President Bush declares a cease-fire. Apart from a few incidents that are quickly handled by U.S. forces, Operation Desert Storm is over [3].

12.2.1.2 Doctrine Situation

At the time of Desert Storm, there was no agreed joint doctrine on operation assessment. Battle Damage Assessment (BDA), which was the current U.S. term for assessing effects of tactical actions, was the only assessment doctrine in use. Further, no systematic training of assessment had been conducted [4], [5].

12.2.1.3 Experiences

Effective assessment was hampered by the deficiency in training indicated above, as well as shortages of staff in BDA positions, an overall shortage of intelligence assets, and information management problems which meant that relevant information did not always reach the right BDA cells. Weapon System Video was found to be an effective source of data for BDA when it was available. Despite that, the overall effect of these shortages was that no comprehensive picture of Iraqi losses could be assembled [5], [6].

The air campaign was aimed at reducing Iraqi ground forces by 50% before the ground offensive could commence. It was never possible to determine if this level of losses had been inflicted, but it was also not known why 50% losses was a good measure of reducing Iraqi fighting ability. The number was essentially arbitrary. Decisions had to be made anyway, and eventually General Schwarzkopf decided that it was time to launch the ground offensive based on qualitative deduction [5].

12.2.2 Operation Allied Force

12.2.2.1 Background

In 1998, open conflict emerged in Kosovo between a Kosovo-Albanian uprising and Serbian military and police forces. In response, the international community launched the Kosovo Verification Mission (KVM) under Organization for Security and Cooperation in Europe (OSCE) auspices. In January 1999, KVM reported a serious deterioration of the security situation in Kosovo. In March, KVM leaves Kosovo as the situation deteriorated further [7], [8].

On 24 March 1999, NATO initiated Operation Allied Force with an air campaign against targets in Kosovo and Serbia. On 10 June after a 78-day air campaign, NATO ceased operations as Serb forces withdraw from Kosovo [8].

12.2.2.2 Doctrine Situation

Since 1990, U.S. assessment doctrine had been updated and defined as Combat Assessment (CA), which was defined as consisting of BDA, Munition Effects Assessment (MEA), and Reattack Recommendations (RR) [5].

Conceptually, BDA was now divided into phases, phase I being physical damage assessment, phase II functional damage assessment, and phase III target system assessment [5].

The author is not aware of any NATO doctrines in use at the time of Operation Allied Force.

12.2.2.3 Experiences

In an Allied Force, the triangulation of intelligence sources and more efficient information management allowed for more effective BDA than in Desert Storm. Still, only about 60% of target hits could be substantiated after the war, and a big debate on how efficient the air campaign had been ensued. BDA was also confounded by Serbian deception. While deception can be considered an effect of the campaign, BDA processes were only able to see through deception to a limited extent. Again, quantitative objectives could not be measured reliably [5].

12.2.3 Operation Iraqi Freedom

12.2.3.1 Background

On 19 March 2003, U.S. forces launched Operation Iraqi Freedom with air strikes and a ground invasion to remove President Saddam Hussein from office. After a hard-fought campaign, U.S. forces reach the outskirts of Bagdad on 3 April 2003, and on 9 April 2003, Baghdad is secured by U.S. forces. While an insurgency is initiated in Iraq, President Bush declared end of combat operations on 1 May [9].

12.2.3.2 Doctrine Situation

In Iraqi Freedom, U.S. doctrine was similar to the situation in Allied Force. The main activity was Combat Assessment (CA), which was defined as consisting of BDA, Munition Effectiveness Analysis (MEA), and Reattack Recommendations (RR) [10].

Conceptually, BDA was still divided into phases, phase I being physical damage assessment, phase II functional damage assessment, and phase III target system assessment [10].

In addition, there was also some influence from Effects-Based Operations (EBO) and specifically Effects-Based Assessment (EBA), which was a developing concept at the time. Success Indicators were used to assess the achievement of operational objectives, and Measures Of Effectiveness (MOEs) were used to measure the achievement of tactical objectives [10].

12.2.3.3 Experiences

"*BDA is broken.*" (Gen. Moseley, quoted in Ref. [10].)

Assessment of the air campaign was mainly conducted by the Operational Assessment Team (OAT) at the Joint Force Air Component Command (JFACC) [10].

Data problems appeared immediately when the operation commenced. Further, organisational friction, other intelligence priorities (or targeting and intelligence being prioritized separately), reporting problems, and information management all hampered assessment. These frictions are similar to those that appeared in Desert Storm, but in addition, assessment was made more difficult by the dynamic battle and mobile targets that kept assessment efforts out of step. In addition, Weapons System Video was no longer widely available, and the doctrine proved to be untested. The culture in the armed forces did not emphasise assessment, and training for assessment was in short supply, especially outside the OAT [10].

Again, all these factors combined rendered assessment largely ineffective [10].

12.2.4 ISAF

12.2.4.1 Background

As a direct response to the Al-Qaeda attacks of 11 September 2001, U.S. forces launched Operation Enduring Freedom in Afghanistan on 2 October. Kabul fell later the same month. ISAF is initiated in December, originally as a stabilising force around Kabul, and is a coalition force initially under British leadership. In August 2003, NATO takes command of ISAF, which is then gradually expanded to cover all of Afghanistan. In 2012, a process of turning over control to Afghan forces is initiated, and in 2014, ISAF hands over operations to Resolute Support Mission [7], [11].

The following discussion of operations assessment in ISAF mainly covers the period from 2008–2012.

12.2.4.2 Doctrine Situation

Assessment doctrine in ISAF was dominated by EBA and results-based management, the main constructs used for assessment being MOEs and Measures Of Performance (MOPs) [12]. While significant doubts had been raised about the efficacy of effects-based operations and EBA [13], doctrine in the U.S. and NATO had mainly adapted by dropping the term "effects-based", but retaining the results-based methods favoured by EBA [12], [14].

12.2.4.3 Experiences

"*Our metrics suck.*" [12]

Connable recognized that the counterinsurgency environment in ISAF provided big challenges for collecting all the data that are required for conducting EBA. This in turn led to problems with accuracy of the assessment. EBA also inferred transparency problems, meaning that there were difficulties in discerning what data the assessment was based on. There were also tendencies of whitewashing or doctoring assessments in order to look successful. Finally, the threshold set for MOEs and MOPs started to affect behaviour in such ways as to design activities to target the threshold and not the intended effect [12].

12.3 SUMMARY OF CASE STUDIES

- None of the approaches used were deemed successful by users.
- Phase III BDA (system assessment) was never even attempted.
- EBA did not help in improving assessments and had detrimental side-effects.

Thus systematic feedback from operations largely failed, and staffs were able to provide little tangible assessment support to commanders. Instead, commanders had to mainly rely on intuition with relatively little analytic underpinning.

12.4 EXAMINING THE FAILURE OF OPERATIONS ASSESSMENT

This section is intended to explore possible explanations for the seeming failure of operations assessment. I look for explanations both in theoretical approaches, in the cognitive area, and in organisational issues.

Bernt Brehmer has identified four requirements for effective decision making [15]:

1) There has to be a goal (the goal condition).
2) It has to be possible to determine the state of the system (the observability condition).

3) It has to be possible to affect the system (the action condition).

4) There has to be a model of the system (the model condition).

Current military planning models have several layers of objectives or goals, but the end state should be sufficient to fulfil the requirement for the goal condition. The system in the context of Brehmer's requirements could be either the enemy fighting system, or the whole conflict system.

These requirements are used in the following discussion of the cases. Operations assessment is mainly intended to provide *observability*, and to assist in creating a *model* of the system. Operations assessment should also include suggestions for improving the effectiveness of the campaign [16], [1], thus operations assessment is also intended to assist in affecting the system in the most effective manner possible.

12.4.1 Theoretical Problems

Assessment doctrine stresses that operations take place in a complex operational environment [2], and that complexity means that inferring causality is extremely challenging [16], [1]. Despite these observations, these publications do not refer to approaches that are based on the academic study of complexity.

Much of the current science of complexity is centred on the concept of Complex Adaptive Systems (CAS) [17], [18], [19], [20]. Some of the problems experienced with assessment can be explained by assuming that conflict is a Complex Adaptive System.

- A CAS consists a large number of elements that in themselves can be fairly simple [17], [18].
- The interaction between CAS elements is characterised by feedback loops and non-linearity [17].
- In a CAS, elements mainly act on local information [17]. This infers a challenge to the *observability* of the CAS.
- In order to analyse a CAS, it is rarely sufficient to analyse the system in parts in order to reach conclusions on the system level. What is instead required is taking a "crude look at the whole" [19].
- CAS are incompressible, i.e., it is not possible to have an accurate description of a complex system that is simpler than the system itself [17], [18], [20]. Incompressibility infers a huge challenge to the ability to *model* a CAS. Incompressibility also challenges the possibility of conducting system assessment (Phase III of BDA).

The results-based methods used in EBA assume that a system can be decomposed into parts that can be studied in isolation and that interact linearly. Further, they assume that a complete model of the system is possible [21]. As we have seen above, none of these assumptions hold for CAS. Results-based methods also suffer from problems with validity [22] and "gaming", i.e., maximising outcomes according to the measurements rather than the goals, or outright falsification [23], [22].

Thus, if we acknowledge that complexity permeates the conduct of war, the methods that have been tried so far make unwarranted assumptions on the observability and modelling of conflicts, which would imply that putting more effort into the methods that have been used so far is unlikely to improve the situation.

12.4.2 Cognitive Problems

A rich field of psychological research looks into how human cognitive abilities are affected by bias in the human psyche. Bias may affect human ability or proclivity for receiving contrary information.

A well-known phenomenon is confirmation bias, the tendency to give greater weight to information that seems to confirm one's beliefs, and a lesser weight to information that seems to disconfirm one's

beliefs [24]. Confirmation bias clearly has the potential to distort assessment, both for assessment staff and commanders.

Anchoring bias is basing further judgements on some known entity in such a way as bias the judgement towards the anchor. The anchor may be modified, but insufficiently so when compared to real data on the situation. [25]. Viewing conflict as a CAS means that only limited and spotty information will be available for operations assessment, thus anchoring bias means that there is a risk that too much value is placed on the information that happens to be available.

Prospect theory indicates a tendency among humans to be risk averse, a loss of a certain amount of money, for instance, is given a higher valuation than the gain of the same amount of money. This indicates a tendency to base decisions on the status quo. By taking the same situation and changing its formulation from a loss to a gain, different preferences emerge. This is known as framing bias, where preferences are affected by how a certain situation is framed [24], [26]. The most obvious example of frames that relates to operations assessment are the set of goals (end states, objectives, decisive conditions, and effects) that military planning creates and that is used as gauges in operations assessment. Exceeding an objective may not be assigned any value at all.

As seen above, commanders have had to mainly rely on their own intuition in order to assess the situation. There is, however, known limits to the value of intuition. Ref. [27] identifies intuition with memory; in order to have effective intuition about something one has to have experienced something similar, at least in training. Intuition thus only works in situations where there are enough regularities for predictions to be made from patterns (even if those predictions are intuitive). Ref. [27] gives firefighters as example of a profession were intuition seems to work, and stockbrokers as an example of a profession where intuition does not work (no one reliably beats the market). Most situations fall somewhere in between, were certain tasks allow effective prediction and others do not, which Kahneman and Klein calls “fractionation” [27]. Further, most people are uncritical about their own intuition; humans are frequently unable to identify the situations in which they should be questioning their own intuition. In a situation of fractionation, behaviour that was successful for one task may be unsuccessful for another, and the feedback from the environment is not strong enough to make professionals question their own intuition. According to Ref. [27], warfare has these characteristics. The best moves may not ensure success, but they do increase the probability of success [27]. Thus trained, professional intuition is a necessary component of military command and control, but there should be scope for some questioning of intuition as well.

These observations of situations of fractionation are also congruent with viewing conflict as a CAS. Some prediction is possible, but it is not possible to always identify those situations beforehand.

12.4.3 Organisational Problems

Various organisational issues also reduce the ability for effective feedback. Operations assessment is essentially an evaluative activity, as operations assessment mainly aims to judge “how well are we doing”. Commanders and staffs, while requiring feedback on the development of the campaign, are also evaluated by the results achieved. Thus, bad news from the front may be resisted by decision makers at various stages of the process [12], [28].

Group-think involves a social process were group members reduce their critical abilities in order to be accepted in a group. This phenomenon assures unity of effort, even for an ill-conceived plan. Group-think may also lead to an escalation of commitment to a course of action that is not working, because reversing policy may look bad. Group-think introduces confirmation bias on the social (rather than individual) level [29], [30].

Assessment can also be seen as an organisational learning activity. Organisations constantly create mental models of their environments, which simplifies the situation and makes prediction possible. These

mental models tend to ossify, and the organisation may resist feedback from the environment in order to preserve the model. Further, effective learning in a part of an organisation (on the tactical level, for instance) may reduce the need for learning in another part of an organisation (on strategic management level, for instance) [31].

All these mechanisms tend to make feedback more difficult in an organisational setting. Feedback is not completely cut off, but it means that receiving effective feedback requires a focused effort.

12.5 CONCLUSION

I am not in a position to suggest assessment approaches that have been tested in the field and found to work, but it seems reasonable to suggest some approaches for operations assessment that show some promise, and that do not have the specific problems identified in previously used approaches.

One point of departure for the discussion of a command and control approach that addresses complexity is the so-called Cynefin framework [32]. In this framework, situations are divided into the *known* domain, where cause and effect relationships are generally known and predictable; the *knowable* domain, where cause and effect relationships may be knowable, but are commonly affected by delays; the *complex* domain, where there may be cause and effect relationships, but they are so intertwined and affected by feedback that reliable prediction becomes challenging; and the *chaotic* domain, where there is no perceivable cause and effect relationships and the system is turbulent. According to the Cynefin framework, the mode of decision making in the known domain is sense-categorise-respond, in the knowable domain it is sense-analyse-respond, in the complex domain it is probe-sense-respond, while in the chaotic domain it is act-sense-respond. Most current doctrine identifies warfighting as belonging in the complex or chaotic domains [33], which would imply that it would be relevant to explore the complex and chaotic domains further.

Developmental evaluation was devised by Patton for giving organisations that need to adapt to constantly changing environments access to continuous evaluative feedback [34]. Based on concepts from complexity theory, developmental evaluation advocates no specific methodical approach, but suggests several different so-called inquiry frameworks that are intended for different types of feedback requirements and situations. Several of those frameworks have the potential of being useful for the assessment problem.

As seen above, incompressibility makes modelling of CAS challenging, and rarely available to support decision making. By relaxing the requirements on generalisability, the model approaches a so-called requisite model [35], which is a model that is sufficient for solving a particular problem, without necessarily being usable in other situations. A requisite model is constructed by stakeholders in a facilitated process. Basically a requisite model may only be relevant for a specific decision, and should require updating and/or redesign as a campaign moves along.

All of the approaches suggested above move away from trying to "measure" progress, towards larger inputs from using judgement for assessment. This idea is further explored in Chapter 5 in this volume, "Assessment and Interpretation" [22].

Finally there is scope for revisiting *who* does the assessment. In using judgement, as in qualitative research, *who* does the judging is crucial [36], [37]. Recruitment and training of assessment staff will have to be revised if there is to be any scope for improving operations assessment. ORSA staff trained in quantitative approaches should be complemented with ORSA with knowledge of "soft" methods, military officers making professional judgements, and subject matter experts that understand the local situation. I would suggest that operations assessment needs to be more like all-source intelligence work rather than something like auditing the books.

12.6 AUTHOR'S BIOGRAPHY

Jan Frelin is an Operations Researcher (OR) currently working at the Policy and Plans department of the Swedish Defence Staff. Previous assignments include OR positions at the Swedish Joint Command, the Nordic Battlegroup, the Swedish Command and Control Command, Swedish Ground Forces Command and at KFOR Headquarters in Kosovo. Frelin has also conducted evaluations of the Swedish Provincial Reconstruction Team (PRT) in Afghanistan as well as contributing to a NATO evaluation of ISAF. With a background in computer science, Frelin initially worked with Modelling and Simulation (M&S), but is currently focused on so-called 'soft' OR, current operations, and evaluative activities.

12.7 REFERENCES

[1] NATO. 2014. NATO Operations Assessment Handbook (Draft) Version 3.0. Norfolk, VA & Mons, Belgium: NATO.

[2] Joint Publication 5-0. 2017. Joint Planning (JP 5-0). Joint Chiefs of Staff.

[3] Wetterqvist, F. and Johansson, L. 1991. Kriget vid Persiska viken 1990–91. Förlopp, lärdomar och slutsatser [The Persian Gulf War 1990–91. Course of events, lessons and conclusions] FOA A10026-1.4. Stockholm, Sweden: Försvarets forskningsanstalt.

[4] Bailey, T.M. 2001. Assess for Success: The Role of Doctrine in Effective Combat Assessment. Master's thesis: Air University.

[5] Janiczek, R.M. 2002. Combat Assessment in MEF Battlespace Shaping. Master's thesis: United States Marine Corps Command and Staff College.

[6] Curry, H. 2004. The current battle damage assessment paradigm is obsolete. Air & Space Power Journal, winter. http://www.airpower.maxwell.af.mil/airchronicles/apj/apj04/win04/vorwin04.html.

[7] Lackenbauer, H. and Langlais, R. (editors). 2013. Review of the Practical Implications of UNSCR 1325 for the Conduct of NATO-led Operations and Missions. 10-2013-EN. NATO. http://www.nato.int/nato_static/assets/pdf/pdf_2013_10/20131021_131023-UNSCR1325-review-final.pdf.

[8] Ljung, B. 2001. Aspects of the Kosovo Operation March–June 1999. FOI-R–0070–SE. Stockholm, Sweden: Swedish Defence Research Agency.

[9] PBS. 2004. Operation Iraqi Freedom. http://www.pbs.org/wgbh/pages/frontline/shows/invasion/cron/.

[10] Allen, N.T. 2005. Effects-Based Assessment in United States Air Force: Rhetoric or Reality? Master's thesis: Air University.

[11] NATO. 2015. Resolute Support Mission in Afghanistan. Retrieved from http://www.nato.int/cps/en/natohq/topics_113694.htm.

[12] Connable. B. 2012. Embracing the Fog of War. Assessment and Metrics in Counterinsurgency. Santa Monica, CA: RAND.

[13] Mattis, J.N. 2008. Assessment of Effects-Based Operations. Unpublished manuscript, August 14.

[14] Frelin, J. and Norén, A. 2012. Recent Developments in Evaluation & Conflict Analysis. Tools for Understanding Complex Conflicts. FOI-R–3406–SE. Stockholm, Sweden: Swedish Defence Research Agency.

[15] Brehmer, B. 1992. Dynamic decision-making: Human control of complex systems. Acta Psychologica, 81:211-241.

[16] NATO. 2011. Operations Assessment Handbook. SHCPPSPL/3055-18/11-280097, Interim version 1.0. Norfolk, VA and Mons, Belgium: NATO.

[17] Cilliers, P. 1998. Complexity and Postmodernism. London and New York: Routledge.

[18] Cilliers, P. 2000. What Can We Learn from a Theory of Complexity? Emergence, 2(1):23-33.

[19] Gell-Mann, M. 1994. Complex Adaptive Systems. G. Cowan, D. Pines & D. Meltzer (Eds.) Complexity: Metaphors, Models and Reality. SFI Studies in the Sciences of Complexity, Proceedings Vol. XIX. Addison-Wesley.

[20] Richardson, K.A., Mathieson, G.L., and Cilliers, P., (2009). Complexity Thinking and Military Operational Analysis. In: K.A. Richardson, (Ed.) Knots, Lace and Tartan. Making Sense of Complex Human Systems in Military Operations Research. Litchfield Park, AZ: ISCE Publishing. pp. 27-69.

[21] Bullock, R. and Deckro, R. 2006. Foundations for system measurement. Measurement, 39:701-709.

[22] Frelin, J. 2013. A Complex Failure: The Introduction of Performance Measurement in the Public Sector. Paper presented at SOAK/NOS6, November. Gothenburg, Sweden.

[23] Bevan, G. and Hood, C. 2006. What's Measured is What Matters: Targets and Gaming in the English Public Health Care System. Public Administration, 84(3):517-538.

[24] Hogarth, R. 1994. Judgement and Choice. The Psychology of Decision, 2nd edition. Chichester, UK: John Wiley and Sons.

[25] Tversky, A. and Kahnemann, D. 1974. Judgement under Uncertainty: Heuristics and Biases. Science, 27 Sep, 185(4157):1124-1131.

[26] Tversky, A. and Kahnemann, D. 1981. The Framing of Decisions and the Psychology of Choice. Science, 30 Jan, 211(4481):453-458.

[27] Kahnemann, D. and Klein, G. 2009. Conditions for Intuitive Expertise. American Psychologist, Sept, 64(6):515-526. doi:10.1037/a0016755.

[28] Vedung, E. 2009. Utvärdering i politik och förvaltning [Evaluation in policy and public administration]. Lund, Sweden: Studentlitteratur.

[29] Goodwin, P. and Wright, G. 2009. Decision Analysis for Management Judgment; 4th edition. Wiley.

[30] Newell, B.R., Lagnado, D.A. and Shanks, D.R. 2007. Straight Choices: The psychology of decision making. Psychology Press.

[31] Levinthal, D.A. and March, J.G. 1993. The Myopia of Learning. Strategic Management Journal, winter, 14:95-112.

[32] Kurtz, C.F. and Snowden, D.J. 2003. The new dynamics of strategy: Sense-making in a complex and complicated world. IBM Systems Journal, 42(3):462-483. doi:10.1147/sj.423.0462.

[33] Joint Publication 5-0. 2011. Joint Operation Planning (JP 5-0). Joint Chiefs of Staff.

[34] Patton, M.Q. 2011. Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use; 1st edition. New York, NY: Guilford Press.

[35] Phillips, L.D. 1984. A theory of requisite decision models. Acta Psychologica, 56:29-48.

[36] NATO. 2012. NATO Guide for Judgement-Based Operational Analysis in Defence Decision Making; Analyst-Oriented Volume: Code of Best Practice for 'Soft' Operational Analysis. AC/323 (SAS-087)TP/345).NATO.

[37] Patton, M.Q. 2008. Utilization Focused Evaluation; 4th edition. Thousand Oaks, CA: Sage.

Chapter 13 – CONNECTING OUTCOMES TO INDICATORS: A MODEL FOR SELECTING INDICATORS

Adam Shilling
Centre for Army Analysis
UNITED STATES

ABSTRACT

The following model is a simple approach to identifying the set of most important indicators for an assessment process. It originated in a conversation in 2015 with the writing team responsible for the update of the capstone United States' Joint Publication (JP) 5-0 (2017) Joint Planning which provides the doctrine for conducting joint planning activities for all military operations. They asked if I had a method for identifying indicators. I said yes, and they asked if I could draw it on a piece of paper. After many attempts, I came up with version one of the model that follows. The team included in Joint Publication 5-0 (2017) a derivation of version one of the model, but this chapter contains an update of the model and an improved explanation of how to use it.

A number of people from various organizations, with different mission sets, have found it useful. It is not based in a particular branch of science. Rather, it is just a common-sense way of breaking down the problem that planners and assessors face as they design the plans that seek to effect change in an operational environment and the processes that assess these plans' execution.

13.1 INTRODUCTION

The model begins by assisting planners in breaking down, iteratively, the endstate that we are trying to achieve into more specific statements that can effectively focus and guide the operations of subordinate organizations. In this way, the model can assist with the design process as we determine the best way to proceed toward the set of desired outcomes [1]. The model identifies relevant indicators for collection, but in practice, I have learned that more importantly, it also *identifies and records the logical links* between the outcomes we seek and the indicators we use to gauge our effectiveness.

Therefore, we become more likely to understand the causal relationships between the operations we pursue and the change in the Operational Environment (OE) that is reflected in our indicators. *Any operational plan that orders a set of activities in pursuit of a set of outcomes is a set of causal hypotheses – that these activities will cause the changes in the OE which are specified as outcomes*. Assessment can be viewed as a testing of these hypotheses. We also better understand the relationships between related indicators. This makes assessment processes more effective, which in turn, makes our operations more effective.

To use the model, we begin with the least specific description of the changes in the OE that we want to effect. At high levels of command, we usually call this an "endstate", but we can start the model with any other outcome statement that is pertinent to our organization given its place in the military hierarchy. For example, a small tactical unit may be assigned a tactical task, and the purpose of that task may be the outcome statement their assessment process begins with.

The term "outcome" may mean any of the following: an endstate, an objective, a sub-objective, an effect, a task/purpose, a condition, a success or termination criterion – or anything else that specifies the change(s) in the operating environment that we are seeking.

13.2 MODEL OVERVIEW

Figure 13-1 shows the complete model. To begin, we evaluate the first outcome statement with Question 1 (Q1): "Is this outcome statement specific enough to guide the activities of subordinate organizations or does it need to be broken up?" If the statement is specific, we can proceed through the assessment loop of the model and begin discovering and recording the links between the outcome we desire and the indicators that will demonstrate our success (or failure). If we judge the outcome is too vague to guide and focus action, we iterate through the design loop of the model by breaking down the less-specific statement into its constituent parts.

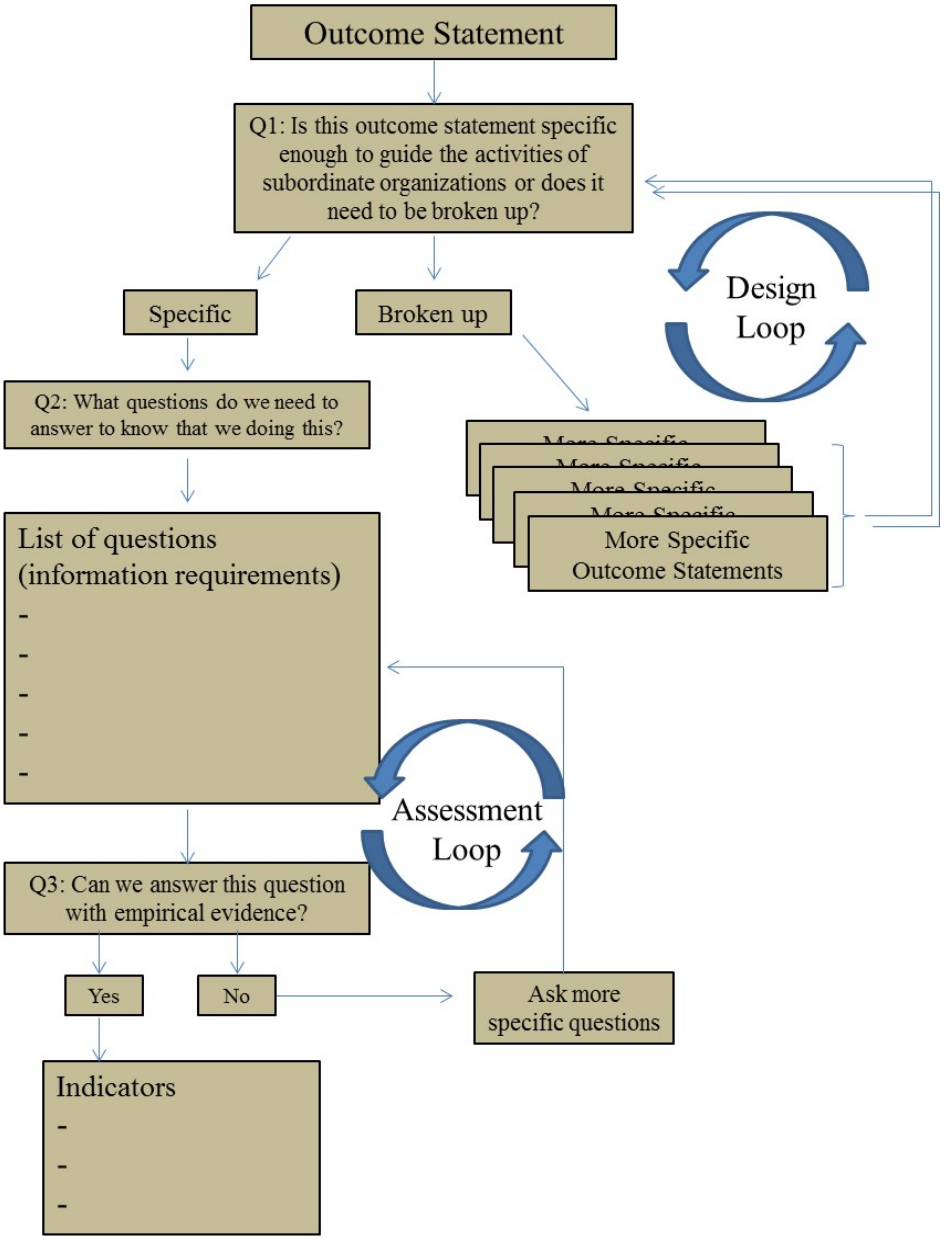


Figure 13-1: The Compete Model.

As we articulate each of the more specific outcome statements, we test each of these with Q1. When each is off sufficient specificity, we can proceed. Until then, we iterate as many times as is useful through the design loop, articulating more specific outcomes each time.

Therefore, if we start with an endstate, the first iteration of the design loop may yield a set of termination or success criteria or may suggest lines of operation or lines of effort. A second iteration may suggest objectives, and a third may suggest effects, although the names we give to these levels of specificity probably do not matter yet.

Once we judge each outcome to be sufficiently specific, we proceed to the assessment loop and evaluate each outcome with Question 2 (Q2): "What questions do we need to answer to know that we are doing this?" We make a list of questions that we might call formally "Information Requirements" (IRs). *These questions provide context and focus to the subsequent list of indicators and provide the logical links between outcome and indicator that helps us to make the case that friendly actions did (or did not) cause the changes in the OE that are reflected in indicators*.

Next, we evaluate each IR with Question 3 (Q3): "Can we answer this question with empirical evidence?" If the answer is "no" we ask more specific questions. If the answer is "yes" we record the information that answers the questions as a potential indicator.

When this process is complete, we can prioritize the questions, and assign collection assets to answer the most important ones.

13.3 SOME TIPS FROM EXPERIENCE (LESSONS LEARNED)

One key to using the model is to use it *loosely*. It is really a linear representation of a more complex thought process that is designed to help assessors get started. As we use the model, we will gain mastery of the process, and will perceive it as non-linear, very elastic, and malleable to the demands of the problem. The model provides an unconstrained look at the problem. It starts with a blank sheet of paper. It has no blanks or cells to fill in. The loops permit and require as many iterations as we need and no more.

I envision the model's product as a loose outline that connects less-specific to more specific outcomes, questions which document the logical links between desired outcomes and the empirical evidence of our achievement of them, and finally the empirical evidence of the change we are seeking. We can impose military hierarchy in terms of the "names" or "labels" we put on elements of the logical chain – such as line of effort, objective, effect, information requirement, and indicator, or something else – after our unconstrained examination of the problem, in a way that makes sense for communicating the commander's desires to those who will execute them.

My experience in teaching the model to assessors is that we frequently attempt to skip the list of questions (IRs), and go straight to the list of indicators. When we do, we typically develop a set of all possible indicators that have bearing on the desired outcomes. This creates several problems:

- The list of indicators is too long, lacks focus, and is hard to prioritize.
- Assessors do not understand the relationship of indicators to each other.
- The command wastes collection assets gathering information it does not need.

The formal step of listing IRs (questions) and then identifying indicators which answer them helps assessors understand the relationship between several related indicators that may answer related questions. Also, most people find the list of questions easier to prioritize than an exhaustive list of potential indicators. In essence, the formality of posing the questions focuses the identification of indicators. Also, a piece of information (an indicator) is only valuable if it answers an important question (an IR).

Questions (IRs) and answers (indicators) need not have a one-to-one correspondence. One indicator may answer several questions; one question may require several indicators; or several related questions may be answered by several related indicators.

Assessment is largely about answering questions about our activities, our OE, and our interaction with it. Once we answer the key questions, we understand the OE better, and can make intelligent decisions to increase our effectiveness. Increasing our effectiveness is the purpose of this model.

13.4 AN EXAMPLE OF USING THE MODEL

Figure 13-2 to Figure 13-13 show an example of using the model to work from a blank sheet of paper (or blank word processor document!) to a fleshed out example. As we work through the model, the details are added to the same sheet of paper. As per Figure 13-2, we start with a blank sheet of paper, insert the endstate at the top of the page.

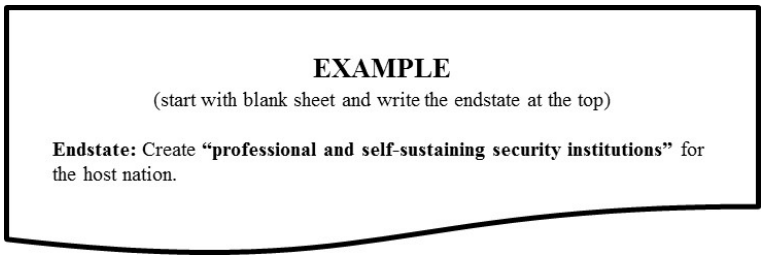


Figure 13-2: Begin with the Endstate.

Assessors, especially at lower levels of command, may also start from an outcome statement, other than an endstate, that is relevant to their unit.

Figure 13-3 shows the beginning portion of the model – the design loop – starting with the endstate. A security cooperation headquarters abroad wanted to aid the partner nation in creating "professional and self-sustaining security institutions", so we write that endstate at the top of the page.

Looking at Figure 13-3, the first step is to evaluate our first outcome statement, in this case, the endstate, with question one (Q1), "Is this outcome statement specific enough to guide the activities of subordinate organizations or does it need to be broken up?" If the statement is specific, we can proceed through the model. If not, we need to break the statement into more specific statements.

In this case, we decide it is useful to break up the endstate into more specific statements. A more specific outcome is any statement that increases the specificity of the outcome desired. For example, for an endstate, we may have several objectives; for an objective, we may have several effects; or for an endstate, we may have several success or termination criteria. At this point, an assessor is helping with design and planning by helping the planners get the outcome statements right to focus, guide, and perhaps restrain subordinates' actions.

We conclude that achieving the five things listed in Figure 13-3 are sufficient to accomplish our endstate. Now our piece of paper looks like Figure 13-4.

The next step is to evaluate each more specific outcome statement with Q1. In real life, we will do this for each of the new outcome statements. For this example, we will do only the one in red type. We will iterate through the design loop as many times as is necessary, and proceed through the model once we have outcome statements we judge to be sufficiently specific.

Going back the model, in Figure 13-5, we evaluate the statement, "security forces are properly trained" with Q1. We conclude it requires more breakdown, because we are not sure what "properly trained" means. After some critical thought, we write several more outcomes that, if achieved, are sufficient (in this example) to conclude our partner forces are properly trained. Now our sheet of paper looks like Figure 13-6.

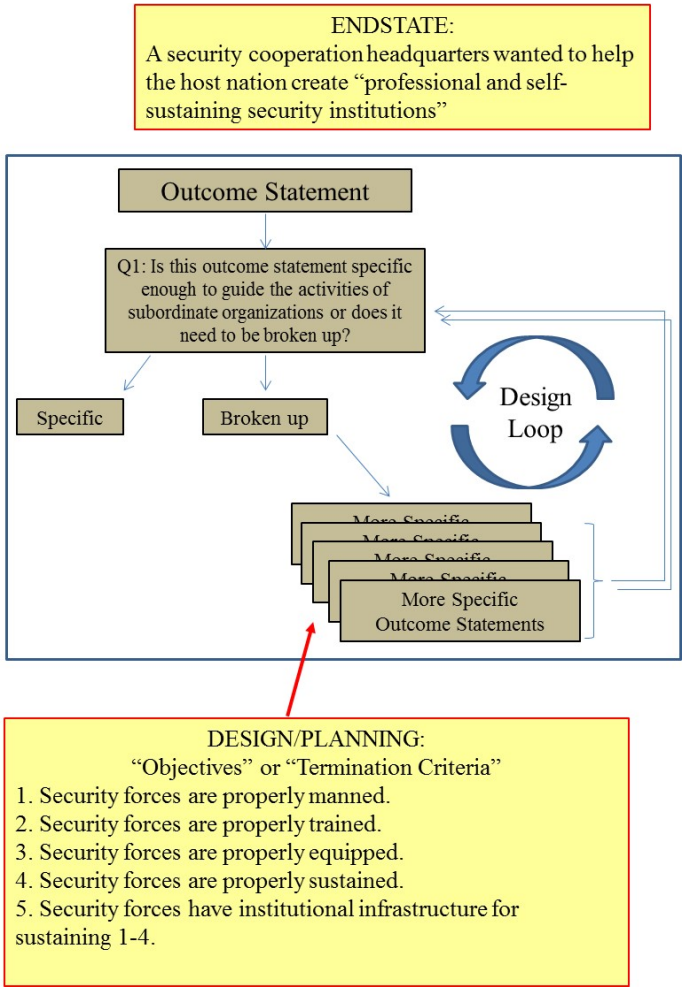


Figure 13-3: One Iteration Through the Design Loop.

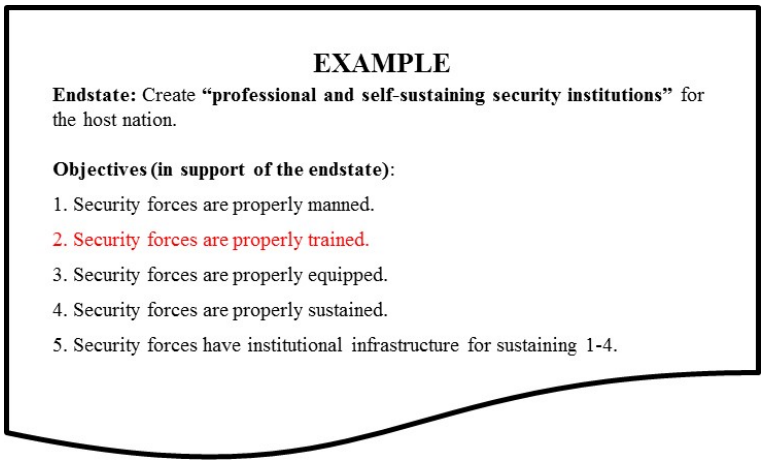


Figure 13-4: Breaking Down Desired Outcomes in Design Loop (First Iteration).

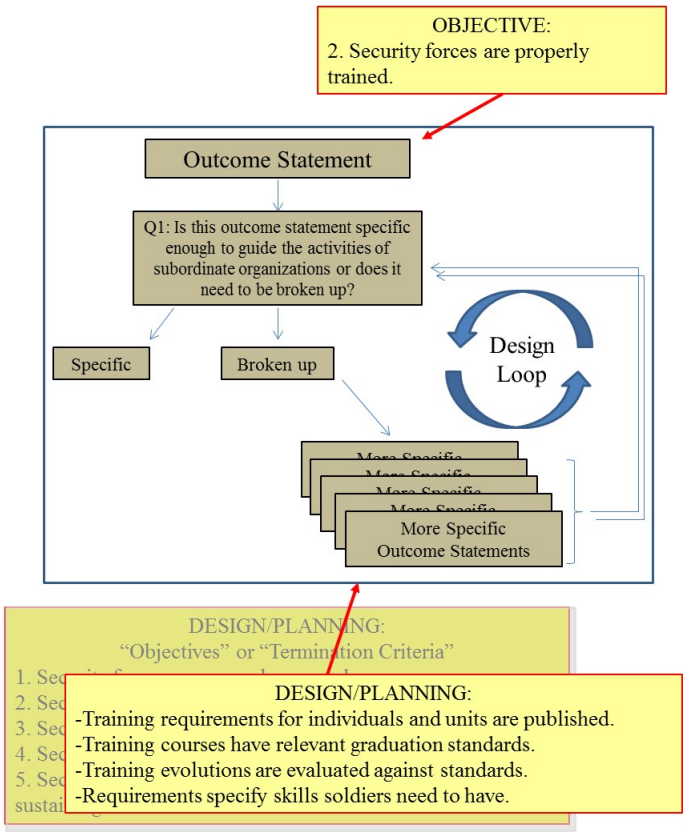


Figure 13-5: Two Iterations Through the Design Loop.

EXAMPLE

Endstate: Create **"professional and self-sustaining security institutions"** for the host nation.

Objectives (in support of the endstate):

1. Security forces are properly manned.
2. Security forces are properly trained.
   - -Training requirements for individuals and units are published.
   - -Training courses have relevant graduation standards.
   - -Training evolutions are evaluated against standards.
   - -Requirements specify skills soldiers need to have.
3. Security forces are properly equipped.
4. Security forces are properly sustained.
5. Security forces have institutional infrastructure for sustaining 1-4.

Figure 13-6: Breaking Down Desired Outcomes in Design Loop (Second Iteration).

In real life, we evaluate each of the outcomes by Q1. For this example, we do the one in red type. As in Figure 13-7, we conclude one more iteration through the design loop is useful. After some critical thought, we conclude that three specific requirements are necessary, and our paper now looks like Figure 13-8.

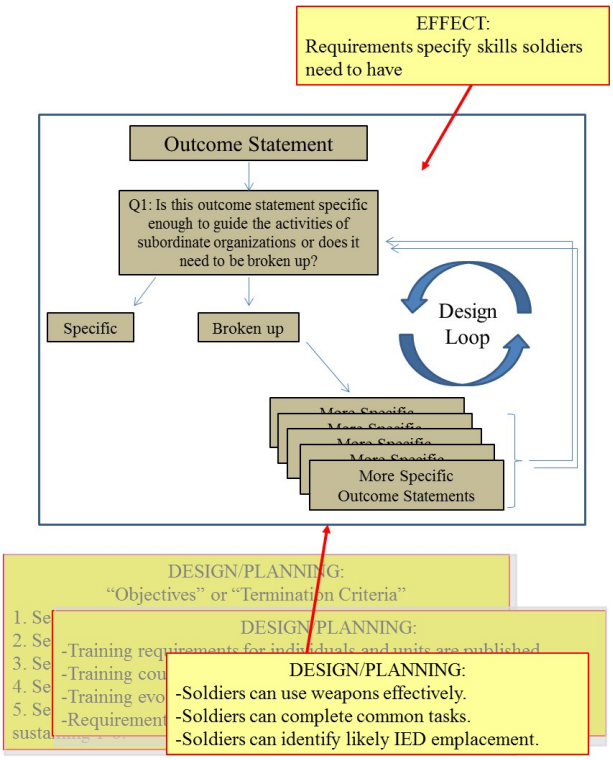


Figure 13-7: Three Iterations through the Design Loop.

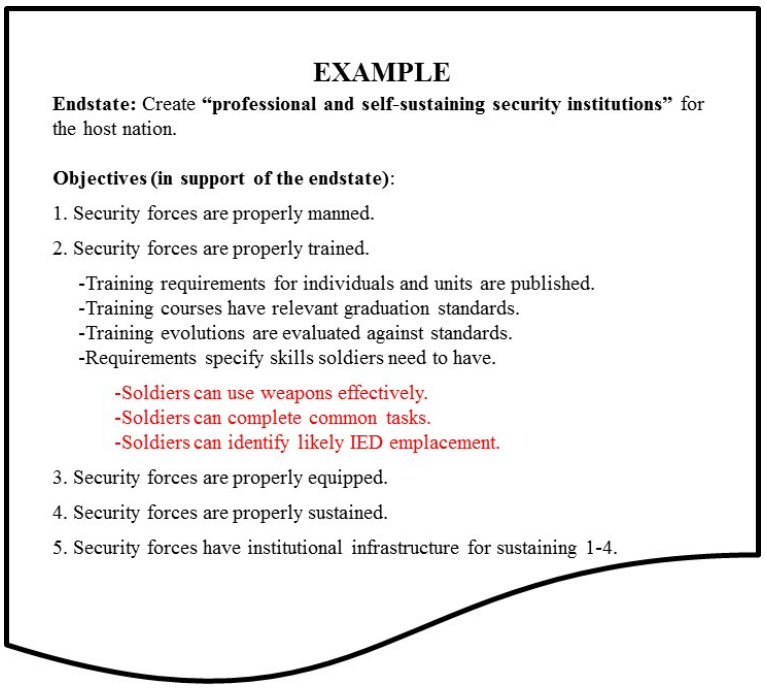


Figure 13-8: Breaking Down Desired Outcomes in Design Loop (Third Iteration).

Now we evaluate these three outcomes with Q1, and we conclude they are sufficiently specific, so that we can proceed through the model as in Figure 13-9.

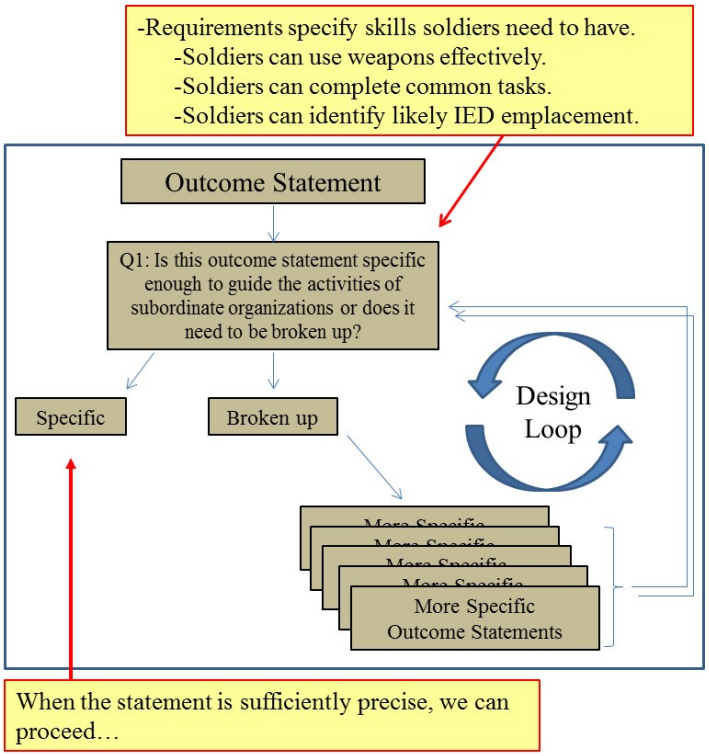


Figure 13-9: Complete Design Loop and Move to the Assessment Loop.

We begin the assessment loop by posing Question 2 (Q2) for each specific outcome statement. We want to determine the specific questions that need to be answered for us to determine how well we are accomplishing each specific outcome.

In this example, as seen in Figure 13-10, we identify one or more questions that need to be answered for each outcome. Our paper now looks like Figure 13-11.

We evaluate each of these questions with the model's Question 3 (Q3) as in Figure 13-12. If each question can be answered with empirical observation, we designate it an Information Requirement (IR), and begin looking for a way to answer it. If it cannot be answered with empirical observation, then we ask more specific questions with the goal of asking questions that can be answered empirically.

Once we have that list of questions, we have our IRs, and the information that answers them empirically are indicators. Now we prioritize our IRs and assign collection assets. Our paper now looks like Figure 13-13.

13.5 CONCLUSION

Once we flesh out the outline seen in Figure 13-13, we have recorded the logical connections between our desired outcomes and the indicators we collect and use to gauge our effectiveness.

Using the above model to illustrate the connections will assist assessors and planners determine specific objectives, effects, etc. for an endstate that will focus units on the things that need to be done to be most effective. It also helps assessors develop appropriate IRs and indicators for collection. By working through this model, the assessor can determine what information is needed for analysis and assessment.

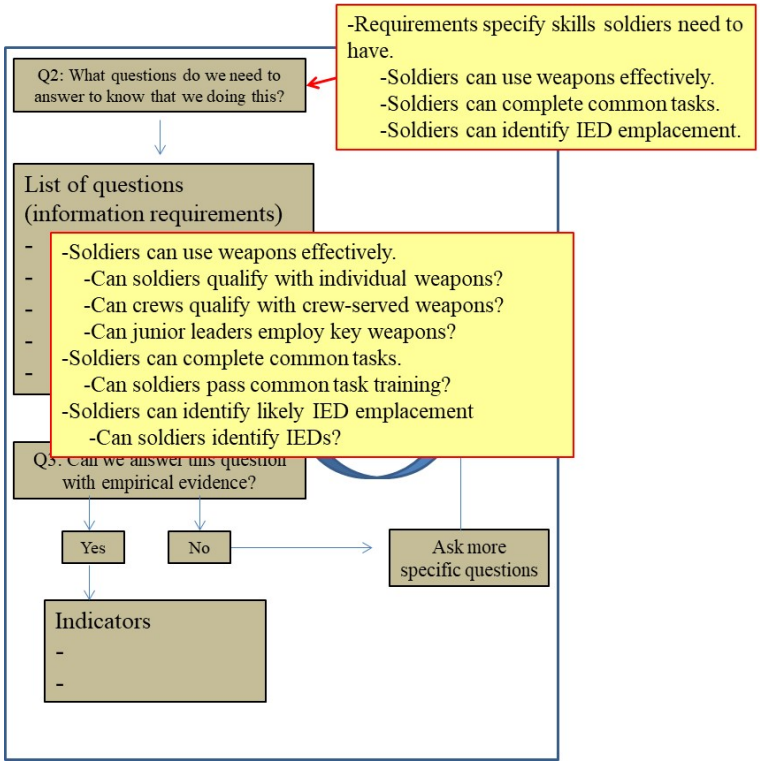


Figure 13-10: Determine Information Requirements.

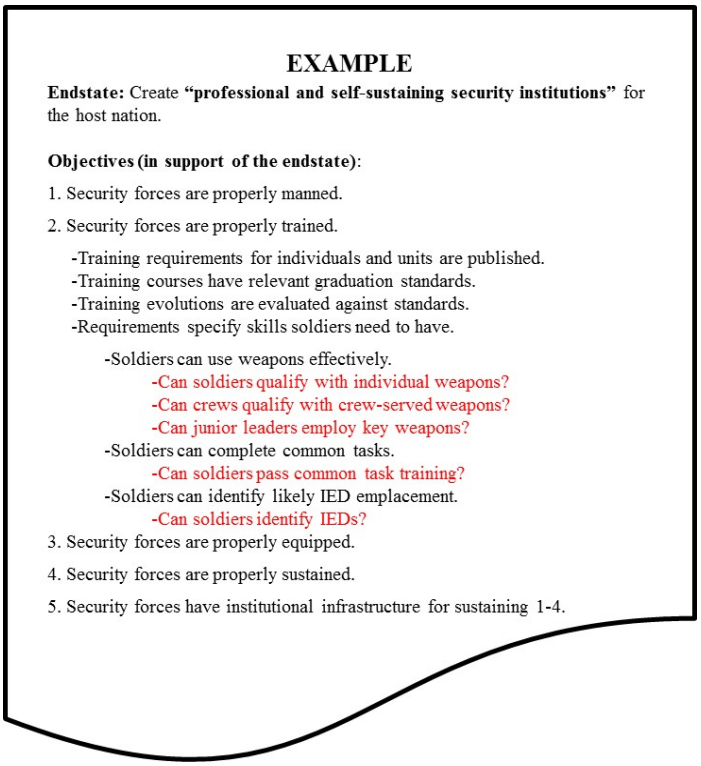


Figure 13-11: Document Information Requirements.

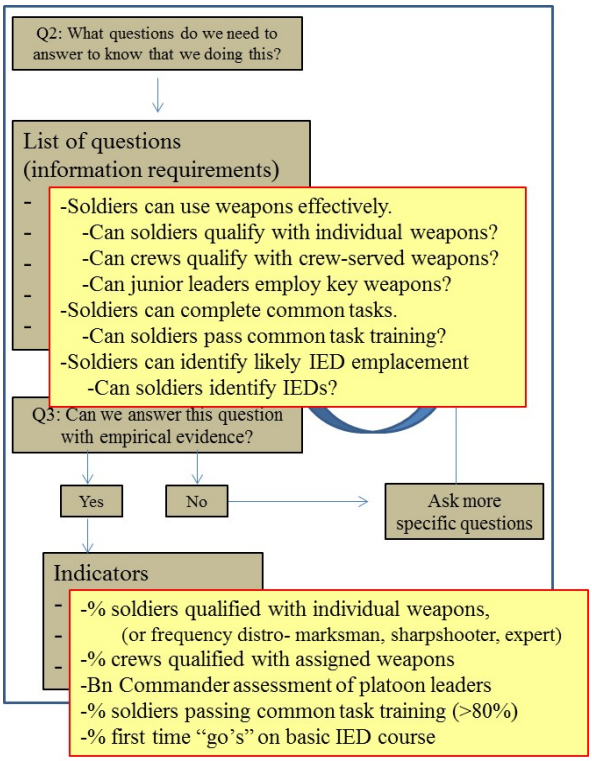


Figure 13-12: Determine Potential Indicators.

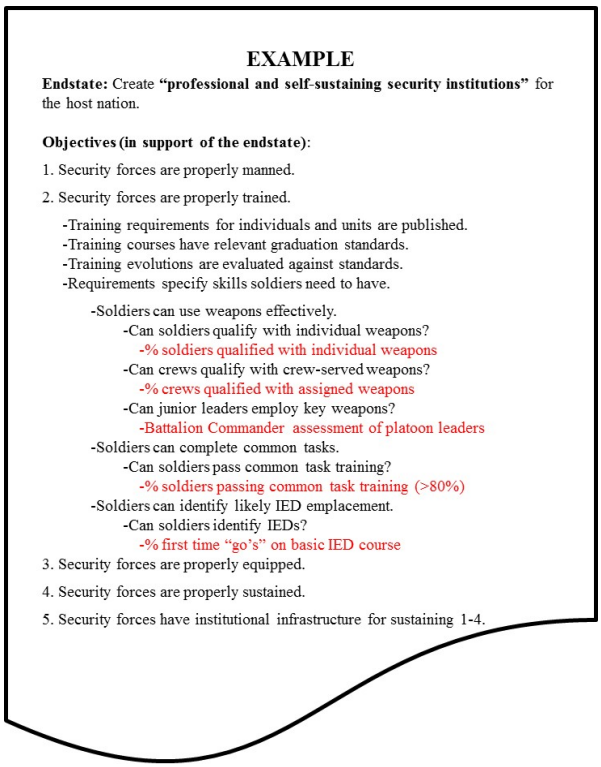


Figure 13-13: Connected! Outcomes to Indicators.

13.6 AUTHOR'S BIOGRAPHY

Adam Shilling, PhD, is a veteran of two operations assessment tours abroad – one to the Afghan Assessment Group at Headquarters, International Security Assistance Force (ISAF), and one at the Combined Joint Task Force – Horn of Africa. He spoke about his experiences and lessons learned at an international meeting. His ideas made sense to analysts and decision makers, and so his superiors selected him to represent his organization, the Centre for Army Analysis, in the community that was writing operations assessment doctrine for the United States military, and to act as senior editor for this report.

13.7 REFERENCES

[1] Army Techniques Publication (ATP). 5-0.1, Army Design Methodology. Washington D.C.: Headquarters, Department of the Army, July 2015.

Chapter 14 – ASSESSING COMPLEX OPERATIONS: A BRIEF LITERATURE REVIEW AND ANNOTATED BIBLIOGRAPHY

Ben Connable
RAND Corporation
UNITED STATES

ABSTRACT

This chapter provides an overview of the literature specifically focused on assessing complex operations and an annotated bibliography. Some of the literature on complex operations assessment is quite good. Collectively, the literature offers a wide menu of theories, methods, and case history lessons. But it is not unified and does not offer a cross-cutting process that would work in every complex operation; there is no defensibly replicable approach. Doctrine periodically reverses itself, bringing into question the generalizability and usefulness of its precepts. Perhaps the greatest benefit of this literature is that, with only a few exceptions, it does not shy away from pointing out the limits of knowledge in the maelstrom of complexity.

Most of our bibliographic selections focus solely on complex operations, while a few have broader purposes but are part of what we perceive to be the necessary canon. Many samples were drawn from existing, published NATO bibliographies. A comprehensive bibliography would be impossibly and unhelpfully long: this represents a sample of professional articles, books, and doctrinal publications chosen by the NATO TR SAS-110 team. It seeks to show both European and American work and a sample of differing opinions. Based on our survey of the literature we believe that more European input to the American debate over assessment is needed, and more American attention to European expertise is warranted.

14.1 INTRODUCTION

The driving purpose of the NATO TR SAS-110 research was to seek a general consensus on theories, guiding principles, and best practices of the operations assessment of complex operations, or *complex operations assessment*. Specific literature on complex operations, peacekeeping operations, and stability operations assessment is anaemic and inconsistent compared to the stable of scientifically grounded and field-proven literature on the technical aspects of conventional war assessment and the tactical applications of operations analysis. Nothing in the collective literature of complex operations assessment recommends it as a distinct field of study or even a consistent field of practice. Many field practitioners are – and will always be – novices with a seemingly thankless task suddenly thrust upon them by a policymaker or military commander. Nevertheless, the demands for complex operations assessment theory, principles, and best practices are probably higher now than at any point in recorded history.

Irregular Wars (IW) and instability remain a constant global phenomenon. Today, western states and militaries engaged in IW and peacekeeping operations have unprecedented access to conflict data. In large-scale irregular wars like Afghanistan, policymakers face unrelenting, real time pressure to show rapid success as military operations play out at a frustratingly languid pace. Policymakers turn to military leaders, who in turn demand assessments that will stand up to the rigor of public scrutiny. Assessment analysts turn to the literature on complex operations assessment that offers, for the most part, a drumbeat of complaints, a helpful but scattershot set of theories and tools, or straightforward and seemingly logical doctrine publications offering methods that, thus far, have not translated into optimal performance. The field is new, and the literature is scattered.

This chapter provides an overview of the literature specifically focused on assessing complex operations and an annotated bibliography. Some of the literature on complex operations assessment is quite good. Collectively, the literature offers a wide menu of theories, methods, and case history lessons. But it is not unified and does not offer a cross-cutting process that would work in every complex operation; there is no defensibly replicable approach. Doctrine periodically reverses itself, bringing into question the generalizability and usefulness of its precepts. Perhaps the greatest benefit of this literature is that, with only a few exceptions, it does not shy away from pointing out the limits of knowledge in the maelstrom of complexity.

14.2 SUMMARY OF THE LITERATURE: IT'S HARD, NO CONSENSUS, UNBALANCED FOCUS

While thousands of historians and strategists have written about the process and outcomes of wars for centuries, there was effectively no literature specific to warfare assessment through the Vietnam War.[1] Arguably, up to the middle 20th Century conventional wars offered fairly clear outcomes and irregular wars and other complex operations attracted relatively little public attention. Demands for scientific measures of progress in war emerged and matured through World War II as operations analysis methods showed what could be accomplished with technical operations assessment: more efficiency, more effectiveness, and perhaps faster victory. As the field of operations analysis evolved over the next half century it would become synonymous with operations assessment. Operations analysts would become the de facto experts in complex operations assessment, and also its most prolific examiners.

Modern war became fuzzier and harder to understand in the second half of the 20th Century. Military leaders felt pressure from political leaders and from the public to show incremental tactical progress towards victory beginning with the Korean War stalemate in the early 1950s.[2] Official U.S. assessment efforts in the Vietnam War set the effective start point for the complex operations assessment field. Thomas C. Thayer's reports on the Vietnam War, published in a 12-part series under *A Systems Analysis View of the Vietnam War, 1965–1972* represented the first substantive, transparent effort to assess a large-scale irregular war [1]. Thayer later summarized his exhaustive work in the landmark *War Without Fronts: The American Experience in Vietnam* [2]. Thayer's arguments were that it was exceptionally hard to understand what was happening in Vietnam; that the available (voluminous) data were of very poor quality; but that the application of scientific methods – and specifically operations analyses – could help cut through the fog of war.

Thayer's observations and his solution set a benchmark in the literature on complex operations assessment. According to Thayer, complex operations assessment can be successful with the right metrics, "good enough" data, and the right methods. While the author of this chapter pushed back against Thayer's conclusions, most of the literature on complex operations assessment accepts his premise and then pursues what has become a tangled mass of relatively generic approaches or effectively *sui generis* technical solutions.[3] Others push back but expend their efforts without offering sufficiently detailed alternatives (see below for several examples).

There is one consensus in the literature on complex operations assessment theory and methods: they are so complex that they all but defy consensus. Otherwise, the intermittent flurries of literature since the Vietnam War tangentially rehash or attack Thayer's arguments, offer sometimes-helpful idiosyncratic case study lessons, bemoan poor policymaking, and reiterate the inherent complexity of the problem. Some of the sources cited here offer specific guidance, including detailed theories, data sources, metrics, and ostensibly generalizable lessons. It possible that one or more of the cited authors has already published a cross-cutting

[1] In their brief 2014 history of operations assessment, Emily Mushen and Jonathan Schroden chose the Vietnam War as a starting point for the history of operations assessment [3].

[2] Certainly there was pressure to show progress before this, but the Korean War marked the turning point from a focus on strategic outcomes to tactical effects and, in turn, technical campaign assessment. See Ref. [4].

[3] For example: see Ref. [5].

theory and method. However, in scientific terms there has been insufficient time to replicate or validate any one approach: the vast majority of literature on complex operations assessment was published between 2009 and 2015. As this report goes to publication the field is still wide open. Worse, it is unbalanced.

At least since the early years of the Afghanistan campaign, European and American assessment experts have worked in close concert to better understand complex operations assessment. Academic experts in the field know and cite each other routinely. European and American military and government analysts have deployed together under a U.S. coalition flag in Iraq, under a NATO headquarters in Afghanistan, and they have long-established relationships both through NATO and between its offices and parts of the U.S. joint community. But much of the publicized literature on complex operations assessment is American, and – in the opinion of this American author – much of it amounts to navel gazing. American assessments of Vietnam, Iraq, and Afghanistan have been described and lamented so many times that they now generate disinterest. Some of the excellent Canadian and European articles cited here are known mostly in the small circles of assessment community; they deserve broader publicity. The NATO *Operations Assessment Handbook* [6] is hobbled by a restricted distribution caveat. Its impact on American assessment experts, and more importantly on their commanders, is uncertain but probably minimal. More European input to the American debate over assessment is needed, and more American attention to European expertise is warranted.

This report, and the relationships fostered during NATO TR SAS-110, are part of the ongoing effort to bridge this divide, to help unify NATO thinking on complex operations assessment, and to bring the community of experts closer to a generalizable, replicable, and scientifically defensible theory and method.

14.3 ANNOTATED BIBLIOGRAPHY

What follows is an annotated bibliography representing a sample of the literature on complex operations assessment. Most selections focus solely on complex operations, while a few have broader purposes but are part of what we perceive to be the necessary canon. Many samples were drawn from existing, published NATO bibliographies. A comprehensive bibliography would be impossibly and unhelpfully long: this represents a sample of professional articles, books, and doctrinal publications chosen by the NATO TR SAS-110 team. It seeks to show both European and American work and a sample of differing opinions. We offer our apologies to the authors of important works not selected here and recommend this only as the starting point for further reading.

14.3.1 *War Without Fronts: The American Experience in Vietnam* (1985) [2]

Thomas C. Thayer (Available Online? No, but see: *A Systems Analysis View of the Vietnam War*).

War Without Fronts stands out as the most in-depth and insightful report on complex operations assessment. Thayer was an operations analyst with in-country experience in Vietnam who directed assessment for the U.S. Secretary of Defence. His book reports his lessons from the Pentagon, where he grappled with the massive and incessant deluge of raw data and often misleading field analysis sent to his team by (primarily) the Military Assistance Command in Vietnam (MACV). Thayer's primary argument is that pattern and trend analysis, while imperfect, can effectively show policymakers what has happened and what is likely to happen in irregular war. He describes challenges with data, and presents selected results from over 50 reports his team published on the Vietnam War.

Reading Thayer – if one can find a copy of his now out-of-publication book – will encourage serious analysts to also read the entire *Systems Analysis View of the Vietnam War* series [1]. These twelve reports are all readily available online and they reveal the detail that Thayer summarizes in his book. Anyone

looking for explicit methods and examples will find a wide menu of choices in the Vietnam-era series. Thayer's original reports also stand out as a baseline for comparative analysis with more recent work. Despite many advances in information technology and analytic methods, the field has not progressed far beyond the Vietnam War.

14.3.2 "Measuring Effectiveness in Complex Operations" (2007) [7]

James Clancy and Chuck Crossett (Available Online? Yes – free).

Clancy and Crossett marked the leading edge of a brief series of irregular warfare assessment articles that would emerge from the Iraq War. If Thayer's work is a baseline for the Vietnam War, then "Measuring Effectiveness in Complex Operations" sets the baseline for the post-9/11 counterinsurgency era. It points out the gap in assessment theory and suggest that something should be done to improve understanding of complex operations. Citing historical examples, Clancy and Crossett suggest that case study can help analysts identify the best metrics to track and, consequently, the best ways to succeed in complex operations.

Analysts seeking new methods or specific examples of historical assessment should look elsewhere. The authors do more to identify gaps than offer solutions. Reading Clancy and Crossett in hindsight reminds us of just how little had been done to progress the field of complex operations assessment since the Vietnam War. Or perhaps it instead reveals how few of the lessons from the Vietnam War assessment efforts had permeated American military discourse.

14.3.3 *Measuring Progress in Afghanistan* (2009) [8]

David Kilcullen (Available Online? Yes – free).

David Kilcullen gained prominence as an advisor to David H. Petraeus in Iraq, but he became well-known in assessment circles for his work on metrics in Afghanistan. His impromptu article captured his take on irregular war assessment from experience in Southeast Asia and Iraq as applied to the American surge in Afghanistan in 2009. His published report paralleled a quieter effort underway at the International Security Assistance Force (ISAF) headquarters under General Stanley A. McChrystal and the oddly-named Afghan Assessment Group (AAG, a group that contained no Afghans). Kilcullen offers practical advice for commanders and assessment analysts, diving right into the debate over specific metrics, or things to measure in complex operations. He offers specific examples relevant to the Afghanistan war including recruitment versus desertion rates and detainee guilt ratio.

Whether or not one agrees with Kilcullen's approach to assessment or his specific metrics, he deserves credit for kick-starting an important public debate. *Measuring Progress* amplified the behind-the-scenes scramble underway at ISAF headquarters and gave the analytic community a public starting point for debate. Are these metrics "right", or are other metrics better? Should we seek to quantify the war or does the overly aggressive pursuit of quantification undermine effectiveness? How does this approach compare with the one taken in Iraq, or Vietnam, or the hundreds (or perhaps thousands) of metrics actually tracked and reported on during the Afghanistan war? Time has not reduced the importance of Kilcullen's direct, if controversial analysis.

14.3.4 *Measuring Effectiveness in Complex Operations: What Is Good Enough?* (2009) [9]

Sarah Jane Meharg (Available Online? Yes – free).

Just as Clancy and Crossett established the need for better assessment of irregular war, Meharg's article reaffirmed the gap in peacekeeping, humanitarian, and stability operations assessment. She brings both a Canadian and United Nations (UN) perspective to the debate, helping to shift some of the community's focus

from Iraq and Afghanistan. Meharg is a firm believer in the need for quantification to assess effectiveness, describing several different but ultimately broad approaches to fill the existing gap in practice. She acknowledges the idiosyncratic nature of complex operations assessment but argues that the community can do better to develop more broadly applicable approaches.

One of Meharg's most important contributions is to highlight the debate over "good enough" objectives in complex operations. In many ways this parallels the debate over good enough approaches to assessment; her own acknowledgements of the limits of quantification serve as examples. While *Measuring Effectiveness in Complex Operations* fails to offer clear and detailed recommendations, it continues to provide insight into the various challenges faced by peacekeepers, humanitarians, and others working to improve governance in less-than-war conditions.

14.3.5 *Progress Assessment in a Multinational Operation – A Norwegian Perspective* [10]

Elin Marthinussen, Bård Eggereide, Frode Rutledal, and Alf Christian Hennum (Available Online? Yes – pay).

Marthinussen et al., describe an important subset of complex operation assessment: development and stability operations. *Progress Assessment* reports on their lessons from assessing progress for a Provincial Reconstruction Team (PRT) in Afghanistan. While NATO infantry units fought the Taliban, the PRTs were responsible for applying hundreds of millions of dollars in development aid across the country. Defining objectives and providing some insight into effectiveness was essential to ensure the money was well spent, and that coalition nations continue to provide assistance over time. Marthinussen, et al., offer practical solutions, tying together standard military technical concepts like Measures Of Performance and Effectiveness (MOPs and MOEs) with more academic applications. They offer examples and guidelines for both theory and application and include specific examples from Afghanistan.

While *Progress Assessment* may not offer a universal solution for assessment, it continues to offer a good one-stop approach for any small assessment team tasked with determining the effectiveness of development aid in irregular war. It also showcases the non-U.S. assessment capabilities resident in NATO, and offers some important lessons for U.S. assessment experts.

14.3.6 *Complex Operations: Countering Irregular Threats Joint Operating Concept* (2010) [11]

U.S. Joint Staff (Available Online? Yes – free).

U.S. doctrine writers have struggled to incorporate assessment into irregular warfare and stability operations manuals since the onset of the wars in Afghanistan and Iraq. In 2006 the U.S. published *Counterinsurgency*, FM 3-24, and also the *Commander's Appreciation and Campaign Design* handbook. *Counterinsurgency* offered a scattershot approach to assessment, suggesting it was added as an afterthought. The handbook was far more deliberate, outlining a detailed effects-based approach to assessment of war. Neither approach was particularly useful to analysts in the field. The *Complex Operations* concept offered little in the way of progress towards furthering the field of assessment, but did provide a buried appendix (D) with a step-by-step breakdown of irregular warfare objectives and suggested measurements. Coupled with a thoughtful methodology and applied by trained analysts, *Complex Operations* might be a useful guideline.

14.3.7 "Operations Assessment in Afghanistan is Broken: What Is to Be Done?" (2011) [13]

Stephen Downes-Martin (Available Online? Yes – free).

Writing from his perspective as a scientist and assessment expert, Downes-Martin delivers a scathing attack on the assessment efforts in Afghanistan. He describes his experiences providing direct support to commanders and assessment cells at the provincial and at the NATO Headquarters level. Downes-Martin decries what he observed to be an unwarranted obsession with scattershot data collection and over-quantification of complexity. His leading sentence sets the tone: "In the absence of a credible numbers-based theory of counterinsurgency there can be no objective, numbers-based assessment for Operation Enduring Freedom" [13] p. 103. At least through 2011, military assessment teams across Afghanistan were applying theories that did not stand up to basic mathematical scrutiny. But they were often doing so at the behest of commanders who had even less understanding of complex operations assessment than their staffs. The ongoing gap in theory, doctrine, and training in assessment was having a direct effect on operations.

Downes-Martin is refreshingly blunt or, for those in his crosshairs, perhaps too aggressive. His arguments are difficult to refute and they show that little progress had been made towards improved assessment ten years into the Afghanistan war and more than 35 years after the end of the Vietnam War. He offers simple solutions: focus on the objectives, be logical, be practical, quantify when it is necessary and when it is possible, but do not commit scientific sins in the pursuit of unobtainable accuracy and precision.

14.3.8 *Embracing the Fog of War: Assessment and Metrics in Counterinsurgency* (2012) [5]

Ben Connable (Available Online? Yes – free).

Embracing the Fog of War reported on three years of research into irregular warfare assessment, focusing on theories and practice in Vietnam, Afghanistan, and Iraq. This 340-page report seeks to rip irregular war assessment down to its foundations, describing and questioning the most commonly applied theories and methods used by operations staffs. It takes a critical view of current practice circa 2012, supporting Downes-Martin's critique of over-quantification and the unrealistic pursuit of accuracy on one hand and good enough precision on the other. *Embracing the Fog of War* offers a detailed chapter on the Vietnam assessment process, including a critique of Thayer's work, body counts, and the famous (or infamous) Hamlet Evaluation System. After ripping apart existing assessment practice, it offers an alternative template for irregular warfare assessment in the appendices.

While *Embracing the Fog of War* seeks to be comprehensive, it fails to address the depth of the European experience in Afghanistan. By design it does not directly address peacekeeping or humanitarian assessment or even small-scale counterinsurgency operations. This is a book about large-scale counterinsurgency assessment.

14.3.9 Assessment: Joint Doctrine Assessment Note 2/12 (2012) [14]

United Kingdom Ministry of Defence (Available Online? Yes – free).

British contributions to complex operations assessment have been extensive but often quiet and unreported, emerging most often in official reports that may not be publicly available. The 2012 Joint Doctrine Assessment Note [14], now archived by the Ministry of Defence, gives a clear if frustratingly brief summary of contemporary British thinking on complex operations assessment. Following the assessment frustrations during the 2009–2011 surge in Afghanistan, the 2012 Note offers a sober, balanced approach to effectiveness evaluation with a specific focus on irregular war. It warns against both over-quantification and the adoption of an excessive menu of metrics, suggesting increased focus on objectives and outcomes rather than a more short-term, effects-based assessment.

Some of the best contributions of the Note are in the appendixes. It describes how techniques more typically applied to peacetime monitoring and evaluation, like interviews and focus groups, can be applied as effective data collection tools in irregular warfare or stability operations assessment. In the main body the Note also suggests that assessment analysts identify a theory of change before applying a methodology, a logical recommendation that is echoed in some of the other contemporaneous literature.

14.3.10 "A Best Practice for Assessment in Counterinsurgency" (2014) [15]

Jonathan Schroden (Available Online? Yes – pay).

Schroden is one of the more prolific American authors on complex operations assessment. His insights come from tours in both Iraq and Afghanistan providing direct support to staffs and commanders. *A best practice for assessment in counterinsurgency* represents the culmination of his experience and efforts: it is a description of the strategic assessment design he built and applied as part of the Afghan Assessment Group at ISAF. Schroden explains how, with guidance from the commanding general, his team set doctrine aside and developed a tailored strategic assessment process designed to meet the commander's needs. It involved a multi-step process drawing from a variety of bottom-up efforts, feeding into a report that contained both data analytics and the subjective inputs from commanders across the country. The results are more narrative than chart-driven. Schroden describes his concerted efforts to avoid the easy but often misleading middle ground offered by colour-coded rating scales, choosing instead to publish more data rather than to aggregate.

While the article is thin on details and does not have any representative images, it offers a good example of a top-level assessment approach in a large, long-term counterinsurgency campaign. This same approach might have been applied to Vietnam or Iraq. However, it would not be universally applicable. Bottom-up strategic assessments of this depth and scope would have little relevance in small footprint operations or in operations with short timeframes.

14.3.11 "The Horizon Framework: Bringing Time into the Assessment of Counterinsurgency Warfare" (2013) [16]

Christian Bayer Tygesen and Kristian Knus Larsen (Available Online? Yes).

Tygesen and Larsen take earlier recommendations to have a "theory of change" to heart. The Horizon Framework is a new theory for complex operations assessment designed to cut across the muddled and (arguably) overly complex efforts that have dominated NATO and other coalition efforts in the past. Their primary argument is that most assessments fail to get to the heart of counterinsurgency theory. If counterinsurgency is a battle for control and support of the population, then assessment should cut to the chase and tell the commander how and when that control might be achieved. Horizon Framework assessments are focused on long-term behavioural change rather than short-term changes in perception.

New theoretical approaches to complex operations assessment are much needed, and Tygesen and Larsen offer something new. Whether or not the Horizon Framework is practical remains to be seen. It is not clear that it has been tested, and there have been no significant efforts to challenge its precepts in the literature.

14.3.12 *Innovation in Operations Assessment: Recent Developments in Measuring Results in Conflict Environments* (2013) [17]

NATO Supreme Allied Commander, Transformation Division (Available Online? Yes).

This must-read volume is the precursor to the present report. Edited by a team of NATO's top operations analysts and assessment experts, it offers nearly 300 pages of articles on effectiveness measurement. While some articles focus on irregular war, *Innovation in Operations Assessment* offers an array of complex

operations assessment perspectives. Diverse authorship ensured both military and non-military perspectives, and also a focus on development, peacekeeping, and humanitarian assessment. Read together, these articles suggest many opportunities for crossover between wartime and normal peacetime applications. Measurement theories from the development community might be useful in military operations. ISAF's efforts to incorporate development assessment tools like MPICE and TCAF suggest reasonable crossover points [18].

One of the most important features of this report is that it brings together both U.S. and non-U.S. NATO experts, showcasing a balance not found in other NATO documents on complex operations assessment. While the U.S. is still probably over-represented, this report makes clear that the community of experts is stronger together.

14.3.13 *Are We Winning? A Brief History of Operations Assessment* (2014) [3]

Emily Mushen and Jonathan Schroden (Available Online? Yes).

Mushen and Schroden put the evolution of irregular warfare assessment into a broader historical context, showing how the Cold War and the advent of information technology and the Revolution in Military Affairs (RMA) made operations assessment simultaneously more accessible and more challenging. Analysts have unprecedented access to information, even considering the volume of reporting delivered during the Vietnam War. But they remain limited by inadequate or inapplicable theories of change, the limits of human analytic capabilities, and the inevitably changing character of warfare. *Are We Winning?* is an excellent starting point for anyone unfamiliar with the literature on complex operations assessment, and for experienced operations analysts looking to put the current literature into historical context.

14.3.14 *Assessing War: The Challenge of Measuring Success and Failure* (2015) [19]

Leo J. Blanken, Hy Rothstein, and Jason J. Lepore, eds. (Available Online? No).

Assessing War is a logical next step from Scott Sigmund Gartner's *Strategic Assessment in War*. Gartner's work on conventional war set a much needed standard for the anaemic literature on assessment theory and methods. Blanken, et al., expand upon Gartner's work, filling in the gaps with both irregular warfare and deep historical cases. *Assessing War* showcases new work by a range of expert academics and practitioners, including Gartner and U.S. Army Major General William Hix. It makes for interesting reading and should be considered part of the canon of assessment literature.

However, where Gartner's *Strategic Assessment in War* offered a linear, in-depth treatise on assessment, this edited volume has a more scattershot feel. It completes the mission suggested in the subtitle by identifying many of the challenges in measuring success and failure. But it does little to offer either a menu or an agreed-upon unitary theory alternative to existing practice. Apparently all of the authors are American, leaving Canadian, European, and other authors unrepresented.

14.3.15 Operations Assessment: Joint Doctrine Note 1-15 (2015) [20]

U.S. Joint Staff (Available Online? Yes).

Operations Assessment is the culmination of over a decade of U.S. military efforts to codify a practical set of lessons and guidelines for assessment. While it is not specifically focused on complex operations, it clearly derives from experience in Iraq and Afghanistan. Where mid-2000s doctrine offered half-hearted and sometimes contradictory advice to assessment staffs and commanders, this Joint Doctrine Note provides clear guidelines in a logical format. It warns against over-quantification, reflecting arguments in recent literature, but it also provides an array of options for data collection and analysis. Most importantly, *Operations Assessment* provides the detail that has been lacking in previous publications. Its appendixes provide forms and templates that staffs can readily adapt to bring a new effort online in short order.

14.4 BACKGROUND READING

Dziedzic, M., Barbara S. and Agoglia, J. (editors). 2008. Measuring Progress In Conflict Environments (MPICE)-A Metrics Framework for Assessing Conflict Transformation and Stabilization. U.S. Institute of Peace; James Derleth. Undated. The Tactical Conflict Assessment Framework, U.S. Agency for International Development.

Williams, A., Bexfield, J., Farina, F.F. and de Nijs, J. (editors). 2013. Innovation in Operations Assessment: Recent Developments in Measuring Results in Conflict Environments. Allied Command Transformation, NATO.

14.5 AUTHOR'S BIOGRAPHY

Dr. Ben Connable is a senior political scientist at the RAND Corporation where he works on warfighting, assessment, regional, and intelligence issues. Ben led a multi-year research effort focused on counterinsurgency assessment methods, resulting in the publication of *Embracing the Fog of War: Assessment and Metrics in Counterinsurgency*. He has also published on assessment in Vietnam and supported a previous SAS panel on assessment in Afghanistan. Ben was the keynote speaker at the 2018 Military Operations Research Society assessment conference. He is a retired Marine Corps intelligence and foreign area officer, an adjunct professor at Georgetown University, and an adjunct professor at the Pardee RAND Graduate School.

14.6 REFERENCES

[1] Thayer, T.C. 1975. A Systems Analysis View of the Vietnam War, 1965-1972.

[2] Thayer, Thomas C. 1985. War Without Fronts: The American Experience in Vietnam, Boulder, CO: Westview Press.

[3] Mushen, E. and Schroden, J. 2014. Are We Winning? A Brief History of Military Operations Assessment. Alexandria, VA: Center for Naval Analyses.

[4] Gartner, S.S. and Edson Myers, M. 1995. Body Counts and 'Success' in the Vietnam and Korean Wars. Journal of Interdisciplinary History, 25(3):337-395.

[5] Connable, B. 2012. Embracing the Fog of War: Assessment and Metrics in Counterinsurgency. Santa Monica, Calif.: RAND Corporation.

[6] NATO. 2015. NATO Operations Assessment Handbook. http://www.natolibguides.info/ld.php?content_id=30192868.

[7] Clancy, J. and Crossett, C. 2007. Measuring Effectiveness in Irregular Warfare. Parameters, pp. 88-100.

[8] Kilcullen, D. 2009. Measuring Progress in Afghanistan. Unpublished report, Kabul, Afghanistan.

[9] Megharg, S.J. 2009. Measuring Effectiveness in Complex Operations: What Is Good Enough? Calgary, Canada: Canadian Defence and Foreign Affairs Institute.

[10] Marthinussen, E., Eggereide, B., Rutledal, F. and Hennum, A.C. 2010. Progress Assessment in a Multinational Operation-A Norwegian Perspective. Conference paper, presented at the 4th IMA Conference on Analysing Conflict Transformation, University of Oxford.

[11] U.S. Joint Forces Command, J9. 2006. Commander's Handbook for an Effects-Based Approach to Joint Operations. Suffolk, VA: Joint Warfighting Center, Standing Joint Force Headquarters.

[12] U.S. Joint Forces. 2010. Irregular Warfare: Countering Irregular Threats. Joint Operating Concept. https://www.jcs.mil/Portals/36/Documents/Doctrine/concepts/joc_iw_v2.pdf?ver=2017-12-28-162021-510.

[13] Downes-Martin, S. 2011. Operations Assessment in Afghanistan is Broken: What Is to Be Done? Naval War College Review, 64(4):103-125.

[14] Ministry of Defence. 2012. Joint Doctrine Assessment Note 2/12: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/432648/20150427-DCDC_JDN_2_12_Archived.pdf.

[15] Schroden, J. 2014. A Best Practice for Assessment in Counterinsurgency. Small Wars and Insurgencies, 25(2):479-486.

[16] Tygesen, C.B. and Larsen, K.K. 2013. The Horizon Framework: Bringing Time into the Assessment of Counterinsurgency Warfare. https://smallwarsjournal.com/jrnl/art/the-horizon-framework.

[17] NATO. 2013. NATO Innovation in Operations Assessment: Recent Developments in Measuring Results in Conflict Environments. NATO Supreme Allied Commander, Transformation Division.

[18] James, D. Briefing, undated. The Tactical Conflict Assessment Framework, U.S. Agency for International Development. http://www.energytoolbox.org/library/infra2009/references/2-30_12-15-09_setting_infrastructure_priorities_for_prts_crnkovich.pdf.

[19] Blanken, L.J., Rothstein, H. and Lepore, J.J. (editors). 2015. Assessing War: The Challenge of Measuring Success and Failure.

[20] Headquarters, U.S. Department of the Army. 2006. Counterinsurgency, Field Manual 3-24. Washington, D.C: Marine Corps Warfare Publication, 3-33.5.

Part 2: EXAMPLES OF OPERATIONS ASSESSMENT PRACTICE

Chapter 15 – SOCPAC ASSESSMENTS: DEVELOPING FUNCTIONALLY-ALIGNED AND COUNTRY-BASED ASSESSMENTS IN THE ASIA-PACIFIC REGION

Sarah Thambidurai
US Army
UNITED STATES

Karla Wayman
US Army
UNITED STATES

ABSTRACT

This chapter discusses Special Operations Command Pacific's (SOCPAC) assessments, to include the process followed, data collected, and resulting products. Beginning in 2012, SOCPAC formally created an assessments team to measure whether progress was made in defined objectives. In 2014, the team introduced a more flexible and dynamic approach to assessments, to account for shifting objectives and functional requirements from higher headquarters.

15.1 BACKGROUND

The U.S. Pacific Command (USPACOM) area of responsibility is the largest and most diverse of all the Geographical Combatant Commands, with the world's three largest national economies, over half the world's population, and seven of the world's ten largest standing militaries. The U.S. Government requires the full range of diplomatic, information, military, economic, financial, intelligence, and law enforcement tools to advance its "Asia-Pacific Rebalance", a strategy launched in 2009 to refocus U.S. Government (USG) attention towards the growing political and economic powers in that region. U.S. Special Operations Forces (SOF), through SOCPAC, play an important role in this endeavour. Given the current fiscally constrained environment, assessing the utility and effectiveness of SOF Operations, Actions, and Activities (OAA) in the Asia-Pacific region is increasingly important to validate, and adjust as necessary, SOCPAC's engagements with partner nations.

15.2 MISSION AND ORGANIZATION

As a Theatre Special Operations Command, SOCPAC is a subordinate command of USPACOM tasked to coordinate, plan, and conduct special operations and related activities in the Asia-Pacific region, to enhance regional security and stability through development and cooperation with allies, partners, US Department of Defence, and other agencies. Additionally, SOCPAC must be postured to respond to crisis in support of the USPACOM commander's objectives.

SOCPAC executes its mission primarily during Phase 0, shape the environment operations, which include missions, tasks, and actions designed to dissuade or deter adversaries and assure friends, as well as set the conditions for contingency plans. These are generally conducted through security cooperation activities [1]. The myriad of OAA that SOCPAC components conduct with partner nations include Joint Combined Exchange Training (JCET); counter narcotics training; subject matter expert exchanges; mobile training teams; military information support operations; and civil affairs operations, among others. OAAs are the foundation through which SOCPAC works to achieve defined objectives. The overall SOCPAC strategy is articulated in the SOCPAC Supporting Campaign Plan (SCP), from which functionally based sub-campaign plans and country-specific sub-campaign orders are derived. The SOJ35, with input from other staff sections, then develops OAA based on the objectives and sub-objectives identified in the sub-campaign plans and orders. These objectives also provide the basic framework off of which the assessments are based.

SOCPAC's assessments team sits in the SOJ 54 (Strategy, Plans, and Analysis Division), and is composed of a Team Lead, three regional Assessment Analysts (South Asia and Oceania, Southeast Asia, Northeast Asia), a Data and Systems Analyst, and Operations Research and Systems Analyst (ORSA) who support the methodology development and refinement for all the assessments. All members of the team, except for the ORSA, are contractors, which provides a significant advantage as they do not face the same frequent turnover rate that military personnel do. Additionally, having a team dedicated to conducting assessments ensures a sustained effort and focus on the assessment process, instead of depending on the various staff sections to coordinate and conduct assessments without a managing force.

15.3 ASSESSMENTS FRAMEWORK AND PROCESS

SOCPAC uses the Plan, Direct, Monitor, and Assess (PDMA) cycle to explain how assessments feed into a continuous cycle of planning and improvement throughout the command (Figure 15-1), and the SOJ54 Assessments Team leads the development and conduct of assessments for stated operational and strategic objectives. SOCPAC assessment requirements are driven by the SOCPAC Commander, (COMSOCPAC), the SCP, and higher headquarters' guidance. The primary focus of SOCPAC Assessments have been the country Sub-Campaign Orders, which identify country-specific objectives and sub-objectives. However, as the command has matured and developed functionally aligned sub-campaign plans, the Assessments Team has been tasked to also assess along functional lines: Contingency Response, Counterterrorism (CT), and Counter Weapons of Mass Destruction (CWMD).

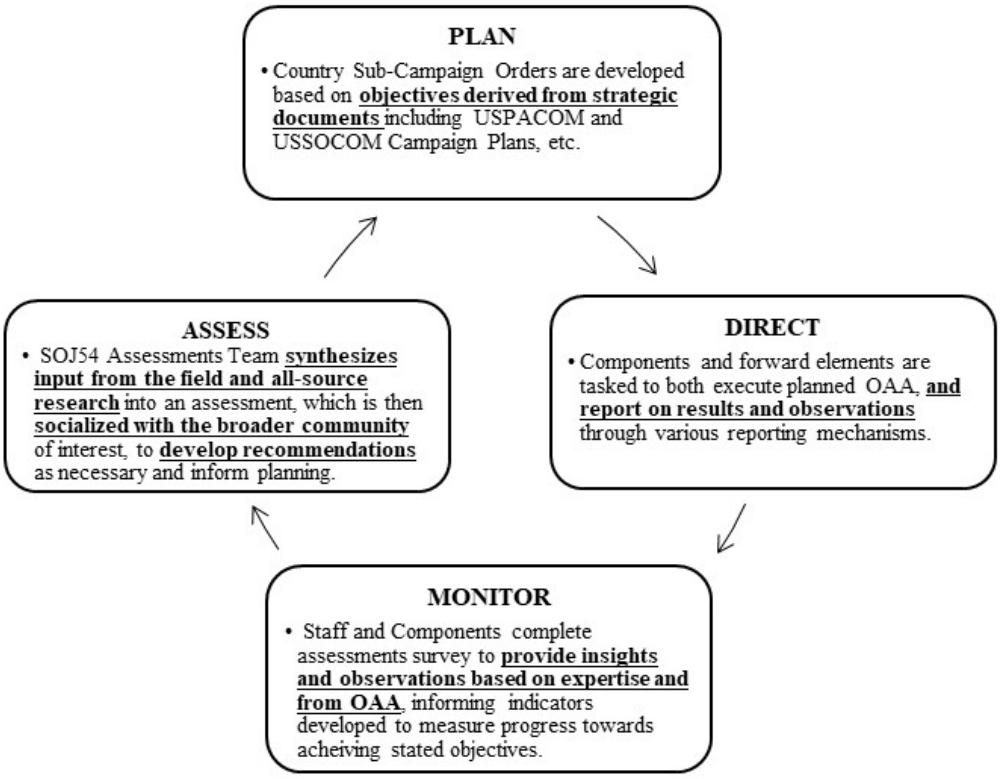


Figure 15-1: Plan, Direct, Monitor, Assess Cycle.

In order to approach assessing a functional subject, such as CT, the assessments team identified a set of enduring mission areas (Deepening Partnerships, Countering Violent Extremism (CVE), Build Partner Capacity (BPC), Enabling Means, CWMD, and Preparation of the Environment (PE)) which captured the themes identified in the sub-campaign plans. The general methodology behind the assessment process itself

is binned in three stages: indicator development, data collection and analysis, and collaborative efforts. The different nature of each mission area necessitated different approaches in the types of analyses conducted, and the sources of data incorporated into each area. As a result, each mission area is represented as a distinct assessment module, allowing flexibility in addressing the specified topic with an approach tailored to work best for that subject.

15.3.1 Indicator Development

Analysts develop assessment criteria for each functional area. These criteria, or indicators, provide a framework that assesses each objective. The indicators are answered through the lens of a specific country, providing information about a wide range of factors including environmental information, level of effort expended by SOCPAC components through various OAA with the partner nation, perceived level of skill in partner nation forces, and sentiment of the partner nation population. Indicators are organized so that they are not only module-specific, but also nest within a Lines Of Operation (LOO), Line Of Effort (LOE), the relevant sub-campaign plan, and the country sub-campaign order (Figure 15-2). This ensures the assessments team is able to use the research and analysis performed on the indicators to develop not only functional and country-based assessments, but also LOE and LOO-based products. These inform the command through the framework of COMSOCPAC's strategic vision.

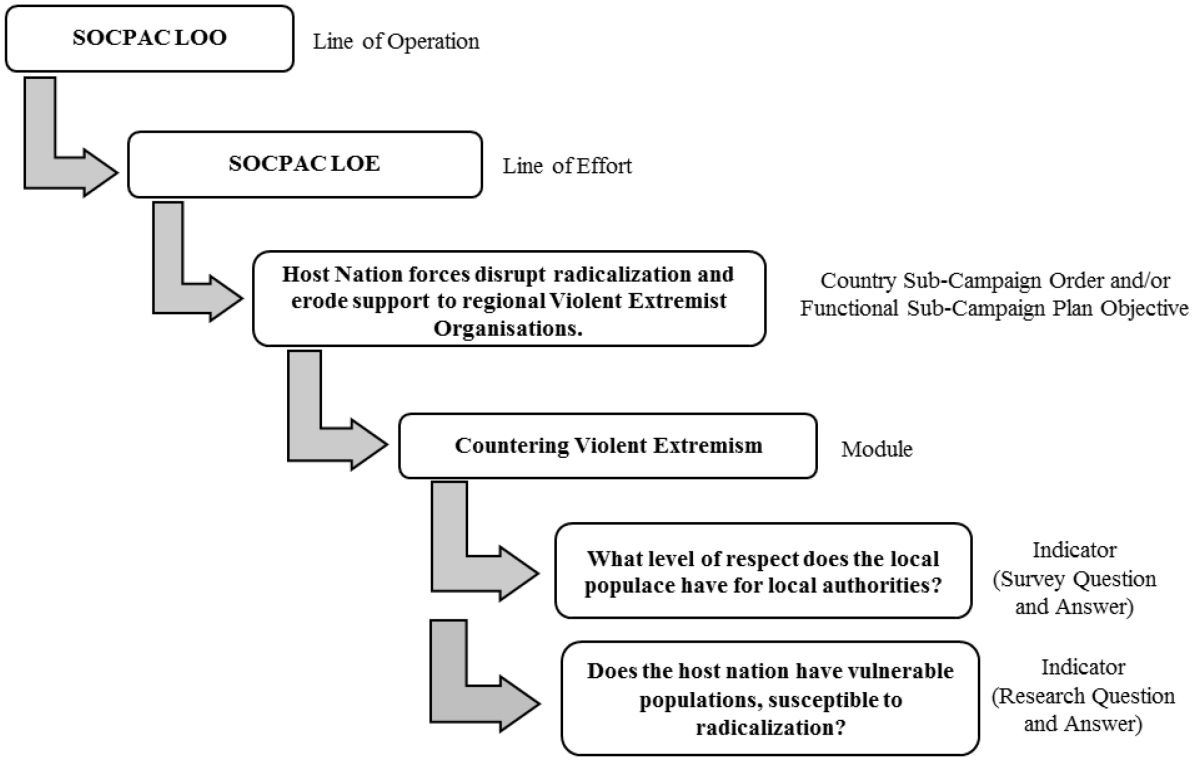


Figure 15-2: Nesting of Indicators within Objectives, Lines of Effort, and Lines of Operation.

The country sub-campaign orders did not fully address the objectives from the functional plans, necessitating a more holistic method to indicator development which assessed both the country and functional plan objectives. The mission areas required different approaches to the development of indicators; modules such as Deepening Partnerships and CWMD overlap significantly between sub-campaign orders and functional plans, so the objectives themselves were used as the basis to develop indicators. The remaining modules had numerous objectives which did not overlap; developing indicators based on all the relevant objectives would

have been infeasible due to the sheer number that would be required. Thus, the team developed alternative methods of indicator development. In approaching the CVE module, the assessment team worked with the Regional Information Support Team and identified relevant elements of the CT plan, and used Centre Of Gravity (COG) analysis to determine indicators. With Violent Extremist Organizations (VEOs) as the COG, the team identified Critical Capabilities (CC), Critical Requirements (CR), and finally, Critical Vulnerabilities (CV), which were used as the foundation for the development of CVE indicators, as shown in the example below:[1]

COG: Violent Extremist Organizations

CC: Recruit Members

CR: There are vulnerable populations in the country.

CV: Government addresses concerns/grievances of vulnerable populations.

Indicator: Does the host nation have vulnerable populations, susceptible to radicalization?

Indicator: What level of respect does the local populace have for local authorities?

The infrastructure development matrices, that capture PE efforts for each country, serves as the basis for the PE assessment module; the enabling means module focuses on counter threat finance, transnational organized crime, and collaborative targeting. The BPC module drew heavily from existing security cooperation data sources such as the Security Cooperation Information Portal and a RAND study on a Security Cooperation Prioritization framework and Propensity Matching Tool [2]. This approach to indicator development maintains flexibility for the given *functional* topic, allowing the analyst to identify indicators capturing all relevant factors to the module, without being constrained by the specificity of *country* objectives.

Indicators, categorised as either research questions or survey questions, are informed either through Subject Matter Expert (SME) surveys or research. Research questions can be answered through all-source research, while survey questions require the input of the operators who executed OAA with partner nations. An example of the two types is shown below.

Research question: Are there successful Violent Extremist Organization propaganda/recruiting efforts ongoing in country?

- There is already reporting about this indicator through available message traffic, after action reports, social media, and open source research.

Survey question: What level of respect does the local populace have for local government representatives?

- Research does not provide enough about the Host Nation regarding local populace support for institutional authority, but a Military Information Support Team (MIST) member would have the on-the-ground experience to provide insights on this indicator.

Each indicator, both survey and research, has an ordinal value from one to four and a fifth option for those unable to answer the question. A narrative block is also included for each indicator to provide context. Indicators were scored as in the sample below.

Sample Indicator: What level of respect does the local populace have for local government representatives?

1) No trust.
2) Low degree of trust; more likely to believe VEO propaganda.

[1] For more information on COG Analysis, see Refs. [3], [4].

3) Moderate degree of trust; less likely to believe VEO propaganda.

4) High degree of trust.

N/A. Not observed or otherwise unable to make an accurate determination.

Once the indicators are developed through these modules, the regional analysts apply the relevant indicators to their countries. For example, in Indonesia, SOCPAC efforts cross all of these mission areas and thus all the indicators would be relevant, but, in Japan, where SOCPAC does not engage in CVE efforts, only a subset of the indicators is necessary.

15.3.2 Data Collection/Analysis

Data collection occurs through surveys and research. The survey is just one mechanism that special operations units report out to, regarding the OAA they conduct; utilizing the web-based platform, Web-Enabled Temporal Analysis System (WebTAS), operators take the surveys online wherever they have access to classified (or secure) networks. Survey respondents include the PACOM Augmentation Teams (PATs), embedded at the U.S. embassy; components that conduct JCETs; Civil Affairs and MIST members in theatre; and subject matter experts in the respective directorates, such as the SOJ5 Environmental Analysis team, SOJ3 Country Desk Officers, and SOJ2 regional intelligence analysts.

The remaining data collection is conducted through research; regional analysts leverage existing U.S. or coalition reporting mechanisms, such as HUMINT Online Tasking and Reporting (HOT-R), Multimedia Messenger (M3), Combined Integrated Data Network Exchange (CIDNE), SOCPAC SharePoint repositories, After Action Reports (AARs), and other data sources in order to minimise reporting requirements for the operators, but also to provide amplifying information for the assessment.

Each indicator is scored; survey questions, by the operators completing the survey, and research questions, by the assessors using all-source research. Survey responses are averaged for the final indicator score, and the indicator scores are averaged across each objective, to get one numeric score for each functional or country objective, depending on what assessment product is being developed. This is not ideal; the numbers should not solely represent how well the command is doing towards achieving an end state, nor is it mathematically correct to average ordinal numbers. This has been communicated throughout the organization and it is understood that the numbers are intended to be a starting point for discussion. Additionally, narratives always accompany the scores, adding depth and context. Both the quantitative and qualitative aspects of each objective's assessments are necessary to provide a comprehensive understanding of each objective.

15.3.3 Collaborative Efforts

Throughout the process of a SOCPAC assessment, there is whole-of-staff integration. During the initial development of indicators, SMEs in the J2, J3 and J5 validate the list of indicators. The research phase ensures the inclusion of products, SME interviews, and insights from the staff. In addition to the input from operators through the assessment surveys, the SOCPAC Assessments Team is able to conduct de-briefs with the outbound MIST, Civil Military Support Element, and PAT leads to get detailed information on their on-the-ground observations of partner nation forces and population, and the overall progress SOCPAC made in the six months they were in country. The integration of a wide range of input from operators not only enhances the assessment product itself by drawing from a broader pool of contributors, but it also increases the buy-in of the wider SOCPAC staff into the final assessment product as well as the process.

All assessment contributors and stakeholders review a completed narrative report and provide comprehensive feedback, identifying any points of contention. During this process, the assessments team and SMEs from the community of interest meet to discuss key issues and provide recommendations at an

Assessments Working Group. Those recommendations are then briefed at an OF-5 level Assessments Board for validation, to ensure those responsible for implementing the recommendations agree they constitute the appropriate course of action.

15.3.4 Assessment Products

The assessments process is continuous and the resulting country assessment products capture a six to twelve-month snapshot of each country assessed. Once the final products are approved and published, the update begins immediately with the regional assessment analyst continuously refining the products as new data, reports, and observations come in. Figure 15-3 outlines the product evolution.

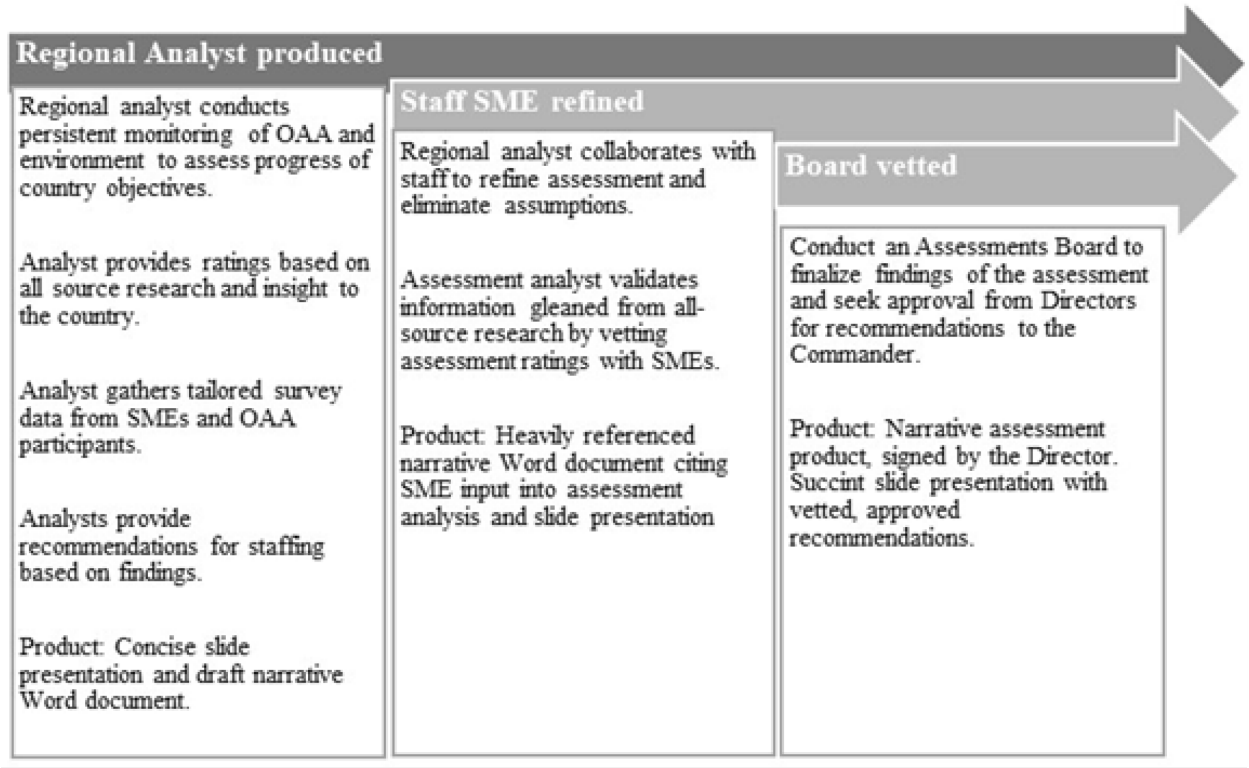


Figure 15-3: Assessment Product Evolution.

There are two primary products that result from the assessments process: an eight to ten-page narrative and a two-slide PowerPoint presentation. The PowerPoint slides provide the graphic overview of the assessment, as shown in Figure 15-4. It shows the status of each objective visually, on a graduated coloured scale, commonly called a status spectrum or a “thermograph” (see Figure 15-4), be it country or function, as well as a brief narrative for each objective, and recommendations for the way ahead. The narrative provides an in-depth look at each country, with key findings organized by each country objective and sub-objective. It also includes the staff-vetted recommendations, information on the primary SOCPAC components who engage in the country, the primary partner nation counterparts, as well as OAA historical analysis providing a comparison of SOCPAC engagement in the country compared to the overall USPACOM effort.

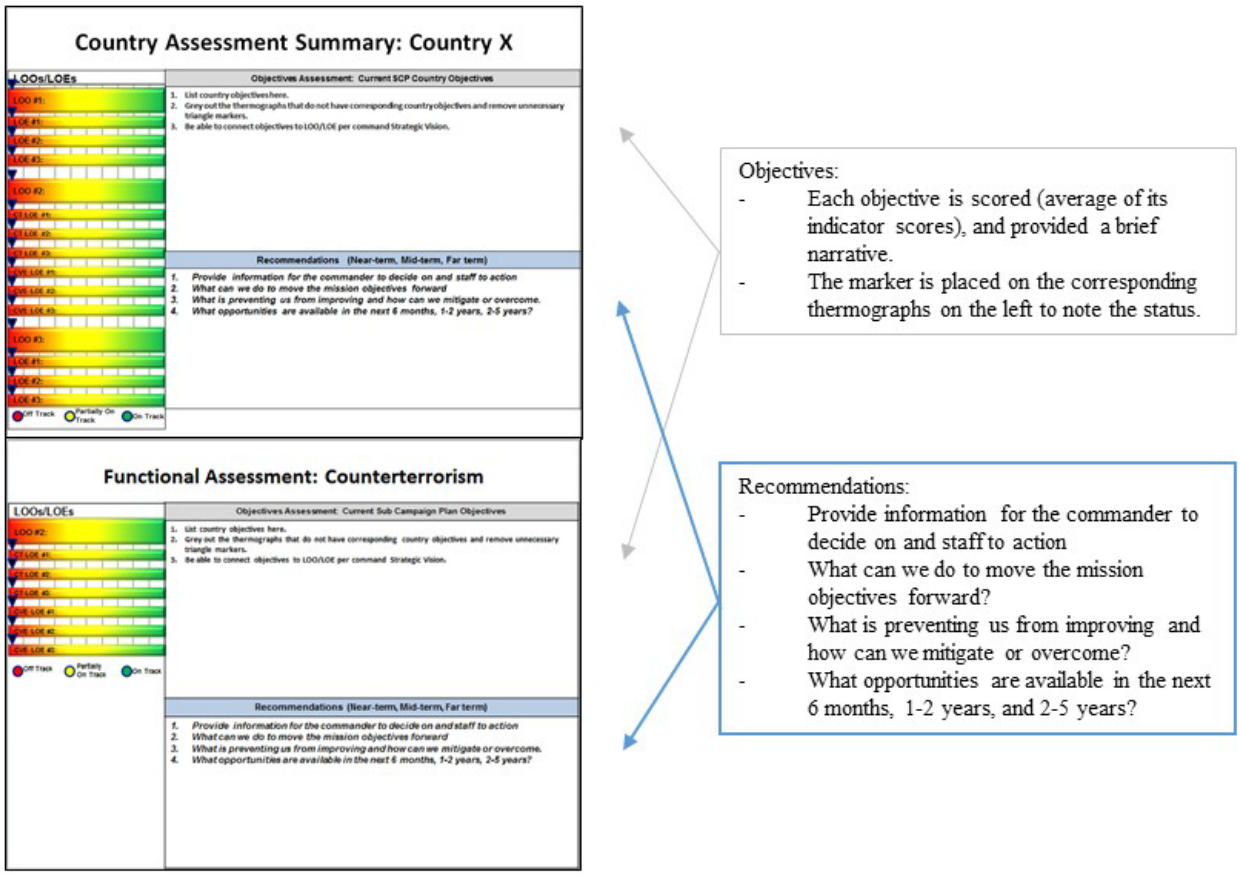


Figure 15-4: Function and Country Assessment Template.

15.4 IMPACT OF THE ASSESSMENT ON THE COMMAND

The staff recognizes that assessments are now the main venue for highlighting key issues and conveying recommendations. The recommendations that result from each assessment are used by planners to inform the annual Joint Planning Groups and Consolidated Special Operations Conference, and are incorporated into the sub-campaign orders for future OAA planning. The narratives provide a "Country-101" overview to incoming PAT members and new staff, including the issues SOCPAC faces in achieving its objectives, and the recommendations for the command's path forward in that country. The current PAT in country also uses the products to re-direct their efforts as necessary and inform resourcing requirements as they arise. Additionally, COMSOCPAC utilizes these products to inform and shape his approach for Key Leader Engagements throughout the region.

15.5 CHALLENGES AND LESSONS IDENTIFIED

The team faced four main challenges throughout its evolution: rapidly changing objectives; nesting within Higher Headquarters (HHQ) assessments; frequent turnover in military personnel; and finally, validation of data sources. The new assessments approach, necessitated by a requirement to assess both country operational plans and functional plans, resulted in a more agile and dynamic process, with the development of enduring indicators. Identifying thematic indicators enabled the team to overcome the challenge of rapidly changing objectives, setting a foundation for longer-term trend analysis, as well as nesting within HHQ assessments. Similarly, the choice of a thermograph to communicate the assessment

was based on maintaining consistency with USPACOM's red-yellow-green assessment products. Thermographs are also a useful shorthand method to portray easily quantifiable information, as long as the information is carefully caveated. A disadvantage of thermographs is that they "create the illusion of science" [5], and do not accurately depict the nuance apparent in the environment, because there is no precise way to locate the marker. In our case, the marker on the thermograph is merely a starting point for discussion, supplemented by accompanying narrative. SOCPAC's rapid turnover of military leadership, something common in many military commands, made it difficult to sustain a product in a format that is appealing to its primary stakeholders. Previous leadership valued a long narrative, typically around 40 pages, which, while very informative could prove cumbersome to some readers. The new products are flexible enough so that they highlight key takeaways in the two-slide thermograph presentation, but also provides further detail for stakeholders in the longer narrative.

The reliability of data sources will always be a challenge. There are gaps in collecting, processing, and disseminating operational reporting. However, SOCPAC's transition to CIDNE has been helpful in consolidating most of the operational reporting in a single database, and the command's reliance on SharePoint is useful for sharing information. However, there are still substantial gaps in reporting. When components conduct an OAA, they do not necessarily come back through SOCPAC to be de-briefed, costing the assessments team valuable insights. Many times a CIDNE entry may not capture all an OAA's detail as much an assessor would require. The assessment surveys themselves are limited not only by a very small sample size, but also by the enthusiasm and time available of the respondent on a given day. However, this is the nature of assessments, especially regarding specific data requirements about a wide range of operational and strategic-level indicators; there are only so many people that can be surveyed, and there will always be a subjective nature to it. As long as the audience for the assessment understands limitations such as these, and every effort is made to find additional data from a variety of sources to depict fairly the whole picture, the products are still very useful and relevant.

15.6 ADVICE TO NEW ASSESSORS

Integrate with the Staff. Assessments reports cannot be produced in a vacuum; their production is a whole-of-staff process and a whole-of-staff product. Assessors require significant input from the staff sections for the assessment to be an accurate reflection of the command's operations. At SOCPAC, the particularly close relationship with the SOJ3 is vital to the success of the Assessments Team. Without the SOJ3, the Assessments Team would not have as strong participation on the surveys by the operators, or be able to implement revisions for data collection systems on OAA and brief incoming staff on assessment processes to ensure continued buy-in and understanding of the assessments. As a part of this integration, the team is able to leverage already ongoing events, processes, and reporting in order to minimise requirements.

Knowledge Management Matters. The principle database SOCPAC uses is CIDNE, which has numerous limitations. It is imperative that the Science and Technology community continues to develop advanced databases that are intuitive for inexperienced users, especially those providing the data from the field, and searchable. The quality of assessments will increase significantly in correlation to database technology. The assessors must maintain a close relationship with the staff section in charge of developing, implementing, and managing the command's knowledge management in order to ensure that the data collection and management needs are understood, and met to the extent possible.

Know your audience. The assessment process is more than just a staff drill. Products should be deliberately crafted to ensure their operational usefulness to stakeholders; otherwise, the products are not serving a purpose. At SOCPAC, assessment products are primarily for senior leadership, but also for operators to plan OAA more effectively and reassess country objectives as necessary. In order to fulfil multiple purposes, the assessment product needs to be something that is both useful and appealing to those audiences. The product has to not only provide value to the operators, but it must be in a format that is usable and digestible to them.

Be Flexible. Assessment is a constantly evolving (and improving) process. There is no one correct way of doing assessments. Each command has different priorities, sources of data, audiences with unique preferences, and methods of planning. As a result, the assessments process will almost always be changing. Presentation preferences will change, and requirements, data sources, and other factors will continue to shift over time. There may be best practices of conducting and presenting an assessment, such as avoiding the use of thermographs or averaging of ordinal scores, but many times the nature and personality of a command necessitate flexibility and adaptability. This, along with rigorous collaboration, is required for a successful product, where success is defined as stakeholders using the product to inform decision making in order to make operations more effective. To this end, a *good* assessment product is one that is tailored to reduce specific uncertainty for a given audience and problem set. Any structure or underlying methodology is acceptable so long as it is pointed, accurate, and analytically sound.

15.7 REFERENCES

[1] United States Department of Defense. 2011. Joint Publication 5-0: Joint Operational Planning. III-42, August.

[2] Paul, C., Nixon, M., Peterson, H., Grill, B. and Yeats, J. 2013. The RAND Security Cooperation Prioritization and Propensity Matching Tool. RAND Corporation.

[3] US Department of Defense. 2014. Joint Publication 3-26 Counterterrorism IV-10. 24 October.

[4] Strange, J. and Iron, R. No date. Understanding Centers of Gravity and Critical Vulnerabilities. Air University. Available from: https://theforge.defence.gov.au/sites/default/files/adfwtc04_centres_of_gravity_and_critical_vulnerabilities_by_strange_and_iron.pdf.

[5] US Department of Defense. 2015. Joint Doctrine Note 1-15, Operation Assessment. C- 3, 15 January.

Chapter 16 – ASSESSING FOREIGN SECURITY FORCES: THE CASE OF AFGHANISTAN NATIONAL ARMY SPECIAL OPERATIONS COMMAND

Kevin Larrabee
US Army
UNITED STATES

ABSTRACT

This chapter showcases an alternative assessment process and framework, which was developed to support the development of the Afghanistan National Army Special Operations Command Special Operations Advisory Group (ANASOC-SOAG) in 2015. The SOAG's mission was To Train, Advise, and Assist (TAA) ANASOC to be an operational headquarters able to man, train, equip, and sustain Special Operations Forces (SOF) operations in support of the strategic objectives set out by the Afghanistan Ministry Of Defence (MOD) and the Government of the Islamic Republic of Afghanistan (GIROA). This assessment process provided a structured way to track progress at the tactical and operational levels within the advisor command to improve ANASOC's effectiveness at command and control, personnel and equipment management, planning, and logistics. This chapter walks the audience through the initial operational design conference, assessment framework development, and implementation of the assessment. ANASOC-SOAG's behaviour positively changed during the assessment development and execution. The assessments process resulted in an increase of situational awareness, unity of effort, mentor continuity, and commander situational awareness.

16.1 BACKGROUND

The decision made on 10 June 2015 to send 450 American advisors to Iraq in order to assist the Iraqi Army against the Islamic State of Iraq and Syria (ISIS) is a direct reflection of the recurring need for U.S. Security Force Assistance (SFA) operations [1]. The ongoing wars in Iraq and Afghanistan are recent historical examples of the U.S. utilizing military advisors to increase foreign military capabilities. Many critics of our advisory efforts ask questions like "why can't the Iraqi and Afghans defend their own country after we trained and equipped them for almost a decade?" or "did the Iraqi and Afghan military units achieve the standards necessary to defend their country?" These commonly asked questions are answered in many ambiguous and politically-charged ways. Most approaches attempt to aggregate tactical-level "report cards" on the Foreign Security Force (FSF) at the operational level. Others attempts focus on the collection of large-scale metrics that are indicative of success, such as manpower fill rates, operational tempo, and equipment utilization. All of these methods appear useful, but in most cases, the overall assessment of progress remains uncertain.

This chapter showcases an alternative assessment process and framework, which was developed to support the growth and development of the Afghanistan National Army Special Operations Command Special Operations Advisory Group (ANASOC-SOAG). The assessment framework was developed by an operations research analyst from the Centre for Army Analysis (CAA). This new assessment process provided a structured way to track progress and improve units' effectiveness at the tactical and operational levels within an advisor command. The chapter presents a detailed description of the development of an SFA assessment framework conducted by ANASOC-SOAG in Afghanistan from May 2014 to January 2015. The assessment process continued past this point, but this chapter presents findings from the initial assessment process creation, implementation, and lessons learned.

The Afghanistan National Army Special Operations Command (ANASOC) is led by a two-star Afghan general and is responsible for the command and control of all SOF generating and operating elements. This includes the nine fighting Special Operations Battalions, known as Kandaks, Military Intelligence Kandak, support battalions, the Special Operations School of Excellence, and two Special Operations Brigade headquarters which serve as the command and control function for the individual SOKs.

As this chapter will show, the ANASOC-SOAG assessment process grew out of multiple attempts to understand if the Afghan Special Forces were prepared to conduct combat operations independently of Special Forces advisors. ANASOC-SOAG began in 2006 in order to grow the Afghan Special Operations Forces (SOF) capability. As the elite force became fully manned and equipped, their mission required them to transition into an operational headquarters tasked to plan and execute direct action operations. Direct action missions are usually short duration small-scale offensive actions by special operations forces to seize, destroy, capture, recover, or inflict damage on designated personnel or materiel.

16.2 ORGANIZATION

In 2015, ANASOC-SOAG was composed of people with a variety of different backgrounds and skill sets. The Special Operations Advisory Group (SOAG) was divided into internal and external staff sections. The command team was composed of U.S. Special Forces active duty officers and Non-Commissioned Officers (NCOs). The majority of the internal staff came from the Special Operations Detachment Korea (SOD-K). The external mentors were made up of Afghanistan-Pakistan (AFPAK) Hands, Special Forces, and individual augmentees who specialized in specific skill sets, such as logistics, intelligence, and maintenance. The Afghanistan Training Program (ATP) rounded out the staff with about 100 contract employees, primarily responsible for advising the support battalions, training the ANASOC School of Excellence (SOE) cadre, and mentoring the various ANASOC staff sections. The SOE housed the Commando School, roughly equivalent to U.S. Army Ranger School, and the Special Forces Qualification Course, which was closely modelled on U.S. Army Special Forces Selection Course. The SOE also taught specialty skills, such as the medic, engineer, and English courses. The SOAG structure provided the necessary life support and security for the mentors to properly mentor all ANASOC staff sections. SOAG's HQ was collocated with the ANASOC's at Camp Commando (ANASOC's Headquarters), allowing the SOAG mentors the opportunity to build strong relationships through daily interaction with the ANASOC staff. ANASOC-SOAG's higher command was NATO Special Operations Component Command (NSOCC). The command structure is highlighted in Figure 16-1.

16.3 ANASOC-SOAG'S MISSION

ANASOC-SOAG's primary mission is to Train, Advise, and Assist (TAA) ANASOC to be an operational headquarters able to man, train, equip, and sustain SOF operations in support of the strategic objectives set out by the Afghanistan Ministry Of Defence (MOD) and the Government of the Islamic Republic of Afghanistan (GIROA).

16.4 ASSESSMENT SCHEME

16.4.1 Origins

Operational approach conferences were conducted at ANASOC-SOAG roughly twice a year beginning in 2013. The conference produced a doctrinal product focusing the staff on specific Lines Of Effort (LOEs) to improve the effectiveness of ANASOC. The experienced SOAG Commander, serving from May 2013 to May 2014, and his staff completed an operational approach framework at the beginning of March prior to the new SOAG Commander arriving in theatre in May 2014. The operational approach was focused on logistical

capabilities, intelligence, and personnel management. Along with the operational approach was a list of tasks associated with each LOE to ensure effective progress was being achieved. These efforts assisted the new command team in cultivating their vision.

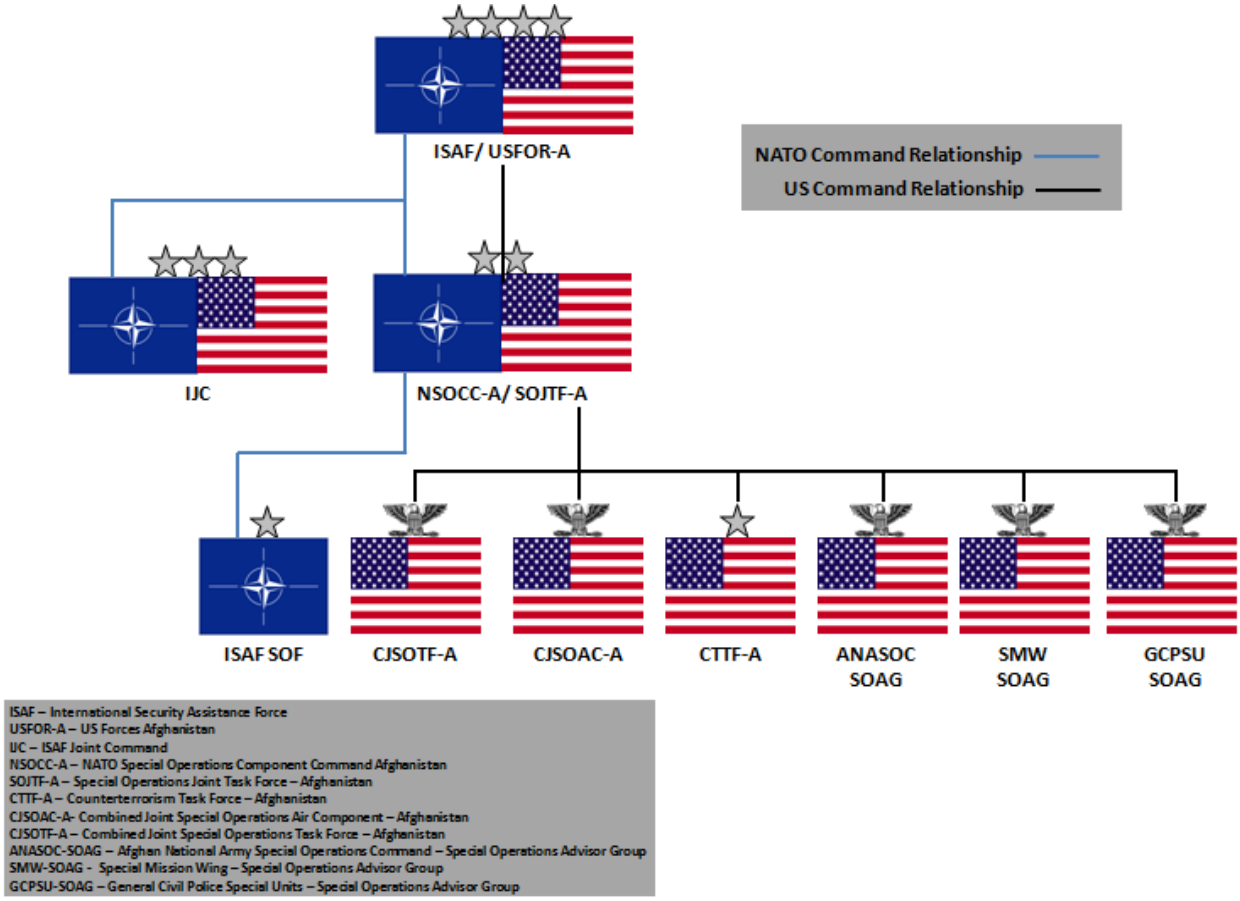


Figure 16-1: ANASOC-SOAG Higher Headquarters Organizational Chart.

The ANASOC-SOAG assessment framework came about as a result of one such operational approach conference that took place in late June 2014 under the new SOAG leadership. The three-day conference brought the entire command together to understand ANASOC's current environment and their desired end state. Before the operational approach conference, the ANASOC commander held a commander's conference attended by all of the ANASOC Battalion and Brigade commanders. The conference highlighted policy changes, successes, and areas of improvement observed over the last 6 months. The output of the ANASOC conference became inputs into the SOAG operational approach conference. The SOAG needed to understand ANASOC's strategic objectives before developing an operational approach to assist them in achieving their objectives effectively.

For the conference, the SOAG was divided into three groups to independently assess ANASOC's current capabilities through three different paradigms. The groups are described below:

- The first was through the seven domains in Doctrine, Organization, Training, Materiel, Leadership and Education, Personnel and Facilities (DOTMLPF). This approach is recommended as part of the assessment guidelines outlined in Ref. [2] Chapter 3.
- The second was through the six warfighting functions of mission command, movement and manoeuvre, intelligence, fires, sustainment, and protection (as outlined in Ref. [3] Chapter 3).

- The last approach was through the eight essential functions established by the International Security Assistance Force (ISAF), which were: Planning, Programming, Budgeting, and Execution (PPBE); Internal Controls: Assuring, Transparency, Accountability, and Oversight; Rule of Law; Force Generation; Force Sustainment; Command and Control; Intelligence and Strategic Communication.

The finalised operational approach stemmed from the warfighting functions paradigm, which resulted in ten LOEs. The first six were dedicated to the ANASOC's operating force and encompassed the six warfighting functions, while the remaining four focused on ANASOC's generating force responsibilities, centred on training, manning, equipping, and developing. The DOTMLPF and eight essential functions paradigms were not directly adopted due to ANASOC's tactical and operational focus. ANASOC's limited ability to implement MOD-wide change resulted in the SOAG focusing on things they could affect. Change within the Afghan MOD was predominately top-down initiated. For example, without MOD approval, ANASOC could not change doctrine or influence the way PPBE was conducted. SOAG advisors had minimal ability to affect changes that occurred at the MOD level. The advisors had to work within their counterpart's Afghan chain of command to make ANASOC more effective. Each one of the operational framework LOEs were assigned to a SOAG staff section to better align responsibilities and functional assessments. Each section identified milestones to improve activities under each unique LOE. Because there was not an established assessment framework, each LOE team had to build the assessment from the ground up. Milestones, enabling milestones, and metrics in support of the LOE needed to be created. The graphic below in Figure 16-2 highlights the process for the creation of the new operational approach through different paradigms, in order to establish the operational plan. The assessment process, shown on the right of the figure, highlights the nested assessment framework starting with the metrics and building to each one of the LOE assessments. A detailed explanation of this process is found in the upcoming sections.

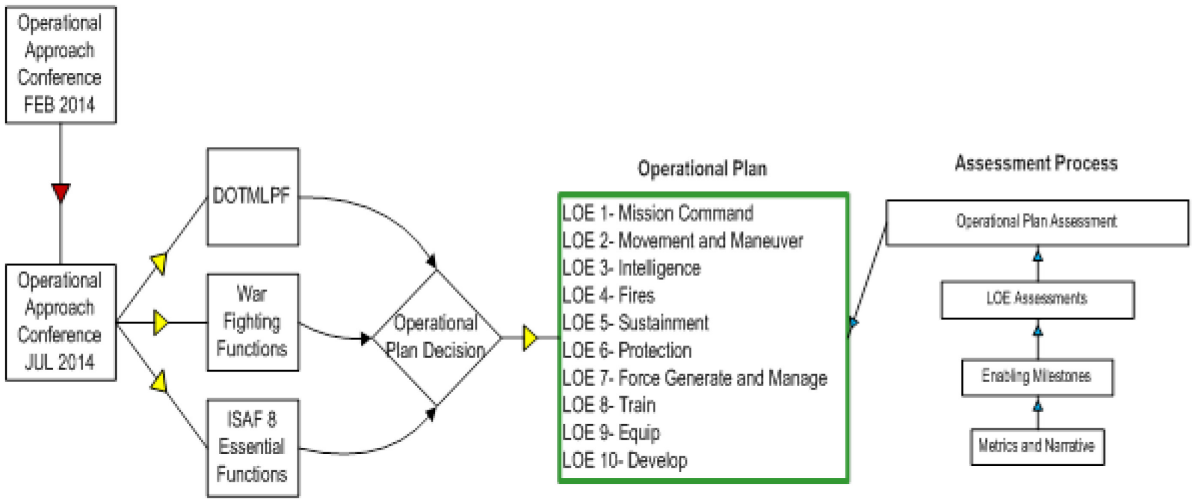


Figure 16-2: SOAG Operational Approach Flow Diagram.

16.4.2 Assessment Framework

An assessment framework was created to understand the progress made under the operational plan. Each SOAG advisor had daily interaction with their ANASOC counterpart, meaning that about 20 different independent exchanges occurred daily in an effort to support the planning, executing, and assessing ANASOC. The assessment framework attempted to centralise all of those efforts.

Assessing ANASOC's effectiveness was not a new idea. Nevertheless, the assessment process was never implemented effectively. The previous command used a task assignment worksheet to ensure the lines of

effort were improving. This worksheet was initially constructed, but never updated or briefed to the command group. Additionally, the SOAG was tasked by NSOCC and ISAF to complete a monthly five pillar assessment worksheet on each of the battalions within ANASOC. The five pillar assessment was not utilized effectively, with many advisors not updating summary fields and ratings from month to month. It also did not have any command emphasis due to that assessment's stagnant nature. ISAF primarily used the five pillar assessments to aggregate data in reports to Congress.

In late July of 2014, the operational approach was approved by the SOAG Commander. Once this occurred, the LOEs were further broken down into a series of milestones. Each LOE was tasked directly to an advisor or team of advisors. For instance, the G3 was responsible for LOE 1 (Mission Command) and LOE 2 (Movement and Manoeuvre), while the G2 was responsible for LOE 3 (Intelligence) and parts of LOE 6 (Protection). Each LOE lead provided the assessment Officer In Charge (OIC) three to four major milestones or objectives in support of that LOE. The team of advisors then developed Enabling Milestones (EMs) in support of the large milestone. In most cases, each LOE had three milestones and two to three EMs. Once the structure was complete, we developed metrics associated with assessing the enabling milestones.

16.4.3 Metrics

Developing metrics was the most arduous portion of the assessment framework because there were challenges to collecting data stemming from communication barriers, cultural differences, and geographic separations. These obstacles sometimes resulted in less confident assessments that were derived from a variety of poor quality sources, or that were sometimes devoid of empirical evidence. These included biases stemming from historical anecdotes passed down from previous mentors, incomplete information from higher headquarters, or miscommunication between the ANASOC and SOAG staffs.

Other things that prevented seamless understanding of the operational environment were lack of rapport between the SOAG advisor and the Afghan counterpart, advisor networks, and transparency. Without rapport between the advisor and his Afghan counterpart it was nearly impossible to assess the unit's current capabilities and future objectives. The advisor was unable to collect usable data without a good working relationship. Rapport was not just important with the Afghans; it was also necessary between higher and subordinate headquarters with advisors that understood ANASOC's role within the MOD. Having this relationship allowed the advisor the opportunity to verify what his counterpart said he was going to do, which gave him better situational awareness to assess effectiveness. The SOAG could not know anything they could not empirically observe themselves or that was not communicated to them by the Afghans. With the limitations present throughout the assessment process, the collection of metrics was essential to evaluating progress. The initial set of the metrics for all 10 LOEs was complete in the middle of August 2014. Data collection was the next step in completing the assessment framework.

16.4.4 Data

Data were collected in a variety of ways to enhance the understanding of ANASOC's effectiveness. While Coalition data sources, such as the Combined Information Data Network Exchange (CIDNE) database, were critical to data collection, one of the largest data sources was the Afghans themselves. Personnel records, operational mission logs, equipment logs, killed in action, and wounded in action datasets were available to the SOAG. The majority of all these reports were collected and disseminated on Excel spreadsheets. These datasets were critical data sources that were available only because of the relationships developed between the SOAG and ANASOC. The assessment Officer In Charge (OIC) was amazed by the amount of data collected by ANASOC. ANASOC collected detailed records of operations and personnel counts because it was required by the Afghanistan MOD. These records provided the basis of an assessment of ANASOC that was reported all the way up to GIROA.

The Afghan data sources frequently had errors. Some of these problems included: late reports (sometimes three to four days behind Coalition units' reports), lack of tactical details, and absence of after action reports that described the "so what" behind an operation. Additionally, some of the data sources were written down manually in notebooks, which had to be transferred to electronic files. This task was completed by roughly 5-7 Afghan Officers and Soldiers in the ANASOC tactical operations centre on a weekly basis. The SOAG used ANASOC data sources carefully to assess ANASOC's effectiveness. Mitigation techniques, such as waiting three days to report operational tempo rates and using monthly versus weekly personnel reports, were implemented to reduce the impact of ANASOC's data problems. Initial reports were usually wrong, so we allowed roughly three days for the data to be updated completely. As the number of Coalition advisors decreased, the information reported by advisors decreased, which increased the importance of ANASOC's data.

Everyone on the SOAG staff was a data collector. However, the assessments OIC collected the repeatable data streams, while the lead LOE advisor collected the data sources needed to assess their milestones. If requested, the assessment OIC would assist in collecting other data sources that would assist in assessing specific milestones.

16.4.5 Assessment Battle Rhythm

The SOAG initiated the weekly assessment battle rhythm briefs the second week of September 2014. The SOAG G3 and assessment OIC briefed three different briefing formats to the Commander, in which he commented that he wanted something straight forward that communicated the story of what was going on and how he could improve progress. The Commander directed that there would be two LOEs briefed per week prior to the weekly NSOCC SOAG brief. This schedule outlined a five week cycle to get through all 10 LOEs. The briefing format turned out to be two products: the first was a spreadsheet, known as the "assessment outline", that included the milestones, EMs, and metrics; while the second was a one-slide status report with information presented in four different quadrants (quad chart).

The spreadsheet provided the details required to support the EM's narrative. This document usually was for backup and was never briefed in its entirety. Table 16-1 illustrates the structure of the assessment outline.

Table 16-1: Example of the SOAG Assessment Outline.

LOE 1 - Mission Command							
Enabling Milestone	Purpose	Metrics	Last Period	This Period	POC	Narrative	Assessment
Milestone 1.1							
EM 1.1.1	Purpose EM 1.1.1	Metric 1.1.1.1	Last Measure 1.1.1.1	This Measure 1.1.1.1	G3	Narrative EM 1.1.1	Assessment EM 1.1.1
		Metric 1.1.1.2	Last Measure 1.1.1.2	This Measure 1.1.1.2			
		Metric 1.1.1.3	Last Measure 1.1.1.3	This Measure 1.1.1.3			
EM 1.1.2	Purpose EM 1.1.2	Metric 1.1.2.1	Last Measure 1.1.2.1	This Measure 1.1.2.1	G3	Narrative EM 1.1.2	Assessment EM 1.1.2
		Metric 1.1.2.2	Last Measure 1.1.2.2	This Measure 1.1.2.2			
		Metric 1.1.2.3	Last Measure 1.1.2.3	This Measure 1.1.2.3			
Milestone 1.2							
EM 1.2.1	Purpose EM 1.2.1	Metric 1.2.1.1	Last Measure 1.2.1.1	This Measure 1.2.1.1	G3	Narrative EM 1.1.1	Assessment EM 1.1.1
		Metric 1.2.1.2	Last Measure 1.2.1.2	This Measure 1.2.1.2			
		Metric 1.2.1.3	Last Measure 1.2.1.3	This Measure 1.2.1.3			
EM 1.2.2	Purpose EM 1.2.2	Metric 1.2.2.1	Last Measure 1.2.2.1	This Measure 1.1.2.1	G3	Narrative EM 1.1.2	Assessment EM 1.1.2
		Metric 1.2.2.2	Last Measure 1.2.2.2	This Measure 1.1.2.2			
		Metric 1.2.2.3	Last Measure 1.2.2.3	This Measure 1.1.2.3			

The assessment column on the far right side of the spreadsheet refers to three different statuses for each EM. Red represented a stalled effort that needed assistance from outside agencies or from the commander, yellow signified an EM with potential issues that required close monitoring of the enabling milestone, and green denoted the enabling milestone was on track with no advisory action needed.

The second assessment product was a one-page status of the LOE. This slide condensed all of the information on the assessment outline in a concise summary for the commander and the rest of the staff. In subsequent briefing cycles, the SOAG printed off the last briefed quad chart to ensure the staff and the commander had awareness of previous accomplishments and challenges. Figure 16-3 illustrates the structure of the LOE assessment product.

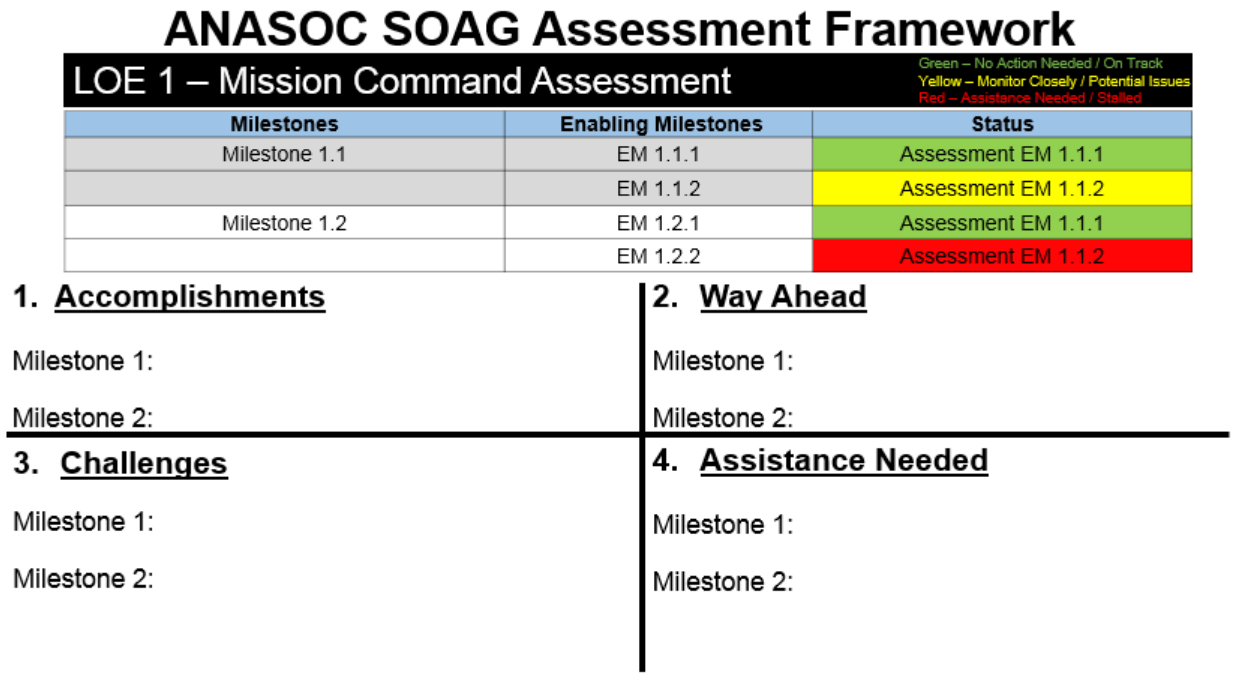


Figure 16-3: Example of the SOAG Assessment Framework Product.

The command assessments focused the SOAG staff on the most essential objectives, while providing a forum for the commander to assess progress, deliver guidance, and clarify specifics to improving the effectiveness of current and future operations. The number of distractions an advisor received on a daily basis could potentially derail efforts to achieve long-term success. Some advisors found themselves visiting the same efforts without any resolution. The assessment process provided the underlying foundation for what that advisor should be trying to achieve with his Afghan counterpart. The daily interactions, discussions, and activities were focused towards improving each of the milestones within each LOE.

Additionally, the assessment briefings provided an opportunity for the advisor to communicate the state of the LOE directly to other staff functions and the commander. The commander provided guidance and acknowledgement, while the other staff sections increased their situational awareness. The commander took a discussion-like approach to the assessment briefings. He asked detailed questions related to critical pieces of information that he needed to confirm or that he needed to test a hypothesis he had concerning an issue. He did not directly focus on the metrics related to each milestone, but understood they were supporting the quad chart discussion. He understood that assessing ANASOC's progress on a five week cycle would be slow and repetitious, but would keep the staff focused. The commander's direct involvement and support of the assessment process stressed importance of the staff's involvement in ensuring it was a quality product. Additionally, the ANASOC-SOAG Commander used the information from the assessments to help guide his mentoring of the ANASOC Commander. He emphasised critical LOEs that were slow in his discussions with senior ANASOC leaders and could reprioritize apply resources when necessary.

16.5 ASSESSMENT IMPACT

The lasting impact of the assessment was that it provided continuity for future commanders and advisors at ANASOC-SOAG. The SOAG assessment framework was built to follow U.S. doctrine so it could be integrated into a new command team with new rotational advisors. Additionally, the framework was flexible enough to be adapted for other advisory commands. The operational approach and assessment framework were briefed to multiple regional TAA Commands (TAACs) in Afghanistan. The approach was well-received and many TAACs were looking to adopt a similar assessment framework.

16.6 LESSONS LEARNED

Developing an original assessment framework was challenging. The biggest lesson learned was ensure the time spent on the assessment process is worthwhile in accomplishing the organization's mission. Roughly 15 to 20 hours a week were spent on collecting data, contemplating assessment statuses, and updating slides. The assessment OIC and the responsible LOE staff collaborated multiple times to ensure a clear and comprehensive look at the assessment. If the direction the assessment provided was not worth the time spent on improving ANASOC, the SOAG's effort was wasted. With that said, the ANASOC-SOAG's assessment was a living process that improved situational awareness and focused the command on what was important versus what was the "daily catastrophe".

Other lessons learned were:

- Be flexible: new advisors, changing command direction, and inconsistent data sources were all challenges. Remaining flexible and staying focused on effectiveness proved vital.
- Qualitative Assessments are not wrong: using written and verbal communication to flush out why things are not progressing is an effective way to communicate assessments. However, they must stay objective and be supported with facts.
- Incorporate your FSF counterparts into the operational approach and assessment framework: Without understanding your counterpart's objectives, constraints, and restrictions, advising FSFs will be a frustrating experience. Communication is critical in understanding slow progress.
- Provide assessment education to the staff: not everyone knows the importance of metrics and how they apply to the larger assessment framework. Provide formal training that assists the staff on how to collect data, manage data, and find assessment references.
- Spend the time validating the metrics and data sources: advisors are routinely overconfident in what types and how much data they can collect, resulting in assessment gaps and qualitative guessing to assess progress. Keeping metrics and objectives simple and to the point would have facilitated the assessment structure. The SOAG tried too hard to develop robust metrics that were unrealistic. In some of the LOEs, such as sustainment, there were way too many metrics that could not be collected by a 3-5 person team over the 6-week assessment cycle. Focusing on the easily collectable metrics that demonstrated progress (or the lack thereof) would have decreased advisors' workloads in order to assess ANASOC.
- Detailed Likert scale methods enhanced our assessment metrics: most of the empirical observations could not be quantified by the advisor or the Afghan counterpart, therefore Likert scale descriptions helped the advisor assess whether things were improving. An example to assess ANASOC's operational planning process is "Never", "Almost Never", "Sometimes", "Almost Always", "Always". This allowed advisors to qualitatively discuss why something changed from "Sometimes" to "Almost Always" through a detailed narrative or specific evidence of progress. This flexible change to metrics allowed advisors to look at the holistic objectives, without having to become overwhelmed with counting occurrences.

- Keep your eyes open and ask a lot of questions: Most activities in an advisory command cannot be collected in an Excel spreadsheet. Advisors must always be seeking truths and asking directed questions. Additionally, physical observations of counterparts can help infer information critical to assessments.

16.7 CONCLUSION

If the last fifteen years of warfare in Iraq and Afghanistan are indications of future conflicts, militaries around the world should plan and train to effectively execute and assess TAA missions. This chapter provides insights into the challenges faced and techniques employed to build a TAA assessment. These challenges included data reliability, mentors' commitment to the assessment process, and communication of the assessment to senior leaders. Incorporating quantitative and qualitative metrics, simplifying the lines of effort, streamlining the assessment's final products, and building the assessment process into the commander's battle rhythm proved to be invaluable techniques that made the assessment process successful. The assessments process provided direction, transparency, and unity of effort for the commander, staff, and subordinate commands. The development of a TAA assessment framework is a team effort that required communication, iteration, and cooperation. Future TAA missions will not just be an enduring mission for Special Forces, but one for conventional forces as well. Advising FSF is a critical mission that builds capacity in partnering security forces and will continue to be the lynchpin of U.S. engagement efforts overseas. This mission requires patient, knowledgeable, and dynamic military leaders. However, without an assessment process that sets conditions for improvement, mentors' energy, time, and effort can be squandered. Advisory commands need to focus on the critical long-term FSF objectives through the use of assessments.

Finally, ANASOC-SOAG's behaviour changed during the implementation of the assessment process. Situational awareness, unity of effort, and mentor continuity greatly increased. Mentors worked smarter through their defined lines of effort, enabling milestones, and metrics, but shared a common vision across functional staff sections. They focused time on the important and achievable, versus the lofty and unmanageable. The process became actionable and better defined, creating an organizational shift in training, mentoring, and advising ANASOC to become an effective special operations fighting force.

16.8 AUTHOR'S BIOGRAPHY

Kevin Larrabee is an operations research analyst with experience in wargaming, modelling and simulation, and operational assessment. His previous experiences at the Centre for Army Analysis and his current assignment at the U.S. Central Command (CENTCOM) provide him opportunities to define problems, analyse data, and communicate solutions to senior leaders across U.S. Department of Defence. In addition to his operations research assignment in Afghanistan, he deployed to Iraq in 2008 as a combat advisor for the 9th Iraqi Army Brigade and to Operation Iraqi Freedom in 2004.

16.9 REFERENCES

[1] Gordon, M. and Davis, J. 2015. In Shift, U.S. Will Send 450 Advisers to Help Iraq Fight ISIS. New York Times, June 10. https://www.nytimes.com/2015/06/11/world/middleeast/us-embracing-a-new-approach-on-battling-isis-in-iraq.html?mcubz=1.

[2] HQDA. 2013. FM 3-22, Army Support to Security Cooperation. http://www.apd.army.mil/epubs/DR_pubs/DR_a/pdf/web/fm3_22.pdf.

[3] HQDA. 2016. ADRP 3-0, Operations. http://www.apd.army.mil/epubs/DR_pubs/DR_a/pdf/web/ADRP%203-0%20FINAL%20WEB.pdf.

Chapter 17 – COMBINED JOINT STAFF HEADQUARTERS, ASSESSMENTS FOR OPERATION INHERENT RESOLVE – CONDUCTING ASSESSMENTS WITH LIMITED DATA

Joseph P. Nowak
Combined Joint Task Force – Operation Inherent Resolve
UNITED STATES

ABSTRACT

The Combined Joint Assessments Cell conducted assessments and analysis for the Combined Joint Task Force – Operation Inherent Resolve (CJTF-OIR) Commander and supporting headquarters staff. The assessments team synthesized multiple data sources to measure and track: long-term and short-term changes in enemy tactics and focus, impacts from friendly operations conducted throughout the theatre of operations, impacts from the Coalition Air campaign, impacts from the Coalition strategic operation, and progress in training the Iraqi Security Force.

When the Assessments Cell was formed, the Coalition was not yet fully manned; availability of relevant data was significantly limited and often lagged behind events by several weeks. This case study will begin with an overview of the CJTF-OIR organization and how the Assessments Cell developed its charter and battle rhythm. Next, the case study will review the Assessment Team's effort to overcome data obstacles and provide timely, relevant, and credible analysis to inform senior leaders in a Coalition conducting operations in a complex environment. The case study will conclude with a discussion of the Assessment Cell's successes, challenges, and lessons learned.

17.1 INTRODUCTION

Conducting assessments at any level requires substantial amounts of data. Assessments for a multinational coalition headquarters requires an understanding of how that data applies tactically, operationally, and strategically. This case study examines the assessments conducted at the onset of Operation Inherent Resolve (OIR) and how analysts overcame the challenges to delivering relevant, credible, and timely analysis to inform senior decision makers.

17.2 BACKGROUND

The United States led the initial invasion and subsequent combat operations in Operation Iraqi Freedom (2003–2010) and eventually in Operation New Dawn (2010–2011) as the Coalition transitioned authority back to the Government of Iraq (GoI). At the conclusion of these operations, the United States and other partner nations withdrew their military forces from Iraq. While the GoI insisted that the Iraq Security Forces (ISF) and the Iraqi Police were prepared to handle security, insurgent groups continued sectarian violence throughout the country. In June 2014, the Islamic State of Iraq and the Levant (ISIL), commonly referred to by its Arabic acronym, "Daesh", launched an offensive to create an independent Islamic Caliphate within the territory of Iraq and Syria. Taking advantage of the ISF's operational inexperience and the overall instability in the region, Daesh established a presence along the border between Iraq and Syria, and extended lines of communication through significant parts of each nation, including the key cities of Mosul and Tikrit. To make the

situation more challenging, multiple, disparate non-Daesh entities were also attempting to gain influence in Syria as the region destabilized, including but not limited to the Government of Syria, Syrian separatists, and Taliban fighters.[1]

To help the ISF halt and eventually roll back Daesh's expansion, the United States led the formation of the Combined Joint Task Force – Operation Inherent Resolve (CJTF-OIR) in October 2014, with the mission to degrade and ultimately defeat Daesh. To accomplish this mission, CJTF-OIR employed two separate lines of effort – a coalition air campaign and a Build Partner Capacity (BPC) program to train, advise, and help Iraq build its military capacity.

17.3 ORGANIZATION/MISSION

The Coalition contained members from more than 60 countries initially commanded by Lieutenant General James Terry from the United States with Major General Robert Bruce from the United Kingdom as Deputy Commander (DCOM). The Coalition was based in Camp Arifjan, Kuwait. The U.S. led the Coalition. The United Kingdom, Australia, New Zealand, Germany, Spain, Italy, and many other nations dedicated troops and resources to the effort. The Coalition Air campaign consisted of aircraft based out of multiple airfields throughout the U.S. Central Command (CENTCOM) Theatre of Operations and involved many of the nations from the surrounding region.

After the Coalition was established, Australian Brigadier Craig Furini became the Director, CJ-5 (Plans, Strategy, and Assessment). The purpose of the CJ-5 was to build a long-term plan to meet the Coalition's objectives, develop the strategy for meeting that plan, and assess the Coalition's progress. The CJTF-OIR CJ-5 Assessment Cell was structured into three divisions: Campaign Assessment, Iraq Security Force Assessment, and Security Trends Assessment. Each division conducted analysis while supporting the other divisions and staff elements.

The Coalition's primary objective was to degrade and defeat Daesh. Enjoying air superiority throughout the theatre, the Coalition engaged in a deliberate target development process to identify possible targets. After a target was developed, the Air campaign provided immediate kinetic effects throughout Syria and Iraq.

The BPC initiative was designed to build the Iraqi's military capacity and capability to enable them to degrade and ultimately defeat Daesh. While the Iraqis maintained control over all military operations and the Coalition agreement with Iraq limited the number of coalition troops allowed in country, the Coalition partnered with the GoI to grow Iraq's combat power. The Coalition developed the BPC construct to allow coalition trainers to go into Iraq at agreed-upon sites and assist the ISF until it was capable of independently sustaining a secure environment.

17.4 ASSESSMENT

The CJ-5 Assessment Cell was chartered to assess coalition progress across all lines of effort throughout each phase of the established campaign plan. The assessment provided leadership with the information and analysis required to make informed decisions and ensure operations were effective at achieving coalition objectives. Each of the three divisions focused on a different element of the security environment. Each division presented insights at the monthly Campaign Assessment Board (CAB) and contributed to the Monthly Campaign Assessment Report (MCAR). CJ-5 Assessments briefed the CAB to the Commanding General (CG), along with coalition staff officers and the Combined Joint Forces Land Component Command (CJFLCC) operating in Baghdad. The DCOM reviewed and approved the MCAR for dissemination within the Coalition.

[1] Actual numbers of Taliban fighters were relatively insignificant in terms of combat power, however the Taliban, namely the Tehreek-e-Taliban from Pakistan, had a vested interest in countering ISIS growth in the region.

The next three sections will describe each division and its role, provide examples of assessments each division conducted, and identify each division's challenges in producing those assessments.

17.4.1 Campaign Assessment

The Campaign Assessment Division (CAD) assessed the Coalition's progress against the ground campaign plan by examining things such as territory reclaimed from Daesh, operational readiness of the ISF, and coordination efforts between the Coalition and surrounding nations. The CAD briefed the assessment to the CG in the monthly CAB. The Campaign Assessment part of the brief highlighted key focus areas based on progress and opportunity. It also provided recommendations to ensure progress continued in the most effective manner.

Most of the challenges experienced assessing the campaign's progress were similar to those the CJ-5 faced in writing the Campaign Plan – an inability to actively collect against the threat and a lack of relevant data. Unlike the previous operations in Iraq, CJTF-OIR ground forces were not actively conducting offensive operations against the enemy they were charged with defeating. The Coalition, being restricted to advise and assist missions on the ground, had to innovate new ways of influencing the area of operations. Key leader engagements and partnering with the ISF as the ISF planned, prepared, executed, and assessed operations became one of the primary methods of establishing positive momentum against Daesh. Innovative (and sometimes unconventional) plans forced the CJ-5 to develop innovative methods to assess the Coalition's progress against those plans.

The CAD examined phases and milestones specified in the Campaign Plan to determine what actions must be completed in order to achieve each respective milestone and complete each phase. Some of the milestones were quantifiable, such as regaining key terrain or cities. These measures provided an occasional progress review on a long-term plan spanning beyond the next 12 months, but were insufficient in gauging campaign progress in a monthly assessment. To provide more timely and robust assessments, the CAD collaborated with the planners regarding progress on short-term goals while incorporating feedback from all of the directorates in the headquarters each month in completing the assessment, and summarizing the understanding gained in the assessment report.

Although quantitative data were limited, staff officers conducted qualitative assessments on the Coalition's progress compared with the published plan. The primary methods for achieving this were surveys of the staff, which were displayed on a "status spectrum", which conveyed how the Coalition's perceived progress compared with its planned headway. The status spectrum charts showed relative progress compared with previous months while considering intended progress. For example, the assessment of a long-term goal of retaking Mosul may be coloured green due to preparations and other successes prerequisite of taking Mosul, even if the ISF has not yet succeeded in that specific objective. The CAD leveraged the staff's open lines of communication to make informed assessments as to how well the Coalition was meeting the objectives stated in the Campaign Plan. This assessment, combined narratives from Subject Matter Experts (SMEs) that provided on-the-ground supporting evidence, supported the insights and recommendations the CAD presented to the CG and coalition staff. The CG approved each recommendation and assigned responsibility for completion of tasks to the appropriate staff sections. The CAD monitored as required and provided monthly assessment updates at the CAB.

17.4.2 ISF Assessment

The ISF Assessment Division (IAD) assessed the progress of the developing security forces in Iraq as the Coalition trained new units in each of the five BPC sites. Like the CAD, the IAD combined a limited amount of quantified data points, such as number of trainees in each course, with more subjective estimates of the ISF's overall progress. The trainers reported their estimates based on the listed tasks and outcomes specified in each respective course. Elements of the training resembled that of the coalition militaries; however, it was

tailored towards the Iraqi culture and military objectives. IAD conducted the assessment based on the ISF's ability to conduct prescribed Army functions as well as ISF effectiveness and progress towards an overall capacity that would allow Iraq to defend itself against threats.

IAD gathered most of the input for the assessment from coalition trainers conducting the ISF training mission at the BPC sites. Trainers assessed tactical and technical competencies (e.g., leadership, basic rifle marksmanship, etc.) at multiple levels for each class at the beginning of training and again at the completion of training to measure the training's effectiveness.

Like the CAD's assessment, the IAD's assessment centred on qualitative determinations that officers with the most knowledge and experience made. In the case of the ISF assessment, more quantitative data was available regarding the BPC training site progress (e.g., number of trainees graduated, average competency score, etc.), but because the training officers were embedded with the trainees and had immediate, first hand operational interactions on a daily basis, the qualitative assessment offered a more accurate and defensible measure of how well the ISF was progressing towards its goals. The assessment highlighted successes and shortfalls, identifying possible areas of focus for future coalition efforts and resources.

Due to the sensitive nature of the ISF assessment, the results were limited in distribution. The IAD conducted its assessments always aware of the potential risk that the assessment could be used against the ISF and the Coalition if it highlighted vulnerabilities. Fortunately, the assessment helped establish open dialogue between senior coalition leadership and the BPC sites to enable an honest, accurate accounting of ISF progress. Beyond the assessment, the established communication channels allowed leadership a clearer understanding of the situation. This understanding enabled leaders to field questions at the strategic level and make informed decisions with greater certainty that the Coalition was progressing effectively towards its objectives.

17.4.3 Security Trends

The Security Trends Division (STD) conducted analysis on the entire CJTF-OIR area of operations, to include Iraq and Syria. In addition to the monthly CAB, the STD briefed assessment trends to the CG and staff elements during the weekly coalition Battle Update Assessment (BUA). The monthly CAB focused on long-term trends and provided in-depth insights. The weekly BUA identified developing trends and looked back at most one quarter. Both analyses provided operational insights and made recommendations to meet near-term goals effectively while considering long-term objectives at the strategic level. The CAB briefings showed the overall campaign in context with the long-term trends. The BUA briefings and the staff's preparation for them generated dialogue and helped the staff think critically about how each respective element influenced the security environment. These briefings showcased ISF's successes against Daesh, identified areas of concern for future focus, and highlighted how various elements of the Coalition contributed to operations, (e.g., Air campaign strikes on key targets, trainers improving ISF Tactics, Techniques, and Procedures (TTP), and trainers helping ISF defeat Daesh TTP).

While internally generated data provided value to assessments, most STD analysis and reporting relied on data obtained from external sources, such as host nation reports and open source data collected by the Joint Improvised Explosive Device Defeat Organization (JIEDDO), now known as the Joint Improvised-Threat Defeat Agency (JIDA). JIEDDO provided the STD with a weekly Microsoft Excel file with updated data on direct fire, indirect fire, and IED attacks. Although the data was supposed to address the entire Area of Operations (AO), data covering Syria was limited due to a lack of coalition personnel on the ground and a lack of availability of open source reporting in the areas surrounding the most intense conflicts. Unlike previous military operations in Iraq, the CJTF-OIR was unable to leverage the thousands of deployed troops in country to reliably report a broad range of incidents and conduct site exploitations.

The reliance on data from external organizations often created a lag in timeliness for reporting. For example, if an IED detonated in a populated area, several news agencies might report on that incident, providing varying, and sometimes conflicting, pieces of information. If the IED was not detonated at or near a coalition-controlled location, the ISF would conduct the investigation and report its findings. That report would not always be sent to the Coalition in a timely fashion, and JIEDDO analysts still had to de-conflict the report with other reports before sending the update to STD.

The JIEDDO data collectors relied on all of these report sources to build the weekly incident tracker. Before JIEDDO published the new file, each report had to be verified for accuracy and de-conflicted with the other reports to ensure each incident was counted only once and the associated data was correct. Many host nation reports took longer than a week to arrive at the STD, and open source reports were considerably less structured. As a result, valid security incident reports could take up to six weeks to be registered and show up in the database of record.

The lag in security incident data proved to be a challenge in presenting timely, relevant, and useful insights to the CG on a weekly basis during the BUAs. To avoid improper comparisons between artificially low data from recent weeks and fully substantiated data from previous months, the STD focused mostly on changes in long-term trends, highlighting only short-term changes if they could be traced back to significant changes of enemy tactics, techniques, and procedures. To minimise incorrect conclusions, the STD made a concerted effort to educate the staff on the nature of the data so they understood that raw numbers from the previous week were less reliable than raw numbers from six weeks ago.

To supplement the security assessments and provide greater operational context to the analysis, the STD began tracking a variety of coalition-generated data contained either with the staff Headquarters (HQ) or on the CJTF-OIR SharePoint portal. The advantage of this data was that it was timely, more reliable, and changes in coalition trends could be quickly explained by coalition staff. This enabled analysts to connect Coalition and ISF operations to successes and challenges in the security environment. Building these relationships with and among the staff further enabled decision makers to confidently dedicate resources to the right areas in order to make effective impacts for the Coalition.

17.5 IMPACT ON THE COMMAND

The greatest impact that assessments had on the coalition staff and commander was in providing a better understanding of the Operating Environment (OE). As an example, the DCOM hypothesized that attack trends would follow the growing seasons in Iraq, similar to Afghanistan. The STD conducted simple data analysis to show that this was not the case in Iraq. By challenging the assumptions, engaging other staff elements, and taking advantage of graphical and geospatial software, STD was able to explain not only that the hypothesis was incorrect, but also why it mattered in relation to the situation. Everyone in the Coalition understood the ultimate objective of CJTF-OIR, and the plan on how the Coalition would enable Iraq to eventually defeat Daesh. The Coalition was challenged, however, in understanding how all of the many variables interacted around and within the environment.

CJ-5 Assessments, through regular interaction with the coalition staff, also helped build relationships and dialogue among the staff elements. These interactions drove better analysis, sharper insights, and stronger recommendations that the entire staff understood and supported in execution with a unity of effort. The result was that tactical observations could be given operational context and highlight strategic implications relevant to the staff and CG. In these ways, assessments positively influenced CJTF-OIR HQ and greatly contributed to making operations more effective.

17.6 LESSONS LEARNED

While gaining a perfect understanding of the OE is impossible, assessments enabled better understanding of the OE, which, in turn, helped decision makers make better decisions. By its nature, assessment relies on a variety of data sources and types. To innovate the right solutions for tough analytic problems, a variety of tools are required. Survey development and analysis, data processing (through Microsoft Excel, Access, or statistical software packages such as R Studio), and geospatial analysis (through software such as Esri's Geospatial Information System – ArcGIS) can all make an analyst more effective.

The following section highlights the primary lessons learned while conducting assessments in support of CJTF-OIR:

- **Establish an effective knowledge management system.** Knowledge management was key to overcoming the internal data challenges in CJTF-OIR headquarters. Initially, each Directorate maintained disparate databases, which were not automatically shared with the other staff elements. A significant amount of qualitative and quantitative data was generated, but analysts without the proper access faced challenges in assimilating that data into knowledge and insights. To remedy this, the STD participated in a knowledge management working group that the knowledge management officer led, to help bring the CJTF-OIR Directorates together and encourage information sharing across the headquarters through established policies and standards. The working group established a means of communication among all of the staff officers that generated or tracked data for the Coalition. Staff officers gained awareness of what the rest of the HQ was doing, and more readily shared their data and analysis through the CJTF-OIR SharePoint Portal. Improving access to data across the Coalition made analysts more able to link operational activities to specific effects within the OE. This in turn helped generate better insights and more impactful recommendations.
- **Deploy with the right computer equipment and software.** In the first several months after CJTF-OIR was established, the Coalition had not yet acquired all of the technical capabilities that would have enabled more effective analysis. For the analysts charged with standing up initial assessment capabilities, or in the case of CJ-5 Assessments, three separate divisions, it would have been significantly easier to get started if analysts could have been able to deploy with the necessary computer equipment and software. Data analysis suites can be expensive and difficult to install in some computer environments; bringing the pre-loaded software on computers sufficiently powerful enough to handle the processing would have immediately enabled CJ-5 Assessments to begin analysis.
- **Challenge assumptions with facts and sound reasoning.** Although the United States has been involved in conflicts for a long time, analysis varies as much as the environment. U.S. forces had only very recently left Iraq, and still maintained a measured presence in Afghanistan when the Coalition to defeat Daesh was formed. Among Operations Iraqi Freedom, New Dawn, Enduring Freedom, and Inherent Resolve, each held a different set of challenges, with different data sets and different questions to answer. The CJTF-OIR coalition staff attempted to apply lessons learned from the previous operations, but analysis proved that not all lessons were directly transferable, even in Iraq between the two different operations. Analysts in CJ-5 Assessments were asked on a regular basis to prove or disprove assumptions about how the enemy would react to ISF and coalition operations, how the civilians would act, how holidays changed behaviours, and more. Analysts quickly learned that no assumption should be taken at face value. Critical thinking about facts, assumptions, and perceptions is imperative when operating with multiple variables.
- **Take advantage of multiple methods of communication to effectively make salient points.** Being able to think creatively and critically is absolutely necessary to conduct good analysis. The ability to effectively communicate insights gained from the analysis (orally, graphically, or through the written word), prove the value of assessments, and maintain the commander's confidence in the analysis.

- **Take advantage of the knowledge and experience resident on the staff.** Any good staff officer, and especially the analysts on a staff headquarters, will be expected to understand a multitude of data types, not all of which will lie in their areas of expertise. In many cases, possessing an understanding of the data trends will not guarantee an understanding of operational impacts. Maintaining strong relationships with other staff members will help an analyst tie operational context and meaning to data and draw appropriate conclusions and insights from the analysis. An analyst will be unable to successfully inform the commander if the analyst is isolated from the staff. Therefore, analysts should engage the other staff officers to gain access to their knowledge and expertise. Building such relationships improves analysis, and opens avenues for other staff officers to be aware of and take advantage of the services and expertise residing in Assessments sections.

17.7 SUMMARY

Assessments can be an integral part of a staff headquarters at any level. Analysts conducting assessments in support of the Coalition effectively helped keep the CG and staff informed, tied strategic decisions and actions to impacts in the OE, challenged assumptions, and brought staff officers together to build a cohesive understanding of CJTF-OIR's operational environment which enabled more effective operations. By making use of the tools available and innovating ways to collect, analyse, and present information in an environment with limited access to data, the CJ-5 Assessments team positively affected operations and demonstrated the value that assessments can add to a staff headquarters.

17.8 AUTHOR'S BIOGRAPHY

Joseph Nowak works at United States Africa Command in Stuttgart, Germany, managing the campaign assessment process for the Command. He previously worked at the United States Army Training and Doctrine Command (TRADOC) Analysis Centre. He has served as an Army Civilian for 13 years. In his time, he has worked on multiple studies requiring analysis of complex systems using statistics, probability, decision theory, optimisation, and simulation. He has deployed twice: as an Operations Research Systems Analyst to Iraq with the XVIII Airborne Corps in the Corps Assessment Cell in 2008, and to Kuwait as the division chief of the Security Trends Division for Combined Joint Task Force – Operation Inherent Resolve in the Combined Joint Assessment Cell in 2015. Mr. Nowak has a Master of Science degree in Operations Research from Kansas State University.

Chapter 18 – ASSESSMENT FOR EFFECTIVE SECURITY COOPERATION EVENTS IN THE HORN OF AFRICA – JANUARY TO JULY 2013

Adam Shilling
Centre for Army Analysis
UNITED STATES

ABSTRACT

The author served as an assessment analyst in the Horn of Africa from January to July 2013. The following is a description of the operations assessment processes and products, and the author's perception of their impact on the command during that time period. The task force commander felt that he received the majority of the information he required through routine processes in the rest of the staff, and that he required little information from the assessment team. Therefore, the assessment team focused on assisting the mission commanders of security cooperation engagements – relatively junior soldiers – to deliver effective training for their own soldiers and for those of international partner states in the region. The lessons the author learned in the process of helping to make security cooperation operations more effective has led to his advocacy of doctrinal change currently occurring in the United States.

18.1 BACKGROUND

Established in October 2002, the Combined Joint Task Force – Horn of Africa (CJTF-HOA) originated as part of the American response to the 11 September 2001 terror attacks on the United States.

The CJTF has persisted in completing minor development projects – building or refurbishing schools or clinics, drilling or repairing water wells – and in collaborating with partner-national militaries to enhance the capabilities of both American and partner militaries. This collaboration, officially known as "security cooperation", and at the level of individual training events, commonly termed "engagements", is the subject of this chapter. The author served in CJTF-HOA from January to July 2013.

18.2 ORGANIZATION

The organization consisted of a headquarters staff to support operations, and an infantry battalion to provide security for Camp Lemonier, Djibouti, the home of the CJTF, and to provide mission personnel for various "engagements" with partner militaries in East Africa. A Civil Affairs battalion had a presence with small teams ranging throughout the joint operations area – most of East Africa – to plan and execute small-scale development projects to benefit local national populations. There were also small aviation and naval detachments, and American liaison officers detailed to the U.S. embassies in several partner nations. The command hosted several liaison officers from African partner states within the command's area of responsibility.

The assessment section resided in the J-3/5/7, which had responsibility for plans, operations, training/exercises, and lessons learned. It was under the "Effects Section", which also included the information operations section. The assessment section consisted of four personnel who assisted planners and led the staff in the production of assessment products.

18.3 MISSION

The CJTF's mission statement read: "Combined Joint Task Force – Horn of Africa, in partnership with our joint, interagency, intergovernmental, and multinational teammates, *conducts theatre security cooperation activities to enable regional actors* to neutralize violent extremist organizations and enables regional access and freedom of movement within East Africa in order to protect and defend United States interests. Be prepared to execute crisis response within East Africa in order to protect and defend United States military, diplomatic, and civilian personnel, facilities, and interests" (author's emphasis).

The bulk of the CJTF's activities were security cooperation. In practice, small teams set out to engage small groups of partner nation service members, or occasionally, selected local national civilians, such as health care workers or others that were part of crisis response communities. Another common mission was the training of partner-national instructors in demining tasks in accordance with international demining standards. The command averaged about two missions each week.

18.4 ASSESSMENT SCHEME

18.4.1 Description of the Framework and the Process

The Chief of Operations, a naval captain (OF-5) in the J3 Operations section, was in charge of the security cooperation engagement program. He was concerned that many missions were not achieving their potential, and that the command could not document the impact of their achievements. He approached the assessment section to enhance the quality of engagement-level assessment.

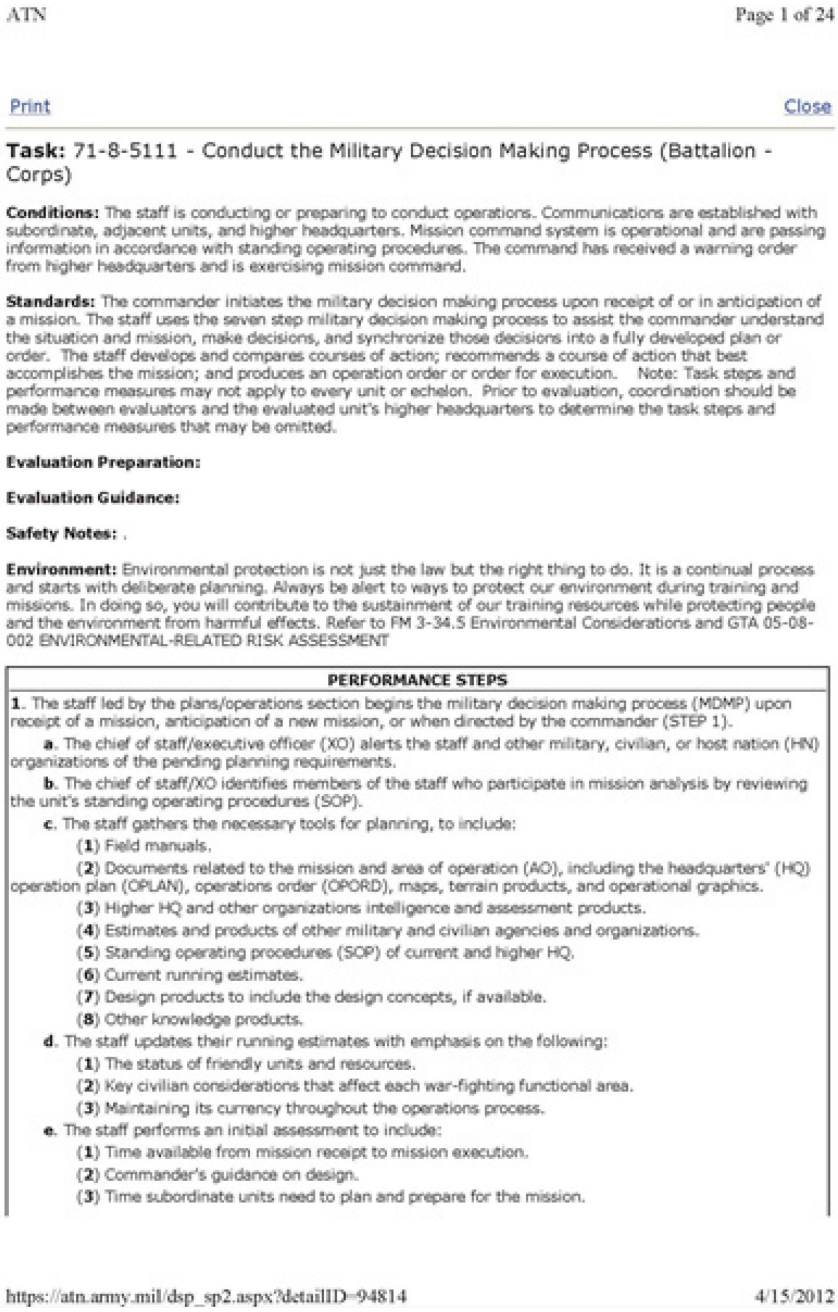
During this time, the assessment section was already assisting Mission Commanders (MCs), the service members assigned to lead these small engagement teams, to clearly articulate their mission objectives as three or four statements on a "desired outcomes" slide in a "confirmation brief" delivered to the Chief of Operations before he would approve the mission. Other items displayed during the brief were the team's mission statement and desired endstate, and a number of items that related to specific issues of travel and the team's safety while abroad in the partner nation.

The development of the "desired outcomes" slide was a collaboration between the mission commander, who had expertise in the mission type, and an assessor, who would help the MC clearly articulate what the mission was trying to accomplish. The command believed that this clear articulation made achieving those outcomes more likely and also made it easier to demonstrate the success of the mission through the assessment process.

The Chief wanted to improve this process by being more specific and by requiring some judgement from the mission commander on the capability of the partner nation's personnel. Indeed, ideally he wanted two observations – one that indicated the partner's capability before the engagement and one after.

The assessment section developed a five-point ordinal scale that leveraged published U.S. military training references, such as Training and Evaluation Outlines (T&EOs) for collective or unit tasks, and soldiers' manuals for individual tasks. A sample page from a T&EO is shown in Figure 18-1.

If U.S. military publications were not adequate for the engagement, MCs could substitute other references. This was commonly true for medical missions and always true for demining missions that were taught using a set of international standards. Assessors asked the MC to be as specific as possible and asked her to complete an "engagement plan" which listed the specific tasks to be trained, the reference which contained the standards for each task, and a rough schedule ensuring a realistic estimation of the time required to complete the training. An example is given in Table 18-1.



ATN Page 1 of 24

Print Close

Task: 71-8-5111 - Conduct the Military Decision Making Process (Battalion - Corps)

Conditions: The staff is conducting or preparing to conduct operations. Communications are established with subordinate, adjacent units, and higher headquarters. Mission command system is operational and are passing information in accordance with standing operating procedures. The command has received a warning order from higher headquarters and is exercising mission command.

Standards: The commander initiates the military decision making process upon receipt of or in anticipation of a mission. The staff uses the seven step military decision making process to assist the commander understand the situation and mission, make decisions, and synchronize those decisions into a fully developed plan or order. The staff develops and compares courses of action; recommends a course of action that best accomplishes the mission; and produces an operation order or order for execution. Note: Task steps and performance measures may not apply to every unit or echelon. Prior to evaluation, coordination should be made between evaluators and the evaluated unit's higher headquarters to determine the task steps and performance measures that may be omitted.

Evaluation Preparation:

Evaluation Guidance:

Safety Notes: .

Environment: Environmental protection is not just the law but the right thing to do. It is a continual process and starts with deliberate planning. Always be alert to ways to protect our environment during training and missions. In doing so, you will contribute to the sustainment of our training resources while protecting people and the environment from harmful effects. Refer to FM 3-34.5 Environmental Considerations and GTA 05-08-002 ENVIRONMENTAL-RELATED RISK ASSESSMENT

PERFORMANCE STEPS

1. The staff led by the plans/operations section begins the military decision making process (MDMP) upon receipt of a mission, anticipation of a new mission, or when directed by the commander (STEP 1).

a. The chief of staff/executive officer (XO) alerts the staff and other military, civilian, or host nation (HN) organizations of the pending planning requirements.

b. The chief of staff/XO identifies members of the staff who participate in mission analysis by reviewing the unit's standing operating procedures (SOP).

c. The staff gathers the necessary tools for planning, to include:

(**1**) Field manuals.

(**2**) Documents related to the mission and area of operation (AO), including the headquarters' (HQ) operation plan (OPLAN), operations order (OPORD), maps, terrain products, and operational graphics.

(**3**) Higher HQ and other organizations intelligence and assessment products.

(**4**) Estimates and products of other military and civilian agencies and organizations.

(**5**) Standing operating procedures (SOP) of current and higher HQ.

(**6**) Current running estimates.

(**7**) Design products to include the design concepts, if available.

(**8**) Other knowledge products.

d. The staff updates their running estimates with emphasis on the following:

(**1**) The status of friendly units and resources.

(**2**) Key civilian considerations that affect each war-fighting functional area.

(**3**) Maintaining its currency throughout the operations process.

e. The staff performs an initial assessment to include:

(**1**) Time available from mission receipt to mission execution.

(**2**) Commander's guidance on design.

(**3**) Time subordinate units need to plan and prepare for the mission.

https://atn.army.mil/dsp_sp2.aspx?detailID=94814 4/15/2012

Figure 18-1: Sample Page from a Training and Evaluation Outline (T&EO).

Table 18-1: Sample Engagement Plan.

Task #/ Reference	Task Name or Terminal Learning Objectives (TLOs)	Date	Before Evaluation (1-5)	After Evaluation (1-5)
41-7-1224	Plan Civil Reconnaissance	01-Jul-13	3	
41-6-6003	Conduct Civil Reconnaissance	01-Jul-13		
41-7-1215	Plan Civil Information Management	02-Jul-13		
41-6-6001	Conduct Civil Information Management	02-Jul-13		
41-6-6002	Integrate Civil Information into the Supported Element Common Operational Picture	03-Jul-13		
71-8-2210	Perform Intelligence Preparation of the Battlefield (Battalion - Corps)	03-Jul-13		

The engagement assessment plan required the MC to plan hands-on activities so that she could observe the partner's service members performing the task alongside hers. To meet the Chief's desire for a before and an after observation, the assessor coached the MC to evaluate the performance of the group at the beginning of the hands-on activity, and to evaluate again as that multinational group completed the hands-on activity. To minimise observer error, the assessment section provided an ordinal scale with definitions to assist the MC in her assessment. This scale is at Table 18-2. Assessors also coached MCs to be unobtrusive in making their observations so that partner service members would not feel as if they were being graded.

Table 18-2: Five-Point Ordinal Assessment Scale for Collective or Individual Tasks.

Evaluation	"Letter Grade"	Numeric Score	Description
Untrained	U	1	Partner soldiers are unfamilar with the task
Needs Practice	Weak "P"	2	Partner soldiers can perform the task only with significant assistance
	Solid "P"	3	Partner soldiers can generally perform the task, perhaps not quite to standard; may require assistance
	Strong "P"	4	Partner soldiers can usually perform the task to standard; may require a little assistance
Trained	T	5	Partner soldiers can perform the task to standard, consistently, without assistance

The scale was a refinement of an ordinal scale common in the U.S. Army, where unit activities, known as "collective tasks", are evaluated by unit leaders or other observers as Trained, Needing Practice, or Untrained. These scores are abbreviated T, P, or U, respectively. The problem with this three-point scale (T, P, U) for assessment purposes, is that it typically creates a small band of U for untrained units, a small band of T for units performing completely to standard, and a wide band of P for units in between. Therefore, assessors broke the wide band of P into three parts – weak P, solid P, and strong P. For simplicity, they enumerated the five points of this scale as one through five, with five being best.

As part of the mission commander's preparation for the engagement mission, he completed a checklist of the staff agencies he needed to visit for a successful and safe mission. When the MC came to the assessment section, an assessor would take about one half hour to coach him on the engagement plan, the assessment plan, After Action Review (AAR) requirements, and help him prepare the "desired outcomes" slide.

18.4.2 Assessors' Relationships with the Rest of the Staff

Over the course of the author's tenure, the assessment section became more integrated into the staff's daily function, and less a separate section that only handled taskings labelled with the word "assessment". Ultimately, assessors became key players in the planning of security cooperation missions by helping the mission commanders articulate their desired outcomes, develop an engagement plan, plan a hands-on activity, prepare to observe partner nation service members at work, and articulate the results of their observations in a standardized way.

The assessment section also led After-Action-Report (AAR) interviews and helped MCs prepare required post-mission lessons learned products. Assessors found that mission personnel were willing to provide more information verbally than they did in writing, and the semi-structured format of the interview allowed assessors the freedom to explore at greater depth issues that mission personnel mentioned during the interview. Assessors also extracted action items from the interviews that required coordination with the staff to resolve. For example, in February 2013, mission commanders were reporting difficulty contacting the Joint Operations Centre (JOC), the operations command post, in the evenings to deliver their accountability reports. The JOC's representative to the AAR interview

committee investigated, and determined that JOC personnel were setting their phones on silent mode during afternoon briefs and not turning them back on consistently. He instituted corrective measures, and mission commanders' complaints on this issue dropped to zero.

Functional area planners (e.g., medical, civil affairs) and "regional engagements" staff members, also known as "country desk officers", developed mission opportunities, received required approvals, and arranged funding for security cooperation missions. When they prepared to pass a planned mission off to a mission commander for detailed planning, they consulted with assessors to ensure draft mission outcomes were articulated well. Before the mission was assigned to a mission commander, the staff wanted to have a solid draft of the mission's objectives to guide the MC. The mission commander was expected to add detail to the plan and complete the articulation of its specific outcomes in cooperation with the assessment section.

The plans section, usually occupied with planning activities other than security cooperation, learned that assessors could also help them be specific in articulating outcomes. Increased specificity enhanced the ability of the command to assess the success or failure of these activities, but more importantly, the increased specificity of operational outcomes statements – mission statements, endstates, and tasks and purposes – also made it easier for tasked unit leaders to understand what was required of them, and therefore, the specified outcomes were more likely to be achieved, meaning the command's operations became more effective.

18.4.3 Data Required

For security cooperation engagements, key data included the manpower and fiscal cost of the operation, its results measured on the five-point scale, and a qualitative description of the MC's perception of the impact of the mission on the relationship of the partner nation with the United States. Assessors also asked a standard series of questions during the AAR to identify problems, threats to safety, best practices, and things that would enhance the effectiveness of future missions. When MCs could do so unobtrusively, they captured the names and ranks of key participants in the engagement or influential partner-national visitors to the engagement. Names and ranks of participants might help build the next, more advanced class, and high-ranking visitors were indicative of the value the partner nation placed on the engagement. Also, today's junior officers participating in an engagement are tomorrow's senior officers, and identifying emerging leaders could help to gain some form of access to the partner nation in the future. The bulk of these data was captured formally during the AAR interview.

18.4.4 Data Gathering and Storage

Data storage at CJTF-HOA was rudimentary at the time. Originally, data were captured by mission personnel answering a questionnaire on PowerPoint slides that served as the basis of the semi-structured AAR interview. The interview was found to be superb at eliciting information from mission commanders, but the capture of information for subsequent use or analysis was inadequate.

The data included some quantitative metrics, such as how much money was spent, how many people trained, and things of that nature; some questions were answered yes/no with a narrative; and other questions required a qualitative description.

During this time, the command was attempting to institute a non-secure internet instance of the Combined Information Data Network Exchange (CIDNE), a database that would be used to capture engagement information and other information important to the command. CIDNE is the database of record for allied efforts in Iraq and Afghanistan, and CJTF-HOA hoped to leverage the known capabilities of CIDNE to capture operational data so that they are available for analysis. At the end of the author's tenure, AAR interviewers captured data in Excel tables with the expectation that CIDNE would be able to ingest the data in the near future. This did occur about eight months after the author left the command.

18.4.5 Analyses Conducted

The analysis of AAR information and engagement-level assessment data required only simple descriptive statistics and data visualization or critical thinking. There was no need and little opportunity for more complex mathematical techniques, nevertheless experienced assessors, planners, and mission commanders managed to extract a lot of value from this analysis – things that enhanced the effectiveness of the next mission going out and the safety of mission personnel. Moreover, the early involvement of assessment personnel in mission preparation improved the mission plan, and its subsequent execution. For example, one MC shared how much individual Africans valued a certificate for any training they completed. They perceived such a certificate as a real credential that made them more employable. It also incentivized their attendance and wholehearted participation in learning activities and further cemented a relationship with American colleagues, making the mission more successful. Assessors began encouraging MCs to devise a certificate to award to their African participants.

18.4.6 Products Required

In addition to proofing the mission commander's endstate, helping her develop the "desired outcomes" slide, and coaching her on the engagement and assessment plans, the assessment section also assisted MCs in the preparation of three AAR products. One was prepared according to U.S. Africa Command (USAFRICOM) format, which was a short report that captured fiscal cost data and three key takeaways or lessons learned. This report was later attached to the entry for that mission in the Theatre Security Cooperation Management Information System (TCSMIS), a database that captures information on proposed and executed security cooperation missions worldwide.

CJTF-HOA, however, judged the information captured in this format inadequate to best enhance the effectiveness of future security cooperation engagements. Thus, the task force required a more robust brief/interview where MCs would answer a questionnaire and then discuss it with assessment personnel and other staff members. AAR information was initially captured in PowerPoint as described above, then in Excel in anticipation of CIDNE coming online, and finally, AARs were input directly into CIDNE beginning about eight months after the author left the command. The Chief of Operations considered the mission incomplete until the AAR was complete.

The MC also provided a short executive summary brief to the weekly engagements synchronization meeting which was attended by 40-50 personnel. He showed his mission, "desired outcomes", a slide showing the fiscal cost of the mission, and the plan of engagement from the confirmation brief. The last was updated with the MC's five-point scale observations on each task practised, as in Table 18-2 above. The final slide was the MC's key lessons learned, which were the same as those reported on the USAFRICOM AAR. This process sounds more complicated than it was. The AAR interview process eased the preparation of this summary and preparation of the other products became trivial.

In addition to the engagement-level assessment of security cooperation activities and related products just described, the assessment section led the staff through a semi-annual assessment of the campaign plan and the production of the associated report.

The assessment section conceived of "the assessment" as the improved understanding of the operational environment, and the command's place in it, that the staff obtained by working its way through the assessment process. "The assessment" did not refer to the assessment report. The report was the device used to communicate the important parts of the improved understanding of the environment ("the assessment") to decision makers.

Therefore, this assessment, and the report that communicated it, was organized by Lines Of Effort (LOEs). The assessment consisted of the analysis of data related to campaign progress and the determination of the

corrective action needed to be more effective. The report indicated the status of the command relative to campaign goals and provided to senior leaders recommendations to improve the effectiveness of operations in East Africa.

This assessment framework suffered from a host of methodological problems, but it still served as a foundation for relevant conversations among the CJTF staff during the preparation of the report. Truthfully, the author, based in Djibouti, was unable to determine if the report subsequently drove relevant conversations among senior leaders, either in Djibouti or at USAFRICOM headquarters in Stuttgart, Germany.

18.5 IMPACT ON THE COMMAND OF ASSESSMENT

The engagement-level assessment process improved the quality of the training that partner nation and American soldiers shared, which the command believed improved learning outcomes for all. It also improved the ability of the command to document the progress made during an engagement, which led to the ability to build upon expertise demonstrated in previous engagements, provided the same groups of soldiers met again, or facilitated the handover of a class of partners from one American set of trainers to another. An example of the latter was the training of partner-national soldiers in demining operations by U.S. Explosive Ordinance Disposal (EOD) technicians. The EOD units rotated out, but they kept excellent records on what was done, so that their successors could build upon their progress. These units kept a "roll book" with individual partner-national soldiers' names and a list of the tasks they had trained, with scores, in anticipation of the same trainees being in the subsequent, more advanced class. This, in turn, helped the partner nation build capacity in demining operations. With sufficient training, these partner soldiers will become the trainers for additional classes of their countrymen.

The assessment process also included a robust AAR interview which captured a lot of lessons that aided the command internally to improve its processes. It highlighted both mistakes to avoid and best practices to follow for subsequent missions. An example of a best practice, brought to assessors' attention by one of the maritime civil affairs units, was the use of printed certificates to reward students – military or civilian – in classes that enhanced the safety of fishermen and other local mariners. They found that Africans perceived a certificate as a real credential that might aid them in the search for a job. Other civil affairs teams confirmed this finding, and assessors began to encourage mission commanders to develop and produce certificates of completion for each trainee who participated in engagement and training events.

18.6 LESSONS LEARNED

The following ideas are a compilation of what the author deemed as the primary lessons learned while serving in the assessments office in CJTF-HOA from January to July 2013:

- **The primary purpose of assessment is to make operations more effective.** In the past, we said assessments were to "measure progress" and "inform decision making". Today, we recognize those things as being instrumental to making operations more effective. A secondary purpose of assessment is to provide appropriate accountability to higher authority.
- **Assessment is a process, not a product**. It is a structured attempt to better understand our Operational Environment (OE), and select future activities that will achieve our desired outcomes most effectively. It is not a structured attempt to produce a report.
- **Therefore, the product of assessment is an effective operation, not a slideshow**. It is a conceptual error to think of "the assessment" as a report. Actually, "*the assessment*" *is the improved understanding of the OE we get from working through the assessment process*. It is a cognitive product, too complex to be written readily, that is analogous to a running staff estimate. The report is

an executive summary of this understanding at a point in time, and should be viewed as a communication device only. Its format is not relevant, and can be changed to meet bureaucratic requirements, provided it communicates the information decision makers need to select activities for our forces that achieve our goals most effectively.

- **Assessment is continuous, not periodic.** It is not performed solely in advance of a report deadline, but it is a continuous monitoring of the OE and a continuous attempt to determine the meaning of the observations we make. A game-changing discovery does not wait for the next assessment report or conference, and need not be broached to the commander by the assessment section.

- **Assessment is not new. We have always done it.** We have always collected information, and always made sense of it. The difference is that we better appreciate the complexity of the current OE.

- **Assessment is everybody's business.** Operations researchers/assessors lack the expertise to make best sense of a lot of data. Assessors help functional area experts on the staff make better observations and improve the rigor of analysis. Neither assessors nor functional area specialists should assess alone; it is a whole-of-staff process.

- **Assessment is a part of, and provides focus to, planning.** Assessors can help the planners better articulate our desired outcomes (*an outcome statement too unclear to assess is too unclear for subordinate commands to take effective action to achieve*). Also, assessment is an activity like any other. To be performed well, it must be planned.

- **Assessment is a part of, and provides timely feedback to, operations.** It is a part of operations because that is where the data come from. Some operator sees something and reports it (which implies we must tell them what to report). And providing feedback to operations is the purpose of assessment.

- **The important question is not *what* the metrics are doing; the important question is *why* they are doing it.** We must attribute change in the OE, reflected in our metrics, to the right causes in order to know what we should do next to be most effective. We cannot duck causation by muttering about correlation.

- **Metrics alone are not sufficient to describe interactions in a complex environment.** Observing a metric to perceive change is not the same as understanding the OE and the changes in it. We must use human judgement to make sense out of the changes we see in our set of metrics. *The numbers don't tell the story; their purpose is to keep the storyteller honest!*

18.7 ADVICE TO NEW ASSESSORS

The most important thing to remember about assessment is that it is not a backward-looking report on the successes (or failures) of a completed operation. It is the forward-looking attempt to make all future operations more effective. It begins ideally at the start of mission planning and persists until operations conclude. Its first task is to help the planner get the plan right, and then assessment monitors key factors of the operating environment until operations end, making course corrections and learning lessons in order to be more effective.

Assessment's primary product, therefore, is not a slideshow or report, but effective operations that best achieve the command's goals.

18.8 AUTHOR'S BIOGRAPHY

Adam Shilling, PhD, is a veteran of two operations assessment tours abroad – one to the Afghan Assessment Group at Headquarters, International Security Assistance Force (ISAF), and one at the Combined Joint Task Force – Horn of Africa. He spoke about his experiences and lessons learned at an international meeting. His ideas made sense to analysts and decision makers, and so his superiors selected him to represent his organization, the Centre for Army Analysis, in the community that was writing operations assessment doctrine for the United States military, and to act as senior editor for this report.

Chapter 19 – ANALYTICAL SUPPORT TO COMBINED JOINT TASK FORCE – HORN OF AFRICA – MAY 2014 TO MAY 2015

Vincent Boncich
US Army
UNITED STATES

Natalie Casey
US Army
UNITED STATES

ABSTRACT

Established in October 2002, the Combined Joint Task Force – Horn of Africa (CJTF-HOA) has been supporting Operation Enduring Freedom, a part of the U.S. response to the terror attacks on 11 September 2001. Beginning in October 2011, the Centre for Army Analysis (CAA) deployed analysts to Djibouti, Africa on six month rotations to support CJTF-HOA.

In the current environment of increasing budget constraints, CJTF-HOA is placing renewed emphasis upon prioritization of Operations, Actions, and Activities (OAA). One of the most common types of OAA is that of military-to-military training and advising, such as pre-deployment assistance to African countries who are members of the African Union Mission in Somalia (AMISOM) in preparation for deployments to Somalia. CJTF-HOA has recently undertaken an effort to identify, assess, and prioritize its OAAs so that it can better align operations and allocate resources accordingly.

Analytic efforts include: integration of operations assessments into the command decision-making process; development of an operations assessment framework for the campaign plan; analysis of the level of impact from OAAs; analysis of SIGACT data from the AMISOM Force Headquarters (HQ); and analysis of Public Perception Surveys. The insights from these analytic efforts support planning and command decisions for current operations conducted not only by CJTF-HOA, but also by U.S. partners in East Africa.

19.1 BACKGROUND

In 2002 The Combined Joint Task Force – Horn of Africa (CJTF-HOA) began operating in support of Operation Enduring Freedom; the only CJTF in the U.S. Africa Command (USARFICOM) Combined/Joint Operations Area (CJOA). CJTF-HOA was established on 19 October 2002 in response to the attacks on 11 September 2001. CJTF-HOA was originally a component of U.S. Central Command (USCENTCOM), but was transferred to USAFRICOM on 1 October 2008. Beginning in October 2011, the Centre for Army Analysis (CAA) deployed analysts to Djibouti, Africa on six month rotations to support CJTF-HOA.

In the current environment of increasing budget constraints, CJTF-HOA is placing renewed emphasis upon prioritization of Operations, Actions, and Activities (OAA). One of the most widespread types of OAA conducted by CJTF-HOA is that of military-to-military training and advising, such as pre-deployment assistance to African countries who are members of the African Union Mission in Somalia (AMISOM) in preparation for deployments to Somalia. CJTF-HOA has recently undertaken an effort to identify, assess, and prioritize its OAAs so that it can better align operations and allocate resources more effectively.

19.2 ORGANIZATION

This chapter pertains to the time period of May 2014 to May 2015; the authors' two successive six month deployments to Djibouti in support of CJTF-HOA. The CJTF-HOA combined/joint operations area

(Figure 19-1) included the East African countries of Eretria, Djibouti, Ethiopia, Uganda, Kenya, Somalia, Rwanda, Burundi, and Tanzania.

Figure 19-1: CJTF-HOA Combined/Joint Operations Area (CJOA).

The CJTF consisted of a mix of members from all service branches, both active and reserve, and consisted of approximately two thousand service members. Within the J3/5/7, the assessment branch consisted of a five-member team led by an OF-3. CJTF-HOA is the Department of Defence (DoD) representative in East Africa and plays a supporting role to the ambassadors or Chief Of Mission (COM) representing the Department of State in each country. This relationship is further delineated by the congressional authorities and appropriations that each department receives to conduct their respective missions. Title 10 of the United States Code (U.S.C.) describes the role of the armed forces. It provides the legal basis for the roles, missions, and organizations of each of the services as well as the overall DoD.[1] In support of DoD's Title 10 U.S.C. responsibilities, CJTF-HOA conducts Theatre Security Cooperation (TSC) activities which support objectives such as building partner capacity by improving host nation military leadership and readiness, strengthening partner relationships, and gaining and maintaining access to land, sea, airspace, or facilities within partner nations' territory that are necessary to accomplish the CJTF's mission. "Access" includes not only physical presence, but also political and diplomatic relationships with respective leaders within these countries.

During the late summer of 2014, the CJTF-HOA staff reorganized to improve both internal and external communications and become a faster, smarter, and more versatile organization. The greatest change was the expansion of the Regional Engagements Branch from one single desk officer to a more robust Fusion Action Cell (FAC) consisting of a 6-8 person team (Figure 19-2) consisting of a team lead, country planner, country analyst, Sociocultural Research And Advisory Team (SCRAT) member, a liaison officer located in the country embassy, a foreign military liaison from the host country, and a U.S. Foreign Area Officer (FAO).

The FACs were reorganized under a new TSC Directorate who managed approximately 100 service members to include members of the East African militaries from several countries in the Combined/Joint Operations Area (CJOA).

[1] Office of the Law Revision Counsel, United States Code: http://uscode.house.gov/browse/prelim@title10&edition=prelim.

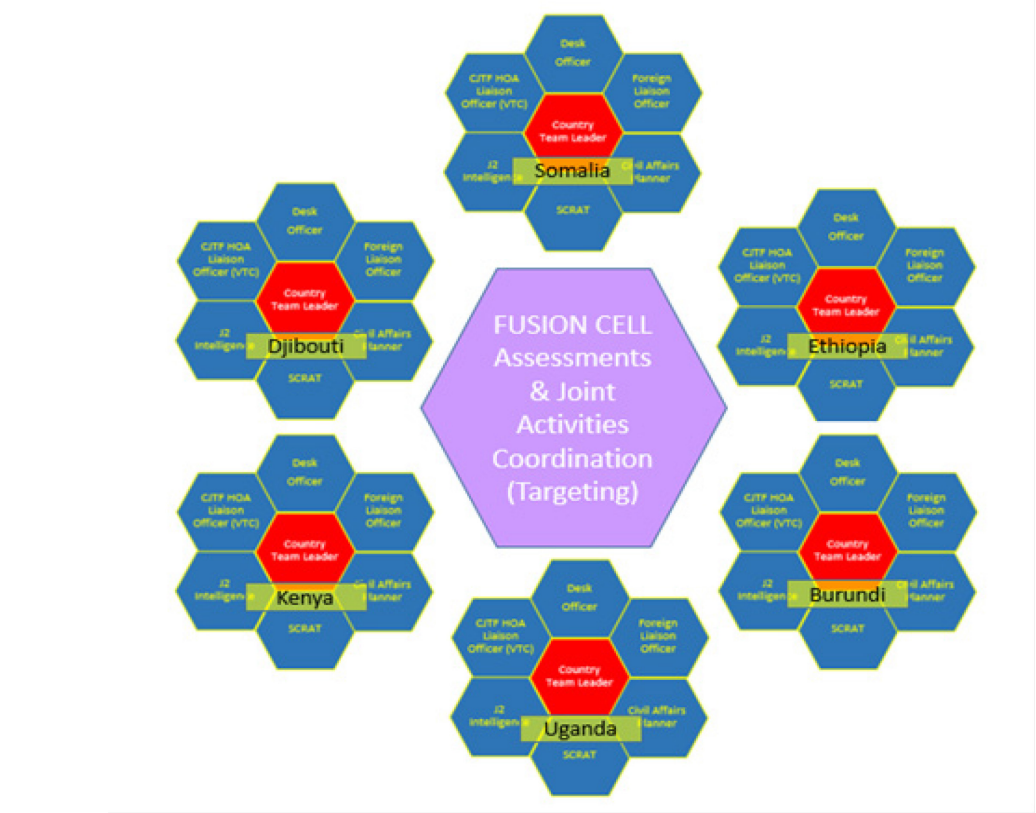


Figure 19-2: Theatre Security Cooperation Directorate FAC Structure.

19.3 ASSESSORS' RELATIONSHIP WITH THE STAFF

Upon arrival to Djibouti in May 2014, the assessment branch was a five-member team consisting of both military officers and DoD civilians. Kinetic operational assessment work was primarily focused on tactical operations in Somalia, while non-kinetic assessments related to military training and capability were being conducted on in Kenya, Uganda, Ethiopia, Burundi or Djibouti. Non-kinetic assessment consisted primarily of After Action Reviews (AARs) based on the TSC planned activities. The CJTF-HOA Campaign Plan articulated a concept of the operation and desired endstate for countries in the CJOA, but did not contain specific objectives for each country or an assessment annex with defined metrics to measure effectiveness of operations. As a result, the demand signal for assessment products was primarily generated from USAFRICOM and not the CJTF-HOA staff. By January 2015, the assessment branch was downsized to only two people, once military officer and one DoD civilian.

With no requirement to assess the CJTF-HOA Campaign Plan, the assessment branch, which was not integrated as a part of the FAC, had a difficult time collecting data outside of the AAR process. With very little command emphasis placed on connecting individual country objectives to TSC activities to help determine desired TSC outcomes, the staff understandably had little interest in making time to implement data collection processes to inform an operations assessment. During the initial weeks of the deployment in 2014, several attempts were made by the assessment branch to establish formal data collection procedures to gather Measures Of Performance (MOPs) and Measures Of Effectiveness (MOEs). Prior to 2014, any element departing CTJF-HOA would be required to check in with the assessment branch before departure to review their plan and develop an information collection plan to support country level operations assessments. Despite these efforts, the absence of a command emphasis on operations assessment made it very difficult to mandate any procedures and enforce them accordingly.

19.4 MISSION

The CJTF-HOA intent of how to carry out TSC activities is best described in the Commander's mission statement:

> *CJTF-HOA, in partnership with our Joint, Interagency, Intergovernmental, and Multinational (JIIM) teammates, conducts theatre security cooperation activities to enable regional actors to neutralize Violent Extremist Organizations (VEOs) and enables regional access and freedom of movement within East Africa in order to protect and defend Unites States interests. Plan, prepare, and on order execute crisis response within East Africa in order to protect and defend United States military, diplomatic, and civilian personnel, facilities, and interests.*

The CJTF-HOA mission statement articulates the four Lines Of Effort (LOEs): develop and strengthen the Joint, Interagency, Intergovernmental, and Multinational (JIIM) team; enable regional actors to neutralize VEOs; enable regional access and freedom of movement; plan, prepare, and conduct crisis response.

CJTF-HOA's mission supports the African Union Mission in Somalia (AMISOM) mission that began in January 2007. The primary mission of AMISOM is to support dialogue and national reconciliation between all political, clan, and religious leaders within Somalia and support the creation of a new Somalian government.[2] AMISOM has military, civilian, and maritime components. The African Union authorizes the member states of Uganda, Burundi, Kenya, Djibouti, and Ethiopia to maintain the deployment of military troops in support of AMISOM's mission; CJTF-HOA refers to these member states as Troop Contributing Countries (TCCs). CJTF-HOA conducts TSC activities within these TCCs to help prepare these militaries for combat operations.

19.5 INITIAL FRAMEWORK AND ASSESSMENT SCHEME

While plans are written at both the strategic and operational level, execution of TSC activities was conducted at the tactical level. Figure 19-3 represents an initial assessment scheme at CJTF-HOA in the summer of 2014. Prior to any service members leaving Camp Lemonnier, Djibouti, to conduct TSC activities, the planned activity was briefed at a biweekly Joint Activity Coordination Board (JACB) chaired by the Commander. The TSC activities were not always linked to operational objectives or long-term outcomes, making them stand-alone events. When teams returned from their missions, AARs were conducted with a tactical focus on describing progress, but did not always provide data to help understand if greater goals were any closer to being achieved.

TSC activities were performed for one of two reasons; either they were a logical continuation of activities conducted in the past, or the activity aligned well with CJTF-HOA's authorities and appropriations. With the absence of defined operational objectives for each country, almost any activity could arguably support the broad mission and endstate for each in country in the campaign plan.

19.5.1 New Assessment Framework

During the months of May 2014 – November 2014, a new assessment framework (Figure 19-4) was adapted from a methodology used by the Afghanistan Assessment Group (AAG) 2012 [1] to help link TSC activities to country level objectives and targeted effects. To gain perspective over time, each of the CJTF-HOA LOEs were assessed quarterly by a rubric using a 1-5 rating scheme (1: LOE not being achieved to 5: LOE has been achieved). The rating was paired with summarized key insights collected over the 90-day assessment period to inform the Commander on progress towards achieving the LOEs in each county in the CJOA.

[2] African Union Mission in Somalia: http://amisom-au.org/.

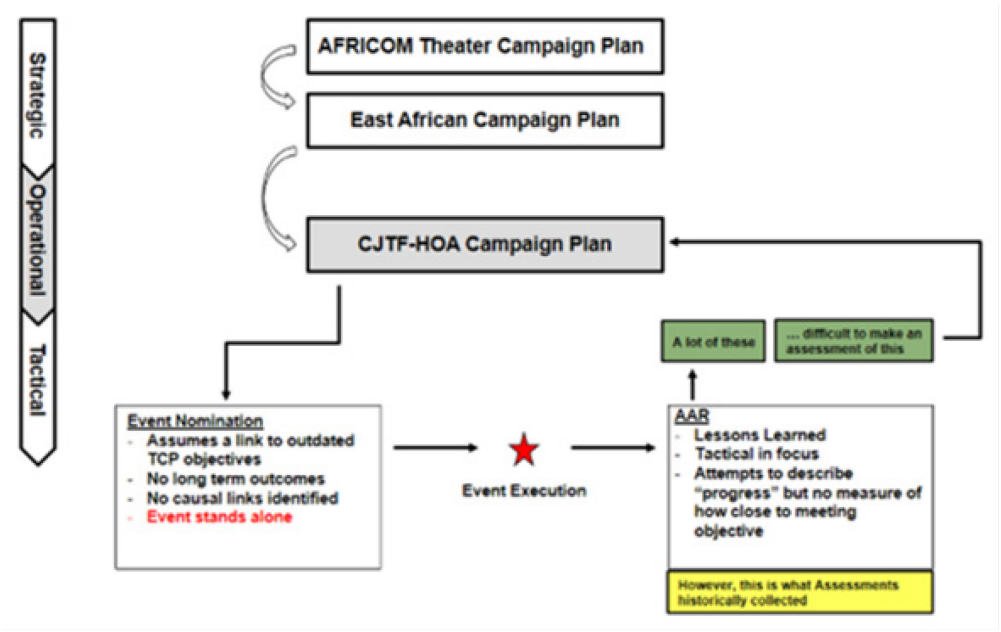


Figure 19-3: Initial Assessment Scheme.

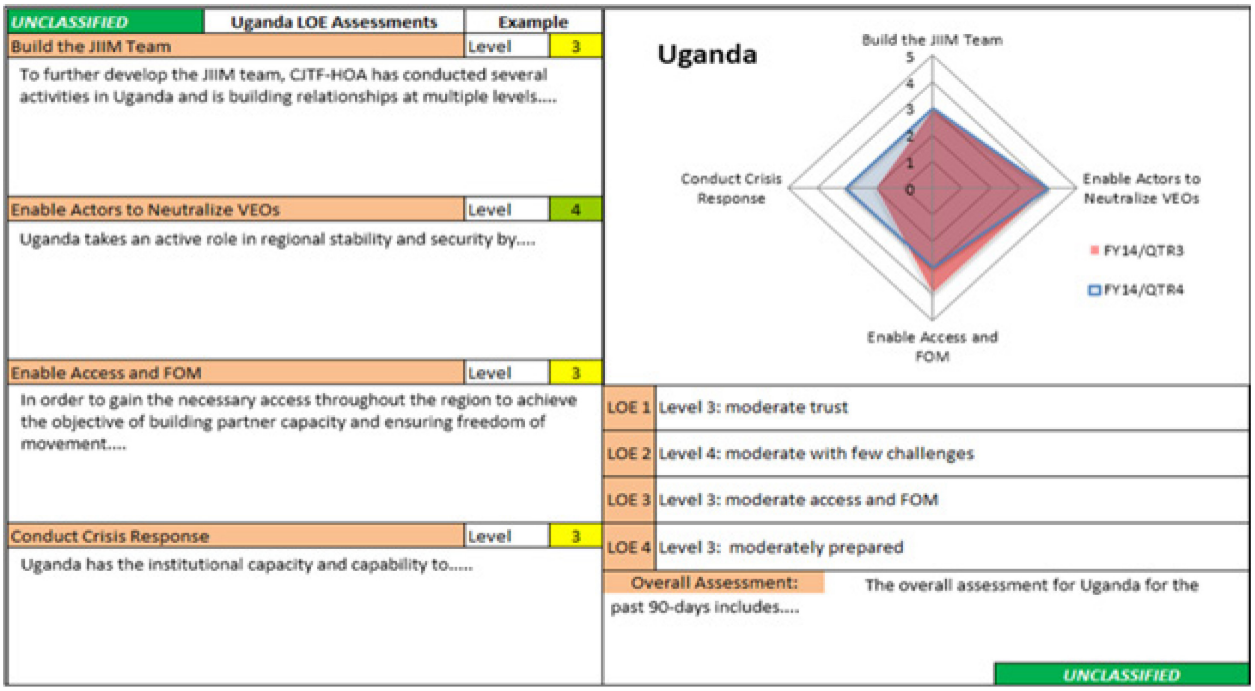


Figure 19-4: New Assessment Scheme.

Staff subject matter experts rated the country they were concerned with. Country team leads would write a qualitative assessment for their country and use that assessment to substantiate the 1-5 rating they provided the country for each LOE. Key leader engagements and training soldiers from troop contributing countries were the main ways CJTF-HOA worked to achieve those objectives. There was also a substantial effort to rewrite country level objectives to ensure they were in line with the lines of effort contained in the East African Campaign Plan and the Theatre Campaign Plan. The staff also assessed legacy activities that were not in line

with current plans and objectives. These assessments were tailored specifically to the activity. They explored the resources required, the effectiveness of the activity, and recommendations for continuation.

19.5.2 Data Collection and Storage

All analysis of ongoing operations requires structured data in order to answer questions, formulate assessments, and identify significant trends, correlations, and possible causations. Most importantly, structured data enables informed analysis which in turn informs the Commander's decision making and helps make operations more effective. The assessment branch leader provided summary statistics and graphs in the quarterly assessment by leveraging the Combined Information Data Network Exchange (CIDNE), the CJTF-HOA database of record for operational data. The assessment team also uncovered other sources of data to provide awareness to the command. Open source information, Public Perception Surveys, after action reports in Power Point and on Microsoft Word documents, and other existing spreadsheets and databases all provided valuable information that collectively helped add detail and context to the assessment. Storage for most of the data was on the shared drives or on SharePoint unless it was input into CIDNE, which is web-based. Data existed on both classified and unclassified networks.

19.5.3 Types of Analysis with Indicators

A variety of indicators were used to answer specific questions from the Commander or staff. One example of analysis conducted during the CJTF-HOA quarterly assessment process sought to determine the level of impact TSC activities were having in the CJOA and to gauge the professional development of the TCCs' military leadership. A key indicator to measure impact was the number of activities conducted in each of the TCCs and the level of participation by each military rank category. Categorisation of the number of TSC activities included; low (1-10), medium (11-20), high (21 or greater). Categorisation of the TCC's military rank structure included: lower enlisted (E1-E4), non-commissioned officers (E5-E9), company grade officers (O1-O3), field grade officers (O4-O6). This analysis allowed the Commander and staff to target future TSC activities and gauge their effects on the objective to professionalize TCC militaries.

Another assessment examined AMISOM's effectiveness against IEDs in Somalia based on pre-deployment TSC activities of TCC units. Key indicators to measure effectiveness were the number and type of attacks that were conducted against TCC military units and the number of casualties during those attacks. The command was interested in determining if the pre-deployment TSC activities pertaining to counter-IED were effective. By analysing the types of attacks over time, it enabled the assessment branch to establish trends to help better understand how to adjust TCC military training prior to deployment to Somalia to be more effective.

An additional example of analysis conducted as part of the assessment explored the degree of information sharing between CJTF-HOA's Joint Operations Centre (JOC) and the TCCs' Tactical Operations Centres (TOCs). A key indicator included the number and type of information transmissions measured via communication equipment given to each TCC headquarters. This allowed the CJTF-HOA command to gain insight on the specific types of information that were most useful to the TCC military leaders and also identified which TCCs were leveraging shared communication capabilities.

A final example explored AMISOM's application of the principles of Civil-Military Cooperation (CIMIC) in combat. CIMIC principles are the means by which a military commander connects with civilian agencies and leaders in the theatre of operations. A key indicator used to measure the application of CIMIC principles included an analysis of public perception survey data collected in Somalia. Analysis of the data allowed the assessment branch to provide the command with insights on the public's view toward various topics related to foreign military assistance, trust in local police and government officials, and the professionalism of the AMISOM troops in Somalia. This enabled CJTF-HOA to tailor TSC activities related to CIMIC operations to be more effective.

19.6 ASSESSMENT IMPACT

The new assessment framework supported a quarterly assessment brief to the Commander from both the assessment branch leader and the FAC team leaders. Prior to this, the only assessments conducted by CJTF-HOA were USAFRICOM-directed and were more focused on specific operations rather than a holistic assessment of all TSC activities. The new assessment framework and subsequent quarterly assessments helped drive operations. Previously each country lead would plan TSC activities independently, while the targeting officer briefed activities for all countries at the JACB. The new quarterly assessment helped change the old paradigm by requiring each FAC leaders to brief their respective proposed TSC events in a context that was driven by assessments and linked to desired effects and country level objectives. By having action officers for specific countries provide their country level objectives, it allowed the assessment staff to synthesize that information and deliver a synthesized regional assessment for the area of interest in East Africa.

During the months of December 2014 to May 2015 the staff produced a new Horn of Africa Supporting Plan that replaced the HOA Campaign Plan. Although CJTF-HOA staff made several attempts, the plan was not synchronized with the release of AFRICOM's Theatre Campaign Plan, which resulted in many gaps and ultimately required a rewrite of the plan about eight months later. With changes in LOEs, priorities and objectives, it was imperative that a new assessment methodology be developed and that assessment working groups and updates be worked into the battle rhythm. When conducting a baseline assessment, it became evident that the targeting process and a lot of CJTF-HOA's resources were tied up in legacy activities to include embassy requirements and operations that were funded or programed over a year prior to the individual activity's start date. The Venn diagram (Figure 19-5) below was developed to depict what component was conducting what operations and who those operations supported. A colour-coded "slant chart" – named for the diagonal slashes – indicated who was conducting the mission. For instance, many of the civil affairs units taught English to Djiboutians. This type of activity did not support the Theatre Campaign Plan nor did it support the HOA Supporting Plan directly, yet it was critical to maintaining a relationship with the ambassador and the Djiboutians. It is in direct support of the Department of State's integrated Country Strategy.

The Venn diagram with a colour-coded slant chart by component would provide a common operating picture for East Africa and would visually communicate who and what DoD efforts in this area of operations are supporting. This type of visualization informs leadership on resource allocation and can be a planning tool that can help identify both risk and opportunity.

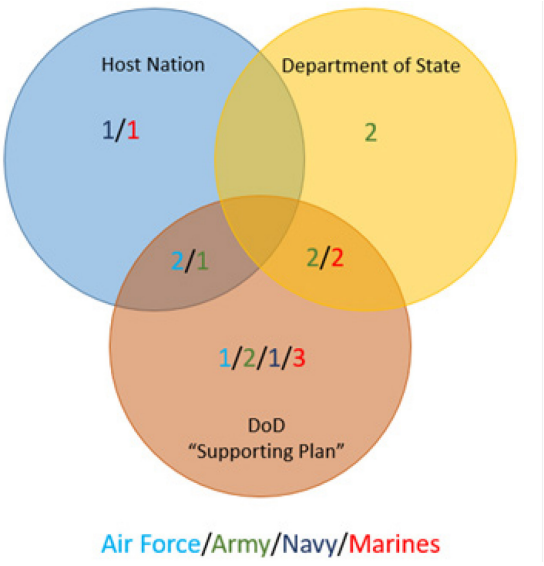


Figure 19-5: Component Venn Diagram.

19.7 LESSONS LEARNED

There were several lessons learned as the plan changed and the commanders changed. It is critical that the assessment framework can stand the test of time. The 1-5 ratings provided for each LOE (Figure 19-4) was a good start at establishing an operations assessment framework for communicating the staff's understanding to leadership, but the ratings were ultimately subjective. As leaders changed over time, a longitudinal look of how the ratings changed provided less insight than was originally anticipated. In addition, changes in leadership resulted in different opinions on how to define the rating scheme and introduced debate questioning if LOEs were the best way to assess operations. Changes in leadership and opinions are unavoidable, yet being more cognizant of this during the initial design and implementation of the assessment framework may have resulted in a different design that considered ways to mitigate these changes. Considering the commander's decision space and understanding their authorities and responsibilities allow the assessment team to assist the staff in developing indicators that can stand the test of time. Authorities and responsibilities are unlikely to change even as people and the plan change.

Assessors need to get buy-in from the commander and staff early when developing an assessment framework. The command's lack of understanding and appetite for operations assessment required a substantial effort by the assessment branch to promote its importance and value in enhancing the command's effectiveness. It is essential to teach the staff about the importance of collecting, analysing, visualizing, and communicating data to inform an operations assessment. By briefing leadership early and often, the assessment team can gain leaders' buy-in and they may mandate staff participation in the assessment process. Once achieved, a desire and sense of ownership from the staff will follow. The Operations Research and Systems Analyst (ORSA) should lead and synchronize efforts to update the assessment through weekly or biweekly meetings and should provide expertise in creating assessment products. However, the ORSA should not be the sole creator of the staff's assessment nor the ORSA brief assessment products to leadership independently. In most cases, the staff's integration in any assessment brief will provide rich context an ORSA cannot provide on his or her own. In addition, the staff can make recommendations to the commander that may otherwise be interpreted as uninformed if provided by the ORSA independently.

19.8 ADVICE TO NEW ASSESSORS

Attend as much training as you can prior to a deployment as an analyst. At a minimum, the Deployed Analyst Course offered by the Centre for Army Analysis (CAA) at Fort Belvoir, Virginia, is an excellent means to expand your analytical network and fine tune analytical skills. The CIDNE training in Colorado Springs, Colorado, provides a pre-deployment course to familiarise users with data management and provides subject matter expertise for ORSAs that are preparing to use CIDNE as the database of record in theatre. The NATO Operations Assessment Course in Oberammergau, Germany, is another excellent course to better understand operations assessment theory and the application of operations assessment processes.

Continue to build your network while deployed by reaching out to the others in the analytical community serving in combatant and component commands to gain situational understanding and share ideas. More than likely, you will be the only ORSA on the staff, which may present you some difficulty integrating initially. It is critical to exercise patience; assess the staff and existing processes to watch for opportunities to introduce new analytical processes. Conversely, it is critical to integrate quickly upon arriving to theatre by contributing to staff work that may not directly relate to analysis in order to gain credibility. Spend time talking with members of the staff who routinely collect data; they may possess datasets that can be used to inform the selection of indicators. Attend meeting and become familiar with the battle rhythm and the commander's priorities, decision space, and his biggest concerns. These are all things that will give you insight into what should be covered in your assessment products.

Data collection can be a taxing and difficult effort for the staff. Do not collect data just to collect data. More is not always better, and it can cause the staff to shut down. A standardized database that is mandated for use is critical to collecting, discovering, accessing, sharing, and storing operational data. There may be more information than time available; develop a plan to target analytical resources effectively.

The change associated with the staff embracing the use of assessments to support decision making for the effectiveness of future operations takes time. It can be difficult to promote the value of analysis and create the space in the staff battle rhythm and commander's calendar to make time for the discussion. However, it is worth the effort to have the commander and staff talk through issues that may plague operations and allow for healthy discussions about ideas to improve operations that are driven by data.

Before briefing the commander and staff, share preliminary analysis with teammates and other staff members to ensure analysis is clear, concise, and communicates the assessment. While briefing the commander and other senior staff members, assume the brief will not go as planned. Don't be caught off-guard by discussions or question that are off-topic and try to steer the conversation in a direction that stimulates discussion and promotes understanding of the assessment.

Operations assessments are a combination of art and science and typically a blend of qualitative and quantitative data. Know your leadership and learn how to best communicate information. Consider using open source data and Public Perception Surveys to inform your assessment. Subject matter experts on the staff are typically a great source of information to enhance any operational assessment. A learning organization will value an assessment process that informs leader's decisions and improves operational efficiencies and effectiveness.

19.9 AUTHORS' BIOGRAPHIES

Lieutenant Colonel Vincent Boncich is an active duty Operations Research and System Analyst (ORSA) and a veteran of one operations assessment tour abroad – the Combined Joint Task Force – Horn of Africa (CJTF-HOA). He established an assessment framework for the CJTF-HOA Campaign Plan utilized to evaluate the progress toward lines of effort, objectives, and end states. He spoke about his experiences and lessons learned at both the Army Operations Research Symposium (AORS) and the Military Operations Research Symposium (MORS).

Lieutenant Colonel Natalie Casey is an active duty Operations Research and System Analyst (ORSA) and a veteran of one operations assessment tour abroad – the Combined Joint Task Force – Horn of Africa (CJTF-HOA). She further developed an assessment framework for the CJTF-HOA supporting plan utilized to evaluate the progress toward lines of effort, objectives, and end states. She spoke about her experiences and lessons learned at both the Army Operations Research Symposium (AORS) and the Military Operations Research Symposium (MORS).

19.10 REFERENCES

[1] Capabilities Engineering and Innovation Division Headquarters. 2013. Innovation in Operations Assessment. Norfolk, VA: Capabilities Engineering and Innovation Division Headquarters, Supreme Allied Commander Transformation. pp. 48-56.

Chapter 20 – COMBINED JOINT TASK FORCE – HORN OF AFRICA MAY 2016 – MAY 2017

Mark Leno
US Army War College
UNITED STATES

ABSTRACT

This chapter provides an overview of the Combined Joint Task Force – Horn of Africa (CJTF-HOA) assessment process from May 2016 – May 2017 [1]. The assessments team implemented a process resulting in a quarterly centralised assessment briefing for the Commander using both quantitative and qualitative data. The centralised assessment informed CJTF-HOA's operations process and provided a baseline for other analyses and reporting. The chapter makes five recommendations for deploying analysts:

1) *Don't forget the big picture;*
2) *Read all you can before you deploy;*
3) *Keep assessment products simple and consistent;*
4) *Don't underestimate what you can contribute to the mission; and*
5) *Doing something is better than doing nothing.*

20.1 BACKGROUND

CJTF-HOA was established following the September 11th attacks for the purpose of counterterrorism within the region. The task force began operating from a U.S. Navy ship (USS Mount Whitney) in November 2002 and transitioned operations to Camp Lemonnier in May 2003. Over time, its mission has expanded from counterterrorism to include civil affairs, regional security cooperation, and counter-piracy.

During this deployment, the CJTF-HOA Area Of Operations (AOR) was nearly half the size of the continental United States, and included the countries of Burundi, Djibouti, Eritrea, Ethiopia, Kenya, Rwanda, Seychelles, Somalia, Tanzania, and Uganda. The CJTF-HOA Area Of Interest (AOI) included Central African Republic, Chad, Comoros, the Democratic Republic of the Congo, Egypt, Mauritius, Madagascar, Mozambique, South Sudan, Sudan, and Yemen.

CJTF-HOA's AOR contains many challenges. Somalia is recovering from decades of internal conflict and is confronting violent extremist groups, particularly Al-Shabaab (AS), and piracy. With United Nations' support, the African Union Mission to Somalia (AMISOM) has deployed a force as large as 22,000 troops from neighbouring and other African countries since 2007 in order to counter Al-Shabaab and create conditions for Somalia to develop. South Sudan, the world's youngest state, also remains fragile with significant internal conflict. Ethiopia and Eritrea have long-standing tensions. Djibouti is a strategic location on a major international shipping route that hosts bases from many other countries, including China, France, Italy, and Japan. Many countries in the region are frequently at risk of food shortages or famine, often due to internal conflict between ethnic groups or other factions.

20.2 ORGANIZATION

CJTF-HOA is commanded by a major general or equivalent (OF-7) with headquarters at Camp Lemonnier, Djibouti, and is subordinate to U.S. Africa Command (USAFRICOM) with headquarters in Stuttgart, Germany. Camp Lemonnier is temporary home to approximately 4,000 deployed military personnel, civilians, and contractors.

The CJTF-HOA headquarters is organized as a traditional staff with seven functional directorates generally led by an OF-5 (colonel or equivalent). CJTF-HOA has several supporting units, including civil affairs, explosive ordnance disposal, and mobile construction. The CJTF-HOA Assessment Team falls under the CJ-5 (plans) director and consists of one U.S. Department of Defense Operations Research Systems Analyst (ORSA, OF-3 major, OF-4 lieutenant colonel, or civilian equivalent) as Assessments Chief assisted by one deputy, usually a U.S. Marine Corps or U.S. Navy OF-2 (captain or lieutenant) who generally is not a trained analyst.

20.3 MISSION

CJTF-HOA's primary missions during my deployment April 2016 – May 2017 included support to regional counter-violent extremist operations, security force assistance, and crisis response. CJTF-HOA's official mission as of March 2017 is:

> *Through unified action with U.S. and international partners in East Africa, CJTF-HOA conducts security force assistance, executes military engagement, provides force protection, and provides military support to regional counter-violent extremist organization operations in order to support aligned regional efforts, ensure regional access and freedom of movement, and protect U.S. interests. Be prepared to execute and/or provide support to crisis response and limited contingency operations in order to protect U.S. military, diplomatic and civilian personnel, facilities and interests*.

20.4 ASSESSMENT SCHEME

20.4.1 CJTF-HOA Operations Process Overview

The CJTF-HOA assessments process was linked with the operations and targeting processes led by the CJ-3 (operations directorate). CJTF-HOA's primary mission was based on the AFRICOM Theatre Campaign Plan (TCP) with CJTF-HOA responsible for one Line Of Effort (LOE) which consisted of three Intermediate Military Objectives (IMOs) and a total of seventeen supporting effects. In order to guide individual activities toward these objectives, CJTF-HOA's CJ-37 (targeting section) ran an operations process consisting of a weekly activity planning meeting chaired by the CJ-3 and CJ-5 directors, a monthly Joint Activities Coordination Board (JACB) chaired by the CJ-3 Director, and a Joint Activities Decision Board (JADB) chaired by the Commanding General (CG). All CJTF-HOA units conducting activities (with exception of preapproved or recurring activities) were generally required to go through this mission approval process.

The assessment tied into the operations process primarily in two areas. First, the assessment team's core product, a quarterly assessment briefing, resulted in the CG's guidance on which TCP effects to prioritize, which informed the CJTF's operations process and made operations more effective. Second, the assessment team managed the Command's post-mission data collection process. The CJTF-HOA Chief of Staff published an instruction, developed by the command knowledge management officer and the assessment team, requiring all teams to provide post-mission reports within five days of completion, generally either via CJTF-HOA's Combined Information Data Network Exchange (CIDNE) or the Global Theatre Security Cooperation Management Information System (G-TSCMIS) databases. The assessments deputy briefed reporting

requirements at pre-mission travel briefs, compiled all post-mission reports into an archive using a standardized naming convention, and analysed the reports using a structured table in order to provide a summary for the CG during the quarterly assessment briefings and improve the effectiveness of future missions.

20.4.2 Assessment Products

20.4.2.1 Quarterly Assessment Overview

The assessment team's core product was a quarterly assessment briefing to the CG and staff section directors. During the deployment, we followed a Calendar Year (CY) schedule for assessments briefings (i.e., Q4 CY16 (Oct – Dec 2016) briefing in January 2017, Q1 CY17 (Jan – Mar 2017) in April 2017, etc.). Figure 20-1 below summarizes the quarterly assessment process. We implemented a centralised assessment process for the entire AOR and any other priority countries. It incorporated both qualitative and quantitative data focused on a few simple core metrics most closely linked to CJTF-HOA's operations and priorities. It provided the Commander a summary of trends and issues and provided analysis to make operations more effective. The core metrics included the number of violent extremist incidents by type and location, number of security force assistance activities by type and location, number of key leader engagements by country, security force assistance funding by country, and number of friendly forces and casualties by location.

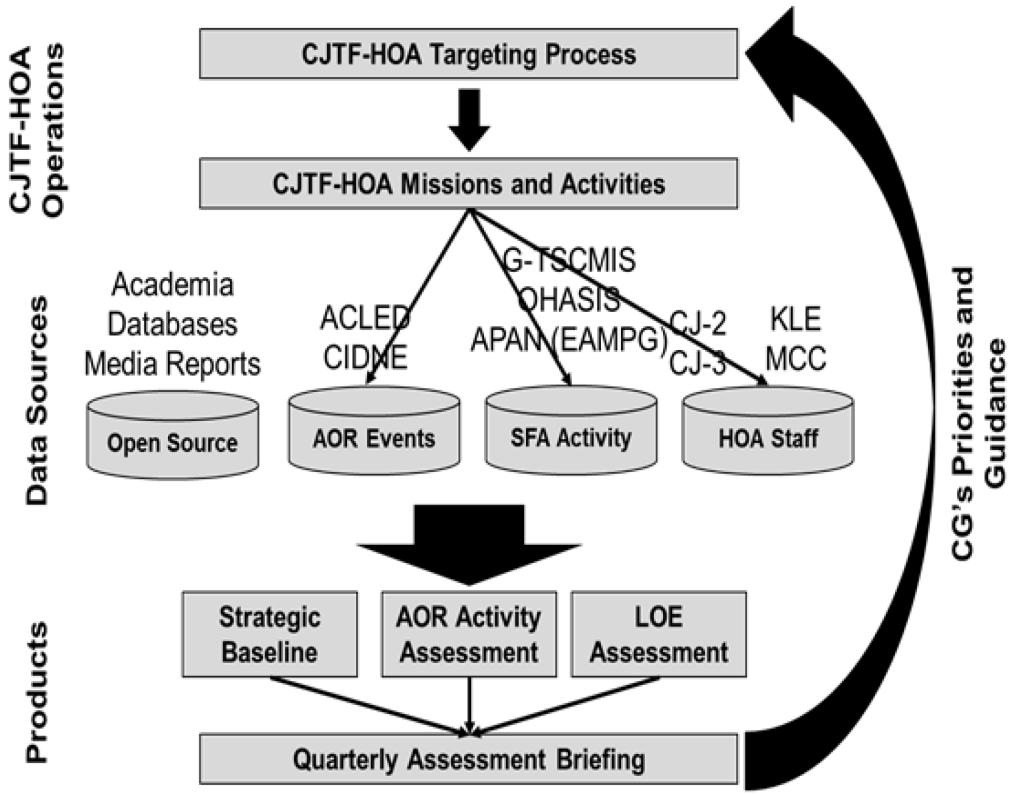


Figure 20-1: CJTF-HOA Assessment Process Overview.

We will provide a detailed rationale for using this assessment approach in this volume's chapter on observations. The centralised assessment process provided the Commander a concise and comprehensive product to interpret the vast quantity of other more frequent and detailed information received and assist the staff with identifying areas for further in-depth planning and analysis. The assessment also measured progress toward CJTF-HOA's assigned objectives and effects and enabled the staff to evaluate and adjust the organization's plan and priorities to make operations more effective.

20.4.2.2 Quarterly Assessment Data Sources

Post-mission reports, an output of the previously described operations process, provided one of the primary data sources for the assessment process. In addition to the mission data, the assessment process drew on data from four other categories of classified and unclassified sources to provide a comprehensive product.

Open sources included news, academic, and media reports to supplement and/or validate other sources as needed. CJTF-HOA contracted for several supporting products, including media analysis and in-country observer reports.

For AOR events, the primary sources were the Armed Conflict Location and Event Data Project (ACLED), an academic database of violent incidents throughout Africa updated weekly by researchers at the University of Sussex, United Kingdom, and CJTF-HOA's CIDNE server, which contains data on violent events within Somalia from AMISOM and other intelligence reporting updated at least monthly.

For Security Force Assistance (SFA) activities, the primary sources of data were the Global Theatre Security Cooperation Management Information System (G-TSCMIS), Overseas Humanitarian Assistance Shared Information System (OHASIS), and the East Africa Multilateral Planning Group (EAMPG) databases. G-TSCMIS is the Office of the Secretary of Defense's system of record for security cooperation activities and includes data for CJTF-HOA and other DoD organizations' activities throughout the AOR, including military-to-military engagements and training. OHASIS is a DoD repository for Humanitarian Assistance (HA) activities, such as building schools and medical assistance. The EAMPG portal on the unclassified All Partners Action Network (APAN) website contains data voluntarily posted by major international donors including European Union, France, Turkey, and United Kingdom on training activities for African partners in the region, primarily Somalia.

Finally, CJTF-HOA staff sections also provided data for the assessment process, particularly the CJ-2 (intelligence) and the Military Coordination Cell (MCC) in Mogadishu. The CJ-39 (information operations section) and Military Information Support Operations (MISO) detachment also provided valuable data, especially surveys on Somali perceptions of key issues.

Data quantity and quality were major challenges. The small team size (two personnel) limited the amount of data we could collect and process. We depended on external sources for much of our assessment data, which often made quality control and timeliness challenging. Two other challenges were how accurately sources measured activity in the operational environment and measuring change in the environment due to variation in collection and reporting timelines. To mitigate these challenges, we focused on a few metrics most closely related to our organization's core missions and obtained two or more sources for each metric to reduce risk of overreliance on one source. We also examined how our sources collected data in order to understand and mitigate distortion, gaps, or other issues. We made maximum use of data the staff was already collecting and minimised any additional collection requirements. We formed close working relationships with supporting units which allowed us to both improve their operations and obtain better data and understanding for the assessment products. Finally, we made extensive use of historical data to provide context to evaluate recent data. We developed baselines using all available historical records for our AOR and AOI from our primary databases.

20.4.2.3 Quarterly Assessment Structure

As depicted in Figure 20-1, the quarterly assessment product consisted of up to three sections. Our first quarterly assessment briefing (Q2 CY16) included a "strategic baseline" covering strategic-level economic, political, and military trends over a 20-year timeframe to provide the Commander and staff a foundational analysis of the AOR early in the Commander's tour.

The second section, "AOR Activity Assessment", consisted of an AOR violent activity assessment and a Security Force Assistance Activity (SFA) assessment. The purpose of this section was to provide an empirical foundation for the line of effort assessment and to provide the Commander and staff a summary of key activities and core metrics in order to focus activities or follow-on analysis so that operations become more effective.

The AOR violent activity assessment started with an animated map to give the leaders a short visualization of violent events by type and location over the current and previous quarters (or past 10 years for the initial briefing) followed by a static map with counts of violent activities by category overlaid with a mini-graph showing the 3-year trend. The following four slides provided a deep-dive into violent activity in Somalia, including activity by Violent Extremist Organizations (VEOs); incident count by quarter; casualties and event types for past two quarters overlaid with mini-graphs for 3-year trends; and a friendly-to-enemy force ratio analysis, including a map overlaid with attack and casualty statistics. The regional assessment concluded with any supporting analysis in response to CG or staff interests, such as specific operations or Centre Army Analysis (CAA) projects on VEO and IED activity trends. The purpose of the AOR assessment was to provide the Commander and staff a consolidated analysis of recent and long-term activities to provide context for more frequent, detailed, and effective operations and intelligence reporting.

The SFA assessment examined friendly forces' activities and resource allocation within the operations area. This provided the staff information on quantity and types of activities by country. For the initial briefing, we analysed all the available applicable historical data for our AOR from G-TSCMIS, OHASIS, and EAMPG to set a baseline for future assessment teams and show the Commander long-term trends to provide context for current activities. For brevity, we blended together the OHASIS (Humanitarian Assistance) and G-TSCMIS (security cooperation) data to provide a single picture of military aid and assistance by country. We developed a process to organize and analyse post-mission reports resulting in a one-slide summary for the Commander on the most commonly-cited issues and recommendations. Based on lessons learned from the first quarterly assessment cycle, we implemented a process to archive large numbers of primarily qualitative post-mission reports and incorporate them into the assessment. The assessments deputy collected all mission-related reports into a folder on the CJTF-HOA shared drive, extracted key information into structured tables in an Excel workbook to facilitate summary analysis, and generated standardized descriptive file names for the archived reports. Since we did not have sufficient resources to analyse all the reports in greater detail, this at least, allowed us to look for high-level trends and ensure reports were easily available for any future analysis or deep dives on a specific topic. To conclude the SFA assessment, we blended several data sources to create a "resource" summary slide showing the allocation of the Command's three primary "resources" by country: SFA activities, SFA program funding, and the Commander's key leader engagements.[1]

The third and final section is the assessment of objectives and effects from CJTF-HOA's Line Of Effort (LOE) in the AFRICOM TCP.[2] CJTF-HOA had historically used an effects-based assessment methodology which our team continued. We refined a metrics spreadsheet and guidelines developed by the previous assessment team. In coordination with the staff, we developed a spreadsheet outlining metrics and indicators for each effect in order to track progress and record supporting data. For each indicator, we established a quantitative rating criterion where possible. If it were inappropriate to quantify a metric, we made a qualitative risk rating (high, medium, or low), and documented the supporting facts and reasoning for the rating. We tried to make this rating structured and transparent, which enabled us to explain and justify the ratings based on empirical evidence or logical thinking, and provides a transparent basis for future analysts to assess change and continue to refine the assessment. Using a formal construct of indicators also helped us and the staff to clarify data requirements and gaps for the Command and evaluate the campaign plan's actual objectives and effects. For summary purposes, we used a Green, Amber, Red (GAR) rating scale for

[1] We did not include funding in the resource slide for subsequent quarters if there was no major change to the data.

[2] Our first quarterly assessment (Q2 CY16) product was different in detail since it was based on the HOA Supporting Plan (HSP) to the AFRICOM TCP. In the summer of 2016, the Command decided to operate off exclusively off the TCP.

communicating the status of objectives and effects. For consistency, we used simple rules shown in Figure 20-2 to rate each effect and objective based on the ratings of its component elements.

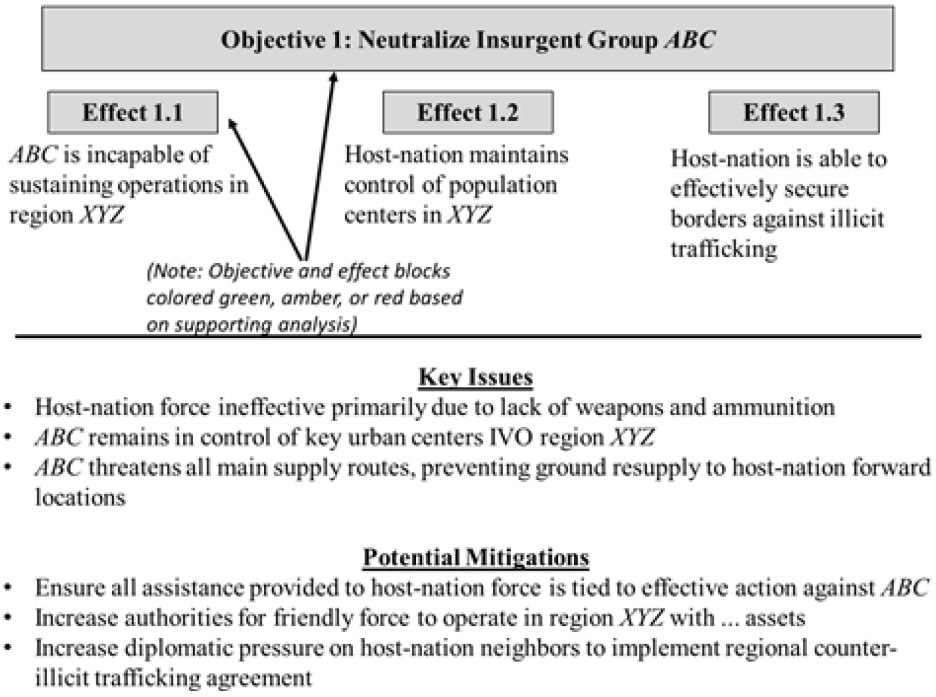


Figure 20-2: Notional Objective Assessment Summary Slide.

The GAR rating scale is an imperfect construct and rightly criticised for several reasons, such as the tendency to converge to "amber" ratings and not showing trends (such as whether an amber item is improving or degrading).[3] However, it has the advantages of being simple and easy to understand and enabling the Commander to quickly review the status of large numbers of items, such as the seventeen effects we were required to assess. Our purpose for using the GAR scale was only to communicate a summary of the relative status of effects and objectives based on the underlying analysis. The GAR scale is not a mathematical model or substitute for analysis. For assessment briefing purposes, we developed a one-slide summary for communicating the rating for each effect that was based upon the data-driven method described above. We also added a narrative of key issues and potential mitigations to provide qualitative context for the ratings and clear takeaways for the decision maker (refer to Figure 20-3 for a notional example). Based on the findings of the assessment process, we concluded the assessment product with recommendations for the CG's decision on the command's priority effects for the next quarter, which in turn informed the operations process so that operations became more effective. The assessment product also served as the source for CJTF-HOA's input to AFRICOM assessment products and other external requests for information.

20.4.1.4 Other Products

In addition to the quarterly assessment, we supported the Commander, senior staff, and other staff sections with specialized analyses upon request. This proved to be very beneficial for both the organization and the assessment process. We were able to directly support operations and decisions, build rapport and support for the assessment process with the wider staff, and obtain better data and understanding of the operating environment for the centralised assessment process. These efforts resulted in more effective staff operations.

[3] For example, see Ref. [2] pp. 75-76 and Ref. [3] pp. 111-115.

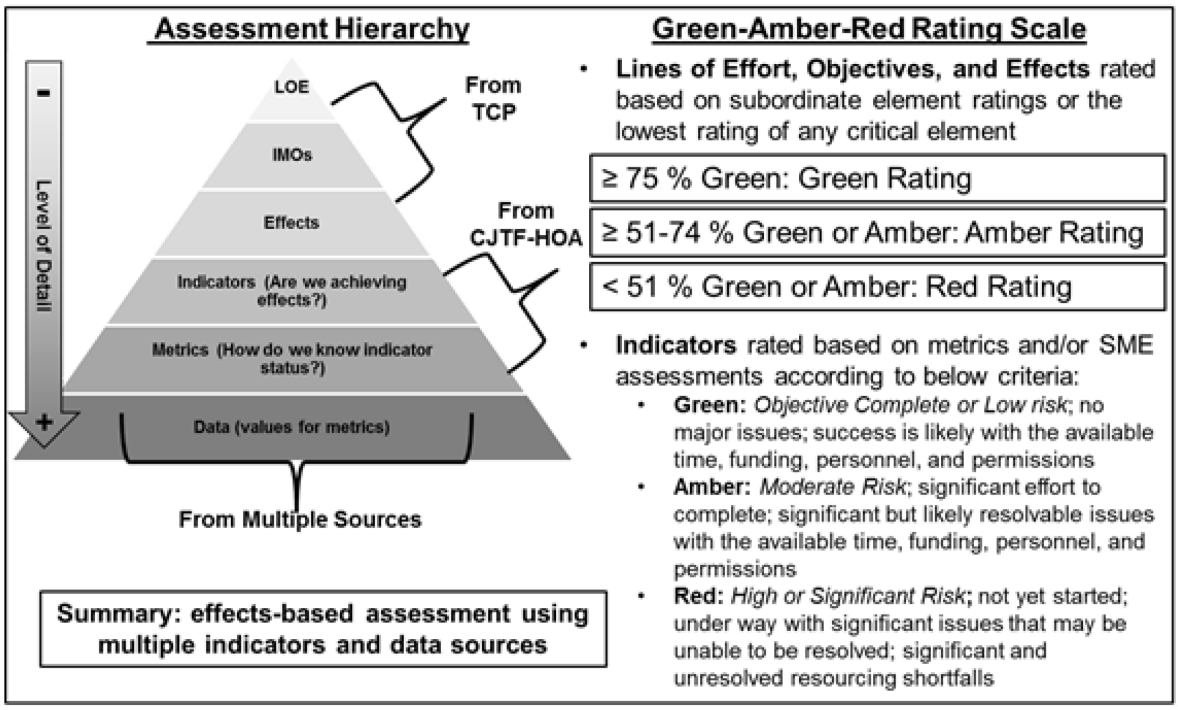


Figure 20-3: Effects-Based Assessment Hierarchy and Rating Scale.[4]

For the command group, we completed several quantitative analyses on violent activity trends in specific areas of interest as well as a project on the volume of staff actions (taskers) and their distribution by staff section. For the key leader engagements team, we created three custom Microsoft Excel tools which worked off their existing spreadsheets, enabling them to track and analyse visit requirements and action officer assignments without having to manually read through various sheets. We also completed numerous analyses in support of targeting and other activities for the operations section and supporting units. Examples of projects included analysing violent activity trends for specific locations, custom calculators for targeting, assisting with survey design, and analysis of large datasets.

20.5 IMPACT ON COMMAND

Our intent was to make the CJTF-HOA assessment relevant and timely in order to substantially contribute to operational effectiveness. First, the quarterly assessment products provided the Commander and staff a summary of an empirical and historical baseline from which to interpret the vast quantity of contextual and other detailed reporting and information the Command received. Second, the Commander used the assessment to evaluate overall progress toward CJTF-HOA's objectives and provide his guidance on priority effects which in turn made operations more effective. Third, the assessment team directly contributed to operational effectiveness through on-demand supporting analysis throughout the command.

20.6 LESSONS LEARNED

There are several areas that I would try to improve on a future deployment. I did not make significant progress in these areas due to insufficient time and personnel, organizational constraints, or decisions to concentrate on other priorities.

[4] This builds on a construct and definitions provided by the previous CJTF-HOA assessments team led by LTC Shawnette Rochelle, Center for Army Analysis.

20.6.1 Travel Within the Operations Area

I recommend that the assessment team visit key locations in the organization's area of responsibility approximately two to four weeks after the start of their deployment to allow them sufficient time to understand the basics of the environment and prepare significant questions. Most of my trips within the AOR were midway or later in the deployment, and information and contacts obtained may have even more useful earlier. As feasible and **useful** for the assessment, the team should occasionally go out again to spend a few days in key operational areas, such as in advance of a quarterly assessment briefing.

However, travel is a double-edged sword. Any assessment team travel must not detract from substantial work or analysis nor should it become "military tourism", too focused on tactical details, or make significant extra or redundant work for outstations. Simply travelling around and talking to people it is not likely to be an effective use of the assessment team's or the host's scarce time; engagements should be mutually beneficial and to satisfy legitimate and specific information requirements that cannot be satisfied by other means. The assessment team's core product should be focused on core metrics and issues as they pertain to overall progress of the operation; the operations, intelligence, and other sections should rightly handle the tactical and contextual details. The assessment team should be able to obtain the majority of its data from what the staff already collects, existing databases and reports, and calls to personnel with specialized knowledge. Travel should only be used sparingly, for major events such as worthwhile external exercises or occasional engagements with key personnel in outlying areas. It is unlikely that members of the assessment team will be experts on contextual details of any aspect of the operation[5]; the goal should be to have a reliable understanding of the major activities and challenges and to focus on the key metrics and issues for the Commander.

20.6.2 Working with External Agencies

Within the operational environment, there are likely to be other analyst teams, including assessment teams for other organizations. Within East Africa, we had assessment teams from various other military organizations, including several assessment teams from our higher headquarters. Other international, governmental, and non-governmental organizations will also likely have teams of analysts. During my deployment, I worked with several higher headquarters teams and made contact with several other U.S. military assessment teams, but did not work closely with other governmental or non-governmental organizations. Of course, an analyst must above all take care of his organization's needs and meet higher headquarter requirements, but collaborating or exchanging data with other organizations may prove beneficial depending on the mission and environment. A deployed military analyst could consider reaching out to diplomatic, intelligence, or other government agencies, and non-governmental organizations to share data and discuss specific issues areas. However, he will likely need to address classification requirements, network restrictions, perceptions, and realities of too many meetings, and other potential issues. Two options I considered but did not execute were a monthly regional interagency analyst secure video teleconference and periodic assessments discussions with regional diplomatic and intelligence analysts.

20.6.3 Program or Activity-Level Assessments

One area where the operations research community might be able to significantly improve the effectiveness of operations is with assessment of cost and benefits of aid, such as security force or humanitarian assistance. While it will likely be difficult to measure effectiveness due to long (potentially multi-year) resourcing process timelines and delayed outcomes, some formal assessment will at least encourage resourcing and program staff to more carefully consider cost and operational effectiveness of these activities. From my perspective, the focus of most SFA and other aid programs was on obtaining and providing resources rather than their effectiveness or opportunity costs and alternatives.

[5] Other than integrating the outputs from databases/sources available to the command as needed for the assessment.

My recommendation is to have the assessment team sized to provide at least one analyst full-time to assess cost and benefits of assistance and related programs with respect to the organization's objectives. The analyst's focus should be assessing how effectively aid and assistance investments actually support the organization's strategic objectives and operational needs as well as provide benefit to aid recipients. This analysis should result in recommendations to use resources better, to include cutting spending that is not productive or that may even be counter-productive. The analyst should also assist teams executing SFA and related activities with developing their own plans to assess program effectiveness at a more tactical level. At the very least, the assessment team can advocate that the staff justify how requested programs will align with the campaign plan and demonstrate a reasonable plan to assess their execution. The requirement for an assessment may at least encourage program teams to think about the return on investment and direct their activities accordingly.

20.7 ADVICE TO NEW ASSESSORS

20.7.1 Summary

Based on my deployment experience, I offer five recommendations:

1) Don't forget the big picture.
2) Read all you can before you deploy.
3) Keep assessment products simple and consistent.
4) Don't underestimate what you can contribute to the mission.
5) Doing something is better than doing nothing.

Don't Forget the Big Picture. Don't forget your organization's strategic purpose or reasons for existence. As much of your time will be spent working with data, make sure you don't lose sight of the plans and intentions of your Commander and other key actors, and emerging issues in the area of operations. I found being attached to the plans directorate to be advantageous, since I was in the same room as CJTF-HOA's lead planners throughout the deployment. Attend key meetings such as operations updates to the Commander and intelligence briefings to maintain situational awareness, but don't disrupt or get in the way of other teams' work. Maintain contact with key staff throughout the organization for understanding of emerging issues and plans, but don't get bogged down in the details. To succeed in assessment, you will need to get and maintain a solid understanding of your organization's objectives, activities, and data. You will need to maintain awareness of key developments in at least operations, intelligence, and plans. For best results, I recommend that the assessment team members deploy for at least six months (ideally on the same timeline as the Commander) and be fully integrated within the staff. Building relationships and a deep understanding of the missions, environment, and Commander's vision also likely takes at least several weeks. It may take a few months to get an assessment process institutionalised across a staff if one does not already exist.

Read All You Can Before You Deploy. Read all you can about your area of operations and the practice of assessment before you deploy. At the very least, get a few recommendations from knowledgeable sources on the best books or articles. This will enable you to more quickly become effective when you arrive. If you are fully engaged in leading the assessment process and providing supporting analysis, you won't have much time to read. Understanding the culture, history, politics, and other issues of your area of operations is important for leading an effective assessment process.

Don't Underestimate What You Can Contribute to the Mission. The staff will likely view the primary contribution or competitive advantage of deployed analysts, particularly operations researchers, as showing the data or as one director put it "telling stories with numbers". Even simple quantitative or other formal

empirical analysis can provide a valuable, different perspective to leadership that they may not be getting from other staff sections, particularly if most of the other analysis in the headquarters is qualitative and/or current-events focused. Quantitative analysis, if done clearly and transparently (with frank discussion of potential limitations), can efficiently summarize large amounts of data, particularly trends over multiple years. This can assist the Commander and staff with validating operational updates and qualitative reporting, and identifying areas for more in-depth review. Rather than taking the easier path of simply writing or aggregating narratives, as a deployed analyst, you should use your specialized training to provide the most rigorous analysis possible. In particular, you should maximise the value of any available quantitative or other empirical data in order to provide the Commander a more complete perspective on the environment for his or her decisions. Work hard to build rapport with other staff sections and units, and find ways you can use your skills to help them, whether with finished products or by improving the effectiveness of their analysis. Building relationships with other teams and units can also build support for the assessment process and provide valuable data and information for your core products.

I encourage all deploying analysts to write, archive, or otherwise share their lessons learned, particularly details of their processes, products, and data sources, for the benefit of future analysts. While there is considerable literature on criticism, theory, and general guidelines for assessments, there are fewer detailed examples and explanations of real world assessments programs written by the analysts who ran them. This chapter is intended to contribute to this gap, and I hope that future analysts will come back with better reports and improved ways to assess operations in order to make them more effective.

Keep Assessment Products Simple and Consistent. Keep your products simple, but effective. An operations researcher's comparative advantage will be the ability to perform quantitative analysis to a higher quality standard than most of the staff. Do not sacrifice your analytical integrity or quality, but learn to only show the Commander what he or she needs for his or her decisions. Leave the rest in backup slides or your archives to reference as needed. Conversely, do not fall into the trap of not doing robust supporting analysis or sacrificing the quality or the integrity of your analysis simply because others may see only the highlights. Highlights will only be ultimately as good as the quality of the analysis they are drawn from. You need to learn to go to deep enough that you can make sure the analysis is sound, but not into so much detail it detracts from other work.

Show or discuss complex mathematics and analytical techniques only as needed. Don't use methods and techniques more complex than necessary, and be able to explain simply and clearly whatever you do use. Possibly only a few people in your deployed organization will really understand advanced techniques and even fewer will have time for them. Ironically, the people who do understand them will probably be some of the busiest on the staff and be the least likely to have the time to appreciate them.

Identify a few core metrics that best fit your organization's operations or problem set and analyse them well. A few metrics strongly linked to the theory of your organization's plan, analysed carefully, are likely more useful than a web of measures of performance and effectiveness or many pages of narrative detail. With any real world problem, there could be a very large number of factors and contextual details to consider, but it is likely that decision makers will use just a few metrics they consider most relevant to the decision, such as cost, risk, casualties, or time. Isolate a few metrics that are closely linked with the strategic or operational objectives (e.g., terrorist activity in a counter-terrorist campaign, SFA activity in an SFA campaign) and analyse them well to provide the Commander a baseline for decisions that make operations more effective.

Focusing on a few metrics facilitates transparency, enables measurement of change for future assessment teams, and enables further aggregated or disaggregated analysis as required. For example, the Commander may require an assessment product at regional level while the operations section needs a product focused on a single city. Maintaining the best possible data on a few metrics closely related to your organization's mission will enable you assess and compare assessments of operations at various levels. If you have significant time or other constraints, focusing on a few metrics may also enable you to improve quality by

using multiple data sources for each metric. Using multiple sources that measure the same or similar things might be one of the best ways to mitigate real world data quality or measurement variation issues. For example, you probably will not have a single ideal data source for civilian casualties, but if you can find two or three independent databases that measure casualties over the same locations and time periods, you can examine whether or not the different sources have similar trends.

Doing Something Is Better Than Doing Nothing; If You Don't Act, Others Will. In my experience, operational headquarters and units, for better or worse, thrive on action. The environment changes rapidly (sometimes in a matter of hours) and decisions are often made quickly. The Commander and staff will always be looking for the best information they can get to support their decisions with the limited time they have. If you don't act, others who may be less analytically rigorous and/or less scrupulous will fill the void. To be effective, you will need to provide useful products in a timely manner.

You will need to focus your efforts on where you can contribute most effectively, providing quality analysis where the other staff sections are unable to do so. Learn to do the best you can with what data you have or can get in a reasonable time. When you make products, it is better to have a complete, coherent, and short answer to the main question(s), rather than a very detailed analysis on a subset of the problem. Start your analysis at a high-level and time-permitting, go into more detail based on the Commander's requirements. Keep methodology and data requirements simple. Better to do simple analysis clearly and well than to get too fixated on theory or complexity and produce little or be too slow. Data quality and quantity will almost certainly be a challenge. Near-perfect data or lengthy and complex academic-style studies are not realistic in a deployed environment nor is the Commander likely to expect them. Be transparent about critical assumptions and potential limitations, but above all, provide the most rigorous analysis and insights you can, with what data you can get, and the time you have available.

20.8 CONCLUSION

Operations assessment in a deployed organization is probably one of the most challenging and significant jobs for an analyst. Your work can directly impact the effectiveness of military operations in ways that are not possible in other analytical roles. Based on experiences at CJTF-HOA, my primary recommendations are to focus on a centralised assessment process using both quantitative and qualitative methods and, as much as feasible, support the staff with specialized analysis and help them improve their own products and processes. Most importantly, regardless of the approach you take, make sure your assessment and other analytical work are based on rigorous analysis, transparent methods, and the best available empirical data.

20.9 AUTHOR'S BIOGRAPHY

Mark Leno is a U.S. Army civilian analyst at the U.S. Army War College. He deployed as Assessments Chief, Combined Joint Task Force – Horn of Africa from May 2016 – May 2017. He is a graduate of the University of Chicago, Gettysburg College, and the U.S. Army Operations Research Systems Analyst Military Applications (ORSA-MAC) course and is a former Presidential Management Fellow.

20.10 REFERENCES

[1] Combined Joint Task Force – Horn of Africa. 2016. From Crisis Action to Campaigning: Lessons and Best Practices. Centre for Army Lessons Learned (CALL).

[2] Connable, B. 2012. Embracing the Fog of War: Assessment and Metrics in Counterinsurgency. Santa Monica: RAND.

[3] Downes-Martin, S. 2011. Operations Assessment in Afghanistan is Broken. Naval War College Review, 64(4):103-125.

Chapter 21 – THE DUTCH APPROACH IN OPSA – LESSONS LEARNED FROM A DECADE OF OPSA IN AFGHANISTAN

A.R. (Aletta) Eikelboom
Netherlands Organisation for Applied Scientific Research (TNO)
NETHERLANDS

R.G.W. (Rudi) Gouweleeuw
Netherlands Organisation for Applied Scientific Research (TNO)
NETHERLANDS

Geert Roseboom
Netherlands Organisation for Applied Scientific Research (TNO)
NETHERLANDS

Jeffrey Schwerzel
Royal Netherlands Army
NETHERLANDS

ABSTRACT

Parallel with Operations Assessment (OPSA) developments within NATO, Dutch OPSA officers have supported the ISAF mission in Afghanistan from 2003 to 2013 in different regions and in different levels of headquarters. From 2003 and onwards the Dutch military OPSA capability and their OPSA methodology has continuously evolved. This chapter describes observations and lessons learned concerning OPSA during the Dutch OPSA involvement in Afghanistan, written by a team of OPSA officers who have contributed personally and have served on multiple deployments.

21.1 INTRODUCTION

Over a decade in the International Security Assistance Force (ISAF) in Afghanistan the Dutch have gained a solid experience in Operations Assessment (OPSA). The foundation for the Dutch OPSA experience was laid in 2003 when two researchers from The Netherlands Organisation for Applied Scientific Research (TNO) were sent to HQ ISAF in Kabul to experiment with OPSA as decision-making support. This practical experiment was deemed so successful, that the Royal Netherlands Army decided to commission a number of scientists as reserve officers to facilitate OPSA at corps and brigade level. These officers were instrumental in designing and implementing OPSA in three Netherlands Armed Forces' contributions to the ISAF mission.

After the pilot programme in Kabul, most of the experience was gained in the team effort of supporting the mission in Uruzgan Province with OPSA for the period when the Netherlands was the lead nation there between 2006 and 2010. In that period, OPSA had proved so valuable that it was deemed a key asset to develop a design for the upcoming mission in Kunduz from 2011 to 2013, see Figure 21-1.

Over the years, as insight into missions and experience with OPSA accumulated, the Dutch OPSA methodology also evolved. This chapter describes the evolution of the Dutch OPSA methodology and the conclusions to be drawn from this experience.

21.2 TASK FORCE URUZGAN (2006 – 2008), EXPERIMENTING TOWARDS AN OPSA RECIPE

Having contributed with Provincial Reconstruction Teams (PRT) in the north of Afghanistan and in the HQ ISAF in Kabul, the Dutch government decided in 2005 to increase its contribution to the ISAF mission by providing a Task Force responsible for leading the ISAF effort in the province of Uruzgan, part of Regional Command South (RCS). Between July 2006 and July 2010, the Task Force Uruzgan (TFU) mission was to

assist the Afghan National Security Forces (ANSF) to set conditions for a secure and stable province, to assist the provincial government in building its capacity, and synchronise overall reconstruction and development efforts.

Figure 21-1: Dutch OPSA in ISAF.

21.2.1 Organisation of TFU

The strength of TFU varied around a mean of 2000 persons. The organisation consisted of a regular brigade Headquarters (HQ); a mechanised Battlegroup (BG) with ground based fire support; a Provincial Reconstruction Team (PRT, a battalion staff and several mission teams); an engineer company; an Intelligence, Surveillance, Target Acquisition, And Reconnaissance (ISTAR) company; a Psychological Operations (PsyOps) element; an Operational Mentoring and Liaison Team (OMLT); an Australian Mentoring and Reconstruction Task Force (MRTF), consisting of a battle group headquarters, a combat team, a combat engineer squadron, a combat service support company; a Special Forces Task Group; and a logistical detachment (supply, maintenance and a Role 2 medical facility). Later, a Dutch AH-64 Apache detachment was stationed at the main base in Tarin Kowt, and the TFU was reinforced with a Slovak guard unit and a Singaporean radar unit. The TFU was located on two main bases and had established several small patrol bases.

In the first two years of its mission, TFU was commanded by a colonel, and in the second two years by a one-star general. TFU staffs and commanders rotated after every six months, which means there have been eight TFU staff rotations in total. TFU was also supported by interdepartmental advisors, like a Political Advisor (POLAD), a Development Advisor (DEVAD) and a Cultural Advisor (CULAD). To better adapt to the complex interests of the mission, in the second term, the decision-making power of TFU was shared by the one-star general and a Civil Representative (CIVREP) from the Ministry of Foreign Affairs. Most of the civilian advisors also rotated in the rhythm of six months.

21.2.2 TFU Mission

The ISAF mission was complex: progress had to be made using all instruments of power across all domains of the engagement space (Political, Military, Economic, Social, Infrastructure and Information). It would take considerable time to accomplish substantial change as multiple parties were involved, and a patient opponent who was difficult to locate and/or distinguish from the population resisted ISAF's efforts.

Given that complexity and the fact that HQ ISAF also used an Effects-Based Approach to Operations (EBAO), the (NLD/AUS) Task Force Uruzgan decided to adopt an Effects-Based Approach to execute its mission. At the time, a considerable amount of writing on EBAO was available. However, internationally, there was little documented experience of an effective implementation of EBAO to steer such a mission on a country or provincial level. Within ISAF there was not one consistent and transparent method to be recognised. So, while the concept of EBAO was not totally new, it was just not incorporated into the Dutch or other participating nations' decision-making process or mission architectures at that time. As the Netherlands Armed Forces had limited experience with the implementation of EBAO in operations, this was the first opportunity where EBAO and OPSA were used to address the question: "how to implement an effective and pragmatic method to steer the Uruzgan mission?"

21.2.3 OPSA Involvement

After TFU formation in 2006, in the months prior to actual deployment, the TFU staff saw itself confronted with the challenge to design a framework to monitor and control the mission. The staff, mostly Army officers, had to deal with different goals across the engagement space and had a rich toolbox with different enablers, ranging from combat and combat support troops to CIMIC and reconstruction capabilities that required close cooperation with personnel and advisors of the Ministry of Foreign Affairs and Development to be most effective.

Moreover, as every region in the area of operations had its own dynamics, a top-down approach clearly had its limitations because these local dynamics were insufficiently known at the higher levels. Therefore bottom-up initiatives had to be incorporated to take this knowledge into account. Apart from defining proper effects to achieve, the staff struggled with the question of how to define and measure progress. Where physical effects can be observed relatively easily, mental effects, like the perceptions or attitudes of actors, are far more difficult to capture. To address all challenges in one suitable framework and build a useful assessment methodology, the staff asked the support of two OPSA officers. Their contribution was so valued that it was decided that the OPSA officers would not only assist in preparing the mission, they would also deploy with the TFU.

21.2.4 OPSA Methodology

The first activity of the OPSA officers was to capture concepts and accompanying definitions of the OPSA methodology, in order to lay a recognisable foundation for the methodology and to counter a certain level of scepticism and uneasiness within the rest of the staff for "academic processes". This was important as OPSA is not just the work of a couple of OPSA officers or just another parallel work-generating process. OPSA thrives by the acceptance and cooperation of the staff and its sub-commanders. To realise this, the aim was to align and integrate OPSA as much as possible within existing procedures of the TFU (Figure 21-2).

To complicate matters, TFU had to incorporate different sets of guidance and Lines Of Operations (LOO). As the HQ ISAF set of effects was already determined, it would seem logical to select those effects applicable to the TFU area of operations. However, RCS, a subcommand of HQ ISAF and the higher echelon command of TFU, had defined its own set of effects, which were different from the HQ ISAF set. Furthermore, TFU had to adhere to the objectives stated in the Dutch national operational guidance, which, although not very specific, incorporated some unique aspects and limitations. Because TFU had to report in

both the ISAF and the Dutch national line, TFU decided to integrate all identified effects and carefully track the headquarters that originated each one.

Key to the first TFU implementation of EBAO was that a breakdown of measurable "effects to achieve" was created and that relations between those effects were identified. This enabled TFU to show how progress on one effect could influence progress or disruption on other effects. It also enabled a better coordination of efforts of several actors and non-kinetic influences which led to more effective operations.

When creating a set of effects, the trade-off is to define a tractable number of effects: sufficient to capture all aspects of the mission and to enable effective control versus the risk of choosing too many and losing the "big picture". This way, TFU ended up with a set of 23 effects.

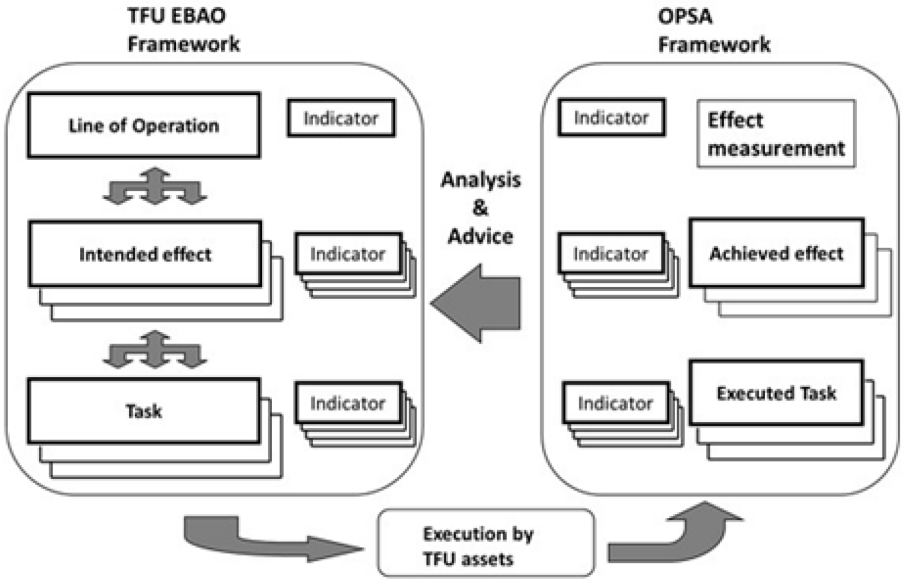


Figure 21-2: Relationship of the Assessment Scheme to the Operational Scheme.

Subsequently, in acknowledgement that reality was far more complex than what could be captured in a structured operational design with a few neat and distinctive LOOs, the OPSA officers created an influence diagram, depicting relationships between effects in one schematic. This diagram marked the effects for which TFU was the primary actor in achieving them, as well as effects that TFU could only contribute to, indicating where other actors would be in the lead. The schematic was soon called "the spaghetti model" (see Figure 21-3), symbolising the interrelationships and awareness that one activity could influence the whole environment.

Essential in TFU's first approach to operations was OPSA, the activity that measures progress and results, in order to develop conclusions and recommendations to support decision making within the Task Force. In order to measure the effects, indicators that characterised the effects had to be defined. Although operation researchers often prefer to work with quantitative indicators, the OPSA officers soon discovered that many indicators were difficult to express in specific numbers. For example, it is easy to measure the number of soldiers at a location or the number of police stations, but the effectiveness of the police, the level of support of the local population, or the perception of security in a certain area are less tangible. This meant that data collection had to settle with mostly qualitative indicators, often based on expert opinion. Provided the substantiation of the judgements was well documented, subsequent measurements could be compared. The added value of independent polling in the TFU Area Of Responsibility (AOR) proved too limited for operational insights as it concerned a small part of the operational area and a small sample of respondents.

Associated with the effects were "levels of ambition", which were basically intermediate goals. For each effect, an "ambition" was estimated for three moments in time: after six months (denoting the rotation of the first TFU staff), after two years (denoting the end of the initial Netherlands commitment to lead the TFU) and at ISAF mission accomplishment (which did not have a specific duration, because it was condition-based). The "levels of ambition" for the short-term determined priorities for the individual effects and were driven by considerations of the current impact on the mission, available TFU capabilities, and expected benefits of other parties' efforts (although in the early phase almost no International Organisations (IO) or Non-Government Organisations (NGOs) were present in the area of operation). By evaluating effects against resources and ambitions, the OPSA process greatly influenced and steered the planning process and decision making of the TFU. In this method, two challenges were encountered. First, in order to define a realistic ambition for a certain timeframe, it is necessary to have a clear understanding of the initial situation. In the preparation phase in the Netherlands, it proved impossible to capture such a detailed baseline for all effects. A best effort was made with the available data, and the first actual assessment, which served as an accurate and satisfactory baseline was established after the first two months in theatre. Second, a typical military habit is to define specific, short-term objectives, which resulted in the definition of six month ambition levels for each effect. These short timeframes caused a minor cultural clash with the fundamental beliefs and experiences of the development partners. Where the military likes to see rapid results reflected in unambiguous indicators, the development community is more narrative and process-driven, using longer-term and more broadly formulated effects. Ultimately, experience showed that progress on many of the stabilisation-related effects, e.g., local support for the provincial authorities, can only be measured reliably over a longer period than a couple of months.

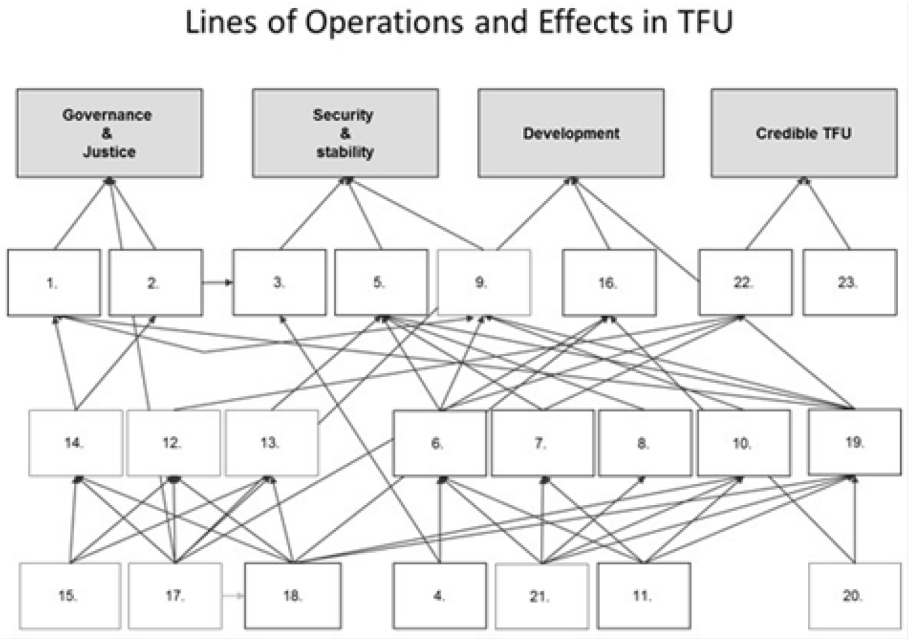


Figure 21-3: Nesting of Effects under Lines of Operations in TFU.

21.2.5 Data Collection

The Measures Of Effectiveness (MOE) were initially incorporated in the Intelligence Collection Plan. Due to the limited capacity of the intelligence gathering elements and their initial focus on terrain and enemy, not much ready-to-use data were generated. To compensate, the OPSA officers had to do much of the data gathering themselves. They read the intelligence summaries and reports of all relevant sections on a daily basis, they attended all briefings of the major sub-units and relevant working groups, and once a week, they

discussed the overall situation during an hour with the commanders of the TFU, battle group, and PRT. When required, they also interacted with the special advisors (POLAD, DEVAD and CULAD), and commanders of OMLT, PsyOps and all other relevant units. Through these interactions, a comprehensive common operational picture on the status of the effects was built and maintained. Measures Of Performance (MOP) were collected similarly, although the people in the TFU Operations Room were soon capable of reporting a large part of the required data.

Part of the data was requested by and provided to the higher headquarters, RCS. After some time, RCS wanted to know why the TFU was less active than other task forces. This annoyed the TFU, because they considered themselves very active. Closer examination showed the cause of this misperception. An operation for the TFU was a coordinated set of activities contributing to one specific goal, so they regarded themselves as executing one or two operations a month. The daily patrols TFU units were executing were considered as normal framework operations and were not reported individually as operations. Meanwhile, other units were labelling each patrol with a unique operation name and thus reported scores of operations a month. An example like this demonstrates that ISAF lacked common definitions of indicators, leading to incomparable data and misconceptions of the status in different areas.

21.2.6 OPSA Forums

Roughly once a month, the OPSA officers organised an Effects Assessment Board (EAB) meeting to discuss the status of effects and the impact on planned operations with key staff personnel. Initially, the idea was that all section chiefs and sub-commanders would participate. In practice, this proved infeasible and also unnecessary as not all sections or units were contributing directly to the mission, but this effort was by nature Task Force oriented instead of mission oriented. In the end the participants of this meeting were the Commander, Chief of Staff, POLAD, DEVAD and the commanders of the BG, PRT and MRTF. When required or on their request, other sections were consulted on an individual basis prior to the EAB. Although all effects were continuously monitored, only the effects most important to the mission were discussed during the EAB meetings. These remained constant throughout the Dutch contribution to the ISAF mission. The added value of these EAB meetings was to create a thorough and well-founded situational awareness at mission level and to provide guidelines to the weekly planners meeting, which in turn, made operations more effective. In that weekly meeting, the planners of all TFU staff sections and subordinate units gathered to plan and synchronise their activities for the coming period. In the early stage of the TFU mission, recommendations from the EAB mainly related to the amount of effort and/or the timing of activities, and no major rescheduling of operations was recommended by the EAB. Such rescheduling only occurred when RCS planned a major operation in a neighbouring province that demanded scarce TFU resources or an operation caused an overspill into Uruzgan and which had an influence on every aspect of planning and the associated levels of ambition.

21.2.7 OPSA Products

The OPSA officers created a range of products. Firstly, they prepared the draft of the weekly commander's assessment that was sent to the higher ISAF command and to the national command line. This report commented on major events of the past week and assessed their impact on the effects to be achieved and on the mission as a whole. Secondly, the input for the EAB meeting consisted of a presentation with customised slides that had the function of a performance dashboard. The status of each effect was graphically depicted on a slide bar[1], providing the current status and the change compared the last assessment. This graphic summary was always accompanied by a description and argumentation of the assessment. Subsequently, the same slides were used to brief the entire staff. The third main product was a three-monthly document containing a comprehensive assessment of all effects, which was a requisite from the national command line.

[1] Editor's Note: Commanders frequently like this – it is simple and communicates the appearance of progress. However, the use of slide bars on a status spectrum is analytically problematic because there is usually no precise way to locate the slide bar, and therefore the temptation for the analyst to nudge the slide bar to the right can become overpowering.

For each effect, positive and negative influences were described and causally-related to efforts of the main actors in the environment. This document was highly appreciated by the TFU staff, the national command line, and especially by the succeeding TFU rotations that were in their preparatory phase, because it provided them with the situational awareness they needed to prepare for their mission.

21.2.8 OPSA Team Strategy

During the entire four years, OPSA was carried out under the direction of the Chief Plans (G5), where a team of two OPSA officers was embedded. The OPSA officers typically deployed for four months, with one person being relieved every two months. This strategy ensured that one OPSA officer was always sufficiently familiar with the dynamics of the mission, the current state of the assessments, and the relational network within the TFU. The system also ensured relative independence of the OPSA officers and moderation between the tendency of incoming staffs to assess the situation more negatively, and outgoing staffs to assess the situation more positively in order to emphasise their own contributions. To be able to support the entire mission with OPSA teams, a pool of around 20 Dutch OPSA specialists was created. Besides physical and motivational aptitude, each of the specialists was ranked on three criteria: ability to function and lead a firm debate in a military staff, advisory skills, and the methodological skills to work with a wide range of quantitative or qualitative methods. The objective was that at any moment of the mission there would be a balanced OPSA team. In this way, also the skills to apply several modelling techniques and software tools to perform analysis could be managed.

21.2.9 OPSA Tool and Methods

Due to the qualitative nature and limited amount of data, the OPSA officers mainly relied on extensive use of basic Microsoft Office software tools. One specific tool that was used by TFU is Method to Analyse Relations between Variables using Enriched Loops (MARVEL) which is based on system dynamics [1]. The strength of MARVEL is that it provides a common picture of the interrelations between interventions and effects in a complex environment.

Also other methods, techniques and tools to elicit information and views of subject matter experts were used. Group facilitation, workshop methods (individual interviews, group discussions, brainstorm sessions) were frequently used to exchange views and to develop a common shared understanding. Apart from group facilitation, there were also several IT tools, such as Excel, PowerPoint, I2 Analyst's Notebook and UCINET, that supported discussions by structuring and visualising data.

21.3 TASK FORCE URUZGAN 2008 – 2010

Initially, the Dutch government signed for a mission of two years, with an operational lead between August 2006 and July 2008. In the early part of 2008 the Dutch government and ISAF had negotiated an extension of another two years, which meant the Dutch would be in the lead of TFU until July 2010. After two years in Uruzgan, the Netherlands developed insight in the complexity of the mission and gained experience with the implementation of EBAO. These two years provided valuable insights for OPSA. An important lesson was how to execute a mission in a complex environment that changes constantly and is characterised by interdependency of many parties with different interests and motives.

With the new term of two years the TFU mission had also changed. TFU was confronted with the increasing complexity of the set of actors. More international enablers fell under the lead of TFU, such as the (AUS) Reconstruction Task Force, the (SGP) Air Defence unit and the (SVK) guard unit, the number of NGOs in Uruzgan had increased considerably from six in 2006 to thirty in 2009. Also the worldwide opinion on the military-in-the-lead role of ISAF (including TFU) had changed into a vision of shared responsibilities between military and civilian organisations, and of explicitly of increasing Afghan involvement through the

cooperation with the local and national Afghan governments, local NGOs and local contractors. Furthermore, the capability of the Afghan security forces had improved, enabling them to be involved in partnered operations with ISAF troops. Besides the growing civilian role and the increased Afghan involvement, the internationalisation of the mission had increased with military and civilian personnel from seven nations and with an expansion of American units in the south of Afghanistan. The differences between civilian, military, and host nation planning, also required a common style of planning to be understood and accepted by all. The increased number of involved actors emphasised the need to align objectives to support, and not contradict, each other's efforts.

21.3.1 Link Between Planning and Assessment

Whereas OPSA has been a continuous effort, the OPSA officers in 2008 identified a gap in the link between two planning products, namely, the long-term EBAO plan and shorter-term OPLAN. In the first TFU staff the overall long-term EBAO plan and Military Operational Plan (OPLAN) (assigning troops to tasks for a period of six months) were well synchronised with OPSA. Each TFU staff had to develop an OPLAN that was derived from the EBAO Campaign Plan that also encompassed civilian interests and objectives. Subsequent rotations seem to have relied on the first EBAO Campaign Plan and focused mostly on developing the compulsory military OPLAN rather than also updating the ambitions of the EBAO Campaign Plan, see Figure 21-4.

Figure 21-4: Assessment and Planning Linkage.

Both the insight in the complex environment and these important OPSA-related lessons learned led to the development of a new process to collaboratively develop an iterative plan to create unity of effort amongst the set of actors. This process, named the Collaborative Decision-Making (CDM) process, also resulted in the Uruzgan Campaign Plan (UCP), a non-rigid civil and military planning document that translated a conceptual integrated long-term ambition into an iterative approach that could be implemented by all participating parties. This plan provided a structure to link military and stabilisation short-, mid-, and long-term objectives using an iterative (step-by-step) approach that could easily be adjusted to changes in the environment (either induced by TFU or external factors). The alignment of objectives made it possible to incorporate the measurement of progress and results. This meant the link between military and overall planning was re-established with OPSA to better contribute to planning, decision making, and more effective operations.

21.3.2 Structure and Function of the UCP

The backbone of the UCP was a conceptual framework that described all relevant factors concerning the mission, divided into development themes, factors to mitigate, and factors to exploit. Development themes are factors that the TFU wants to contribute to, like governance, healthcare or economic development. Long-term progress on these factors is desired and therefore enabling and contributing to progress on the development themes was the main effort of TFU. Progress on the development themes could be hampered by things such as local conflicts or narcotics, which are factors to mitigate in the conceptual framework. Factors to mitigate should be prevented from having an influence on the progress of development themes. Factors with a positive influence on the progress of the development themes are the factors to exploit. These factors enhance the progress of development themes, e.g., new and more IO/GO/NGO activities or national (Afghan-led) programmes, and should be exploited to achieve a positive influence. For all factors in the conceptual framework short-, mid-, and long-term objectives were developed, see Figure 21-5.

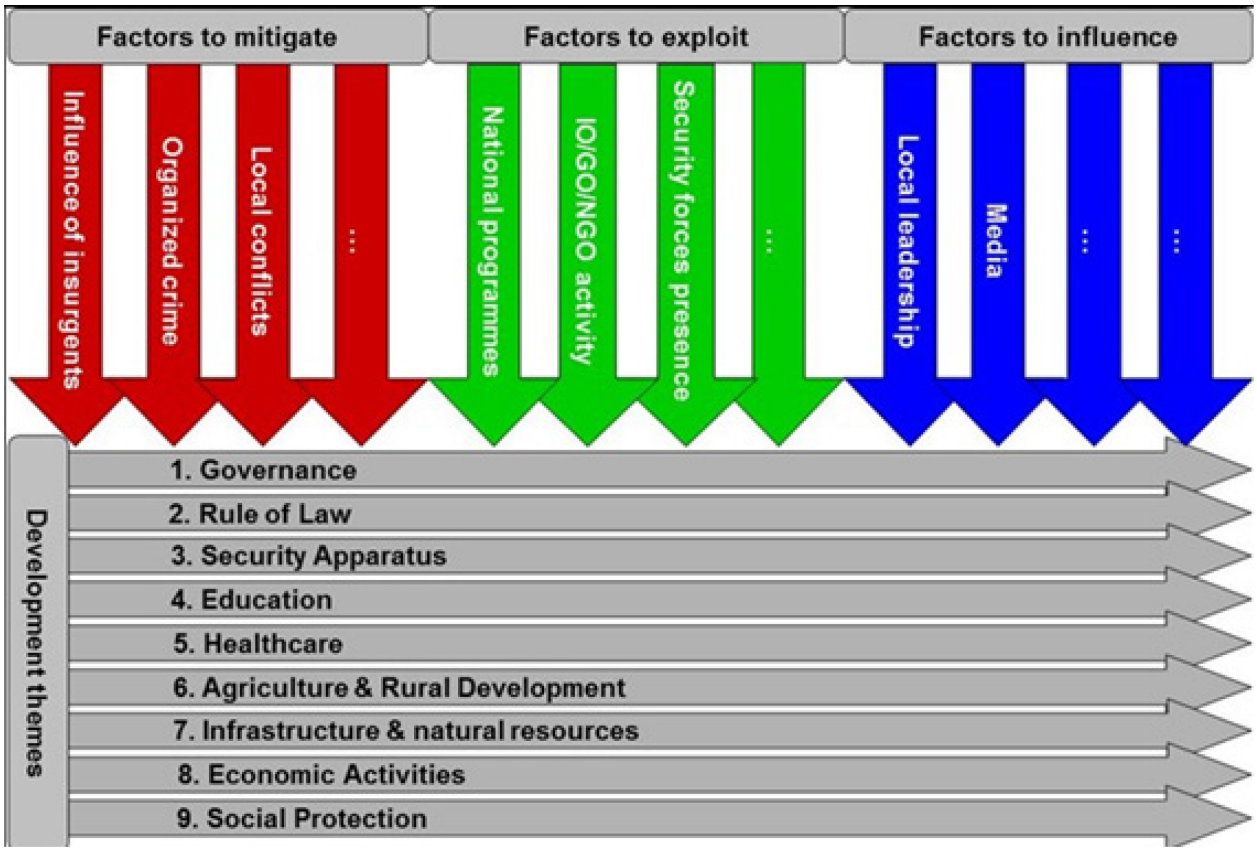


Figure 21-5: Structure of the Uruzgan Campaign Plan.

What made the UCP different compared to traditional military plans? And how was the incorporation of assessment ensured? As the end state in a complex multiparty environment cannot be defined unambiguously, the UCP ambitions were condition-based instead of time-based. It offered rough objectives for the long-term (decades), desired situations for the mid-term (years), and realistic and specific effects for the short-term (months). This increased the agility of the plan and subsequent activities and operations to (un)expected changes in the mission environment. The UCP was updated in regular iterations, at least with each new TFU rotation. Since its initial design, the plan went through two updates until the last Dutch TFU rotation.

21.3.3 Reframing of TFU Effects

Because of the impacts of the contributions of other actors, a distinction was made between effects to achieve and effects to contribute to. Effects to achieve are effects that TFU could realise through its own operations and activities, such as providing a course in sustainable entrepreneurship, through which entrepreneurship in the province could be stimulated. Effects to achieve were easy to translate in specific activities that would contribute to the long- and mid-term objectives. Effects to contribute to are effects that TFU could only realise by supporting other actors or through small contributions. An example is "Branch office of the Kabul Bank opened in Tarin Kowt" to develop the finance sector. Both the objectives for the long-term and mid-term were regarded as overarching to TFU staff rotations. The effects to achieve and the effects to contribute to were revised and updated every six months by a new rotation of staff. These effects were translated into operational plans with specific tasks to military units and civilian advisors.

21.3.4 Measurement of Progress

To measure progress toward the effects and the objectives, a measurement plan was developed. The development of this plan was similar to the OPSA process described for the start of the TFU. The most important difference was the framework of the UCP that related all stabilisation and military objectives for the long-, mid-, and short-term. To measure progress, metrics were defined. This was not only a job for OPSA. Expert opinions from all relevant actors were collected and integrated into a draft measurement plan. OPSA's job was to refine the set of metrics to be measurable. These metrics could be quantitative (e.g., number of entrepreneurship lessons provided) or qualitative (e.g., motivation of the students), with most of the metrics in the latter category. Subsequently, the subject matter experts were involved to discuss whether and how the metrics could be observed and how the combination of metrics could provide valuable information in order to make operations more effective. The draft measurement plan was then presented and finalised by a discussion within the TFU.

21.3.5 Data Collection and Analysis

Data were collected and monitored continuously, and every three months, OPSA published a report with an analysis of the collected data related to the TFU's activities. Since information could be little or highly dependent upon subject matter experts' opinions, the line of reasoning behind the analysis was very important. Staff meetings were held to gain input upon and discuss the line of reasoning that made sense of the changes in the environment reflected in the metrics. Conclusions drawn were input for new planning processes or adjustment of existing plans to make operations more effective.

21.3.6 OPSA Products

The main difference between the OPSA products of the first two years and those of the second period was that the OPSA officers also provided input for the measurement of mission progress on each factor in the conceptual framework. This analysis was based upon the measurement of change on the effects to achieve and the effects to contribute to, providing information on the difference between planned and executed activities (measurement of performance), and other external influences. The sense-making process that gave meaning to the observed changes in the mission environment was one of the most important outcomes of the OPSA process. These discussions led to valuable insights throughout the TFU staff that influenced planning and decision making to optimise mission effectiveness.

21.3.7 OPSA and Overall Mission Evaluation

At the beginning of 2010, it became clear that the Dutch government would end their leading role in the province of Uruzgan. Again, this gave a new dynamic to the role of OPSA and the OPSA team. Whereas the role of OPSA in the beginning was mostly to collect data to look forward, it was now a time to manage the

transfer of all knowledge that has been built up in four years and to provide useful information packages to the coalition that would succeed the Dutch contribution. Another important role for OPSA was to create a comprehensive snapshot that captured four years of the Dutch contribution. Senior leaders, such as the TFU commanders and the Commander of the Dutch Armed Forces, assumed that these tasks should fall naturally to the OPSA team. This demonstrated the value of OPSA to the mission, and reflected leaders' faith that the OPSA team would do the best possible job. This is an achievement that could only be established by being involved in the heart of the mission for the full mission length.

21.4 A NEW MISSION: THE POLICE TRAINING GROUP IN KUNDUZ

As the Netherlands was closing down its contribution to the TFU, the Dutch government decided in January 2011 to maintain a contribution to ISAF through the establishment of a police training mission in the northern province of Kunduz. The decision to commence with a new mission in a different part of Afghanistan was much discussed in Parliament. As a result, the mission incorporated many national caveats that justified and required another OPSA team.

The mission of the Police Training Group (PTG) was to strengthen and improve all institutions involved in the rule of law in the province of Kunduz. This included training the local police (from patrolmen to senior levels), educating the judiciary (courts, judges as well as prosecutors), and strengthening civil society groups active in the field of rule of law. Thus, the mission was much more comprehensive than the relatively straightforward task of training police officers.

This clearly was a mission that could not be handled solely by the military. Building on the experience gained in Uruzgan, the mission was designed to be executed through the close cooperation of the ministries of Defence, Foreign Affairs, and Justice and Security. From the outset, it was decided that the mission would have a shared leadership, consisting of the three contributing government ministries. In practice, this meant regular meetings between representatives of the three ministries in the Netherlands, and similar meetings in theatre.

21.4.1 Design Phase

As the mission required the integrated efforts of three ministries, the usual (military) approach was not applicable. The question was how to design a complex mission, orchestrating the short- and long-term efforts of three ministries, along widely divergent lines of operation. Based on the experiences and knowledge gained in the mission in Uruzgan, the mission in Kunduz was designed using a civil-military design process (the collaborative decision-making process). This led to a long-term plan for the mission in Kunduz and familiarised the participants with each other to create a measure of like-mindedness that was of benefit during the entire mission.

21.4.2 In Theatre

The staff of the Dutch PTG was basically an expanded battalion staff, with a relatively large intelligence section, and a number of civilians from the Ministry of Foreign Affairs. The military commander was a colonel. The OPSA effort was thus not solely for the benefit of the military commander, instead OPSA officers reported to all three members of the integrated command team. To facilitate cooperation between representatives of the three ministries in theatre, a number of boards were created which met regularly. Once a month, the outcomes of these board meetings were discussed by the highest representatives and adjustments were made, if necessary, to make operations more effective.

21.4.3 Data Collection and Analysis

Data collection was as complicated as in Uruzgan. As in Uruzgan, the metrics required for the OPSA effort were incorporated into the Intelligence Collection Plan. Again, the intelligence units were more focused on terrain and enemy. Thus, OPSA officers needed to collect most mission effectiveness data by themselves, through cultivating contacts and regular interviews. As Kunduz was a less hostile environment than Uruzgan, there was a broad range of NGOs that provided interesting sources of information. Furthermore, there were American, German, and Belgian units in the area that could provide valuable insights into the functioning of local police and judiciary system. In all, the OPSA officers faced similar challenges in collecting data as in Uruzgan. Qualitative data, often through expert opinion, was of vital importance and quantitative data were absent, incomplete, or unreliable. For example, it was virtually impossible to determine the exact number of judges in Kunduz province, and their educational backgrounds. It was possible to construct an overview of the status of judges in Kunduz only through interviews.

21.4.4 OPSA Forums and Products

In the (mostly) military staff of the PTG, the OPSA officers were tasked with assessing the entire PTG mission. As such, they were the officers with the broadest overview of the mission, where others were focused solely on their efforts to train the police, judiciary, or civil society institutions. Thus, the OPSA officers brought data together and constructed an overall operating picture in support of the three members of the integrated command team. The OPSA officers were members of the various boards that managed the effectiveness of day-to-day activities.

21.4.5 Comparison to OPSA Within TFU

Kunduz was a very different mission from Uruzgan, and based on experiences gained in Uruzgan it was decided to deploy, from the outset, an integrated mission, led jointly by representatives from the three participating ministries. Although this complicated the mission and created frictions in the day-to-day direction of efforts, the OPSA effort was similar to that in Uruzgan. The OPSA officers were performing the same tasks, employing a similar methodology, but reporting to a triumvirate rather than single military commander. As in Uruzgan, the OPSA officers were tasked with writing the commanders' weekly assessment report and the larger three-monthly report. They sat in on the various briefings and staff meetings determined in the battle rhythm, had access to a wide range of reports, and frequently met with commanders, representatives, and key staff officers. The hefty three-monthly report was especially important, and sensitive, as all three ministries wanted to report good progress. This did put some pressure on the integrity of the assessment products, but this was mitigated by finding the appropriate words and formulations that all parties could agree with, without changing underlying facts or essence of the conclusions. The document became somewhat more "political" as a consequence. If anything can be said to characterise the OPSA effort in Kunduz, it is that in a complex, whole-of-government mission such as the PTG, the importance and impact of the OPSA effort is only enlarged.

21.5 OVERALL OBSERVATIONS AND LESSONS LEARNED

The most important lesson from the Dutch missions in Afghanistan is that the ideal of a largely quantitative approach to OPSA must be abandoned. In complex missions, set in a non-permissive environment, data collection is problematic. Data are absent, incomplete, or not to be trusted, and there is a tendency that numerical data may become politically charged and subject to manipulation by senior officers wishing to report "good numbers". Offsetting these shortcomings, collecting qualitative data is an absolute requirement for any OPSA effort. As a consequence, OPSA officers need to be well-versed in methodologies associated with qualitative data collection, processing, and analysis. Besides data collection plans being made, also the awareness and focus on tasking assets for the collection of mission data, needs improvement.

With the need to steer a complex mission, OPSA officers are instrumental in developing a long-term mission design and supporting OPSA products. A mission design should encompass the objectives of all main actors (civil and military) and blend in with the different planning cultures. In this manner, the understanding and commitment to the mission among all is maximised. The OPSA products and the way they are visualised should facilitate discussions among the decision makers of all actors.

Where OPSA was instrumental in monitoring and steering the long-term perspective during the mission, at the end of the TFU it also provided the most important database to look backward and report on the noticeable change and results of the four-year Dutch effort.

As OPSA doctrine was non-existent in 2006, the current Dutch OPSA doctrine has been established in a learning-by-doing manner, most of which was developed in the TFU era and elaborated on later. The NATO Operations Assessment Handbook, with the first version published in early 2011, was developed in parallel. The TFU-based OPSA process very much resembled the process later described in this Handbook – data were collected and analysed, discussed with key staff personnel, and the resulting conclusions were input to the commander's decision-making process.

Although NATO doctrine states that all aspects of OPSA mainly occur at the operational level of command, and that the tactical level should merely focus on measurement of activities [2], the Dutch experience proved that OPSA can also be useful at lower (tactical) levels. Our experiences are that OPSA was a valuable tool for a unit with the size of a small brigade. Applicability of OPSA is not solely dependent on hierarchy or unit size, rather the criteria for a successful OPSA application are the complexity of the environment, and the capability and mandate to plan and execute operations while employing different instruments of power.

The strategy to maintain an independent rotation cycle with two-person OPSA teams has been one of the key success factors for the long-term OPSA contribution. Two persons with the right competences have proven to be the minimum effective team. Also the fact that OPSA has been organised from reservist positions is seen as a benefit, since it guarantees a level of independence to all actors as no career perspectives or other affiliations are at stake.

Even though OPSA contributions are well appreciated, for every new mission the challenge for OPSA remains to be involved at the right moment and at the right level. Maintaining OPSA knowledge and expertise during the peacetime establishment with exercises is difficult. The complexity in exercise scenarios and available data on the exercise mission environments cannot be compared with the volume of data in a real mission environment. This shortage means that OPSA efforts during exercises can only be practised technically, as the artificiality of an exercise, including the played time line, leads to incidents being marked as trends and so on. A risk of this role in exercises is that it might negatively affect the opinions on OPSA contribution in real missions.

21.6 AUTHORS' BIOGRAPHIES

All authors have been deployed at least twice as Operational Analyst as part of Dutch contingents in Afghanistan in the timeframe of 2003 to 2011. All are still active as reserve officers in 1 (NLD) Civil Military Interaction Command, providing OPSA support to the Netherlands Army manoeuvre brigades, the Netherlands Maritime Forces (NLMARFOR) and 1 (German/Netherlands) Corps.

Aletta Eikelboom studied Artificial Intelligence and works at The Netherlands Organisation for Applied Scientific Research (TNO) in the department of Military Operations as project leader and consultant focused on the topics of command and control, collaborative decision making and intelligence. She has deployed twice in Task Force Uruzgan (TFU) and facilitated the development of the Comprehensive Mission Design of the Police Training Group (PTG) in Kunduz, Afghanistan.

Rudi Gouweleeuw studied Computer Science and works at TNO in the department of Military Operations as programme manager in the field of land-based operations and cyber operations. He has been deployed to HQ ISAF and three times in the TFU. He acts as a subject matter expert in assessment at the Netherlands Defence Staff College and the NATO School Oberammergau.

Geert Roseboom studied Mathematics and Computer Science. He works as research manager of the department Modelling, Simulation & Gaming in the TNO Unit Defence, Safety & Security. He has been deployed twice in the TFU and also acts as a subject matter expert in assessment at the NATO School Oberammergau.

Jeffrey Schwerzel studied Cultural Anthropology. He has been deployed twice to Uruzgan, Afghanistan, and was the first Operational Analyst of the PTG Kunduz, Afghanistan.

21.7 REFERENCES

[1] Veldhuis, G.A., van Scheepstal, P., Rouwette, E. And Logtens, T. 2015. Collaborative problem structuring using MARVEL. EURO Journal on Decision Processes, 3(3-4):249-273.

[2] NATO. 2014. NATO Operations Assessment handbook; Version 3.0, September.

Chapter 22 – OPERATIONS ASSESSMENTS WITH A DECISIVE ACTION FOCUS

Andrew Swedberg
US Army
UNITED STATES

ABSTRACT

Although doctrine and best practices exist for conducting assessments in long-term stability operations, the complexities of conducting an operational assessment in a fast-paced decisive action fight, involving intense conventional combat, such as that represented in an U.S. Army Warfighter Exercise (WFX), can be a challenging endeavour. This chapter outlines a framework for an assessment method for U.S. Army Corps and Division assessors as their organizations conduct decisive action operations. It explains a method to integrate the warfighting staff with the purpose of assessing the operation's plan. The focus of the integration is to answer: "is the plan still valid?" and develop recommendations for actions to take as the conditions change. This method seeks to encompass the entire staff into the often least emphasised aspect of the operations process – assessment.

22.1 BACKGROUND

Within the Department of Defense, an "assessment" has a few different definitions. Likely the most comprehensive definition exists within U.S. Joint Publication 5-0, which defines assessment as "a continuous activity that supports decision making by ascertaining progress toward accomplishing a task, creating an effect, achieving an objective, or attaining an end state for the purpose of developing, adapting, and refining plans and for making campaigns and operations more effective" [1]. The publication "Operation Assessment – Multi-Service Tactics, Techniques, and Procedures for Operation Assessment", defines assessment as "the continuous cycle of observing and empirically evaluating the ever-changing Operational Environment (OE) to inform decisions about the future and make operations effective" [2].

Commanders develop and maintain a sense of the operation's progress through battlefield circulation, conversations with senior and subordinate commanders, and key leader engagements. A supporting operations assessment complements the commander's awareness by identifying changes in the operating OE, analysing upcoming risks and opportunities, and providing recommendations to improve progress towards mission accomplishment [1]. Although doctrine and best practices exist for conducting assessments in long-term stability operations, the complexities of conducting an operational assessment in a fast-paced decisive action fight, involving intense conventional combat, such as that represented in an Army Warfighter Exercise (WFX), can be a challenging endeavour. Recently, an Army Corps headquarters executed a WFX through a decisive action operation with an offensive operation focus. Recent assessment practices, such as using Measures Of Effectiveness (MOEs) and Measures Of Performance (MOPs) as indicators, is largely based on stability operations and is meant to show operational effectiveness and progress towards attaining a long-term end state based on Lines Of Effort (LOEs), Lines Of Operation (LOOs), and associated objectives and effects. For many Army exercises and other training designed for future missions, decisive action is the main focus of operations, and LOEs or LOOs may not be necessary. This chapter is meant to outline a framework for an assessment method for Army Corps and Division assessors as their organizations conduct decisive action operations. In particular, Operations Research – Systems Analysis (ORSAs) on Corps and Division staffs, who find themselves in assessment roles, may be able to consider the method explained in this chapter. Translating assessment doctrine into a decisive action environment can be challenging, but multiple exercises from the author's organization has polished a rough assessments process into one that can help other Corps and Division-level ORSAs and others in assessment billets.

22.2 PLAN TO ASSESS

One U.S. Army Corps recently conducted a series of high profile decisive action training exercises in 2017, including a Forces Command [3]. FORSCOM trains and prepares a combat ready, globally responsive Total Force in order to build and sustain readiness to meet Combatant Command requirements [3]. The purpose of the WFX was to develop core warfighting competencies in accordance with Corps' Mission Essential Task Lists (METL), which are a combat unit's wartime mission and related tasks. The exercise simulated a corps-level operation which exercised the Corps staff and command team's ability to command and control subordinate units during an offensive operation. Given the nature of the decisive action exercise, planners used a planning framework that led to a mission statement, key tasks, and end state along with a comprehensive Operation Order (OPORD). They did not plan to use LOEs, LOOs, objectives, or effects that would be found in a longer-term stability type of operation. As such, rather than building an assessment framework that linked objectives and effects to indicators, the assessment team decomposed the end state into measurable objectives using the approved mission statement and key tasks. These objectives aligned with indicators by phase, in order to demonstrate progress toward the end state, and provide insights that enable the Corps to improve the effectiveness of operations. Critical to this concept was establishing success criteria by continually asking "How can we know we are achieving the end state?" Each indicator was assigned to one or more Offices of Primary Responsibility (OPR), who were the staff elements best able to provide a status on each indicator. The OPR compared the current value of any indicator to the relevant success criterion, which enabled an OPR to determine if the indicators were "on track" or "off track" and provided an opportunity to adjust operations to be more effective.

22.3 ASSESSMENT WORKING GROUP

An Assessment Working Group (AWG) is necessary for the assessment lead to understand the state of the plan from the perspective of each Warfighting Function (WFF). For the U.S. Army, there are six WFF – intelligence, fires, movement and manoeuvre, sustainment, protection, and mission command [4] pp. 4-3. This battle rhythm event is the one focused opportunity that the assessment lead has to gather staff officers in order to analyse the changing OE and cultivate recommendations to the plan. During the WFX, the AWG met daily for one hour. The initial focus of the group was to measure operation effectiveness and progress towards the end state through the data collection framework. For areas that were not "on track," the group identified specific reasons why they were off plan. For the WFX battle rhythm, the assessment was briefed "as required" as opportunities for improving effectiveness emerged. Table 22-1 shows a generic example of each how indicators can be collected in order to provide an "assessment" of if the plan is on track or not.

22.4 FIRST ASSESSMENT CYCLE AND ADJUSTMENT

During the first briefing with the Commanding General (CG), the assessment team gained insightful feedback on the assessment product. Essentially, the developed framework was sufficient, but only went as far as the status of the objectives, which did not provide the assessment the commander needed. He stated that he wanted to first know if the plan was still valid. If the plan is not valid, he wanted recommendations to get back on track, along with associated risks and opportunities. While there was a good set of defined colour-coded bubble charts to communicate the status of indicators, but there was a poor synthesis of these into a recommendation. In this operation, the commander knew the end state. We learned that the recommendation should be framed in terms of if the operation was on plan in time and space, and if not, what planning efforts (branches or sequels) would get us back on plan, assuming the plan is still valid. During this first cycle, the AWG conducted an assessment of past events in the operation, but learned that this did not aid in decision making or suggest where adjustments to operation were required. Figure 22-1 shows a generic format with objectives and an assessment of if the objective was "on plan" (green), "off plan – minor" (yellow), or "off plan – major" (red). Although this product communicated the status of indicators in a simple way, it did not enable the commander with a

recommendation to improve the plan. Initially, the AWG also tried to roll up colour-assigned indicators into larger on or off track categories. As a result, that became the focus of the AWG, which was not useful, as it left little time to develop clear and focused recommendations, which were really what the CG needed for an assessment. Following this briefing to the CG, the assessment lead recalibrated the AWG and assessment focus.

Table 22-1: Data Collection Framework for Indicator Collection.

Objective	Sub-objective	Indicator	OPR	Success Criteria
Defeat XYZ	Defeat enemy will to fight	Is enemy moving in a certain direction?	G2, G33	No enemy movement west of Phase Line ABC.
		Are there abandoned enemy combat vehicles discovered?	G2, G33	Enemy tanks and combat vehicles abandoned along MSRs.
		Is the enemy surrendering? Rank, Size, Unit	G2, G33	Company sized elements from enemy unit N surrendering.
		Are enemy elements getting smaller as we progress?	G2, G39	Enemy elements reduced to squad size.
	Decrease enemy maneuver element the ability to fight	What is the enemy combat strength percent?	G2	35% aggregate or less
		Are there any High Value Targets (HVT) destroyed?	G2	Command and control high value targets destroyed (enemy unable to C2).
		Are subordinate blue units on plan with location and time?	G33, G35	Friendly unit A at Objective M by 0800L.

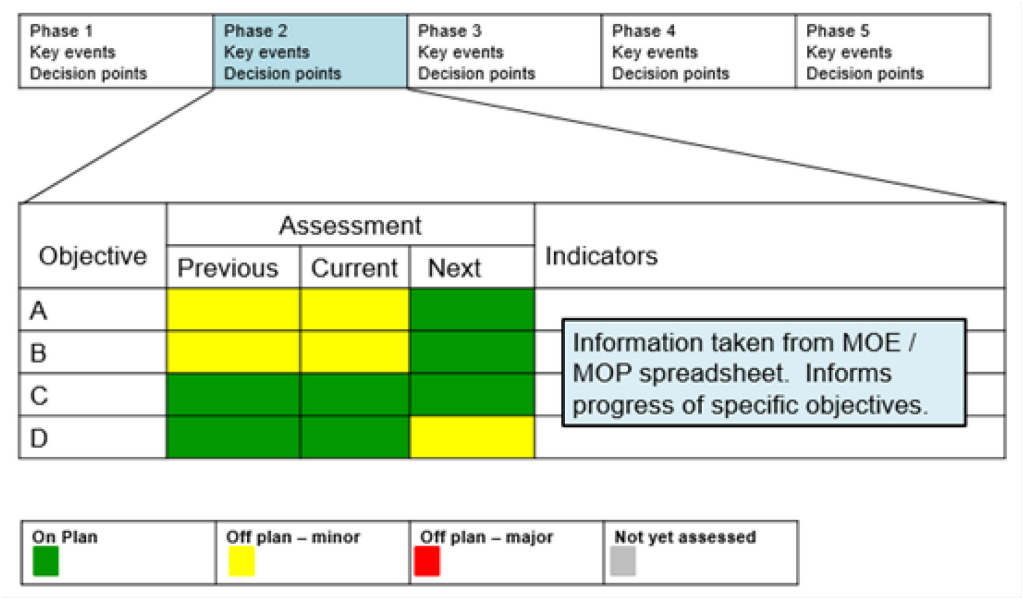


Figure 22-1: A Generic "Operation Assessment" Communication Product with Criteria.

22.5 REFRAMED ASSESSMENT PROCESS

After guidance from the CG, the assessment lead reframed the AWG into a different format for participants. It moved away from a detailed step-by-step walk through of each MOE and MOP with the goal of assigning each with a colour. Although detailed analysis of indicators seemed to provide the "science" behind the

assessment, it did little to provide a recommendation for a way ahead. In the reframed assessment process, each AWG attendee was empowered to answer the question – "under current or expected conditions, is the plan still valid?" Rather than devote most of the AWG time to coding the indicators, each staff section was asked to consider their warfighting function and staff section, and then identify opportunities and risks in the form of gaps, seams, and shortfalls. Indicators (MOEs and MOPs) would still be used, but they would be evidence, rather than the primary focus of the assessment.

The AWG used the definitions in Table 22-2 order to guide the AWG attendees in discovering opportunities and risks, and thereby developing recommendations for a way ahead. Using definitions from previous ORSA best practices, the assessments team developed a set of ratings to guide the AWG during the valuable time they had together daily [5]. According to their specific staff element, the AWG attendees could judge if they were on a trajectory to see success within the plan, or if there should be some type of adjustment.

Table 22-2: Definitions of Assessment Criteria.

Opportunity	A possibility exists to exceed the end state by time or effect; an exploitable set of circumstances. Certain conditions have created an opportunity to accelerate the plan or increase effects if we additionally resource or conduct these activities.
Risk	The gap, seam, or shortfall (see below) has an associated risk if we do not address it. The probability is high/med/low and the consequence would be this. Recommend mitigation by doing an action.
Gap	A void in the plan. There is no plan, or the plan is not valid. Conditions have created a significant gap in the plan that will change our end state. We need to create a sequel in order to get back on track.
Seam	The plan is valid but requires staff synchronization. Additional coordination is needed for success. Conditions have created a point of friction in the plan. The plan is still valid, but we need to create a FRAGORD or branch plan in order to get back on track.
Shortfall	The plan is valid, but resources are insufficient (time, funding, resourcing, commodity). We have a shortfall of this resource impacting achievement of the objective. Recommend requesting additional resources.

We also organized our assessment to reflect a new pattern – by organization of unified land operations. This was an adjustment to consider deep, close, and support operations in terms of the WFF, who were already represented by the same AWG attendees. This adjustment allowed for considering operations in space and time. The corps deep area generally extends beyond subordinate division boundaries to the limits of the area of operation. As the operation progresses, the deep area transitions to the close area. The close area was considered to be within the immediate timeframe and within subordinate division's fight. The support area was considered to be operations behind the corps, and would include operations in the expanding rear area as the decisive action fight progressed (see Ref. [4], pp. 4-6).

During the reframed AWG, intelligence representatives set the stage by discussing risks or opportunities with the templated enemy set and disposition over the next two phases of the operation. Next, the Fires WFF and shaping communities, such as air, information operations, and civil affairs components, were able to discuss risks and opportunities with lethal and nonlethal fires and effects that should occur within the three battlefield areas. Movement and Manoeuvre WFF (including the current operations and future operations sections) was next, followed by the Sustainment, Protection, and Mission Command WFFs. Each WFF and staff element had the opportunity to indicate risks and opportunities to their specific piece of the plan, and the intention was to allow the discussion to generate into a larger collection of recommendations. Through this process, the AWG

participants developed recommendations that would get the plan on track or capitalise on the current situation. Figure 22-2 shows a methodology for uncovering risks and opportunities within the AWG.

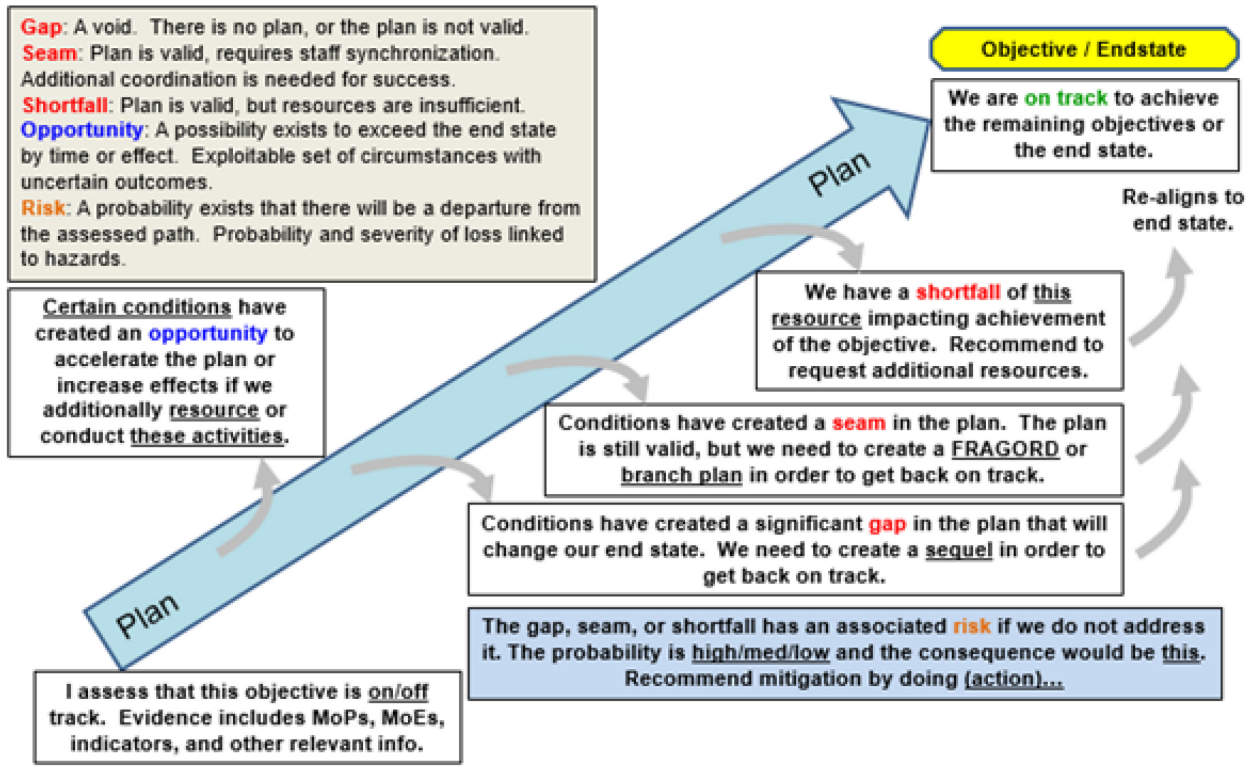


Figure 22-2: A Method for Uncovering Risks and Opportunities Within the Plan.

The assessment lead should prepare questions to spur conversation. Some examples may be:

- What is the enemy's most likely behaviour at a specific objective?
- How much combat power is needed to secure/fix/destroy/seize an objective?
- Does the Fires WFF see a problem with radar coverage? Any concerns of having enough rounds of the right type? Are there security requirements that are not yet organized?
- What nonlethal effects worked and what did not? How do you know? What evidence do you have?
- Are there any unplanned for civil considerations that are needed for displaced civilians?
- Is there an unexpected security risk for sustainment operations?
- Do Main Supply Routes (MSRs) require unexpected maintenance or security?
- Will there be any issue of maintaining communication retransmittal security?

The discussion findings were combined into collections of risks and opportunities by WFF, with recommendations on how to proceed. These recommendations were colour-coded for ease of reference in the briefing product. The colours describe the AWG's recommendation about how to mitigate the risk:

- Blue (overwhelming success and opportunity);
- Green (on plan and desired);

- Yellow (recommend a refinement to the plan via FRAGORD or other staff action);
- Orange (recommend a branch plan); and
- Red (recommend a sequel).

Table 22-3 shows a notional briefing product that summarizes the risks and opportunities by WFF and phase. There would also be another slide that explains the risks and opportunities in a concise manner. The branches and sequels would require G35 Future Operations and G5 Plan planning efforts, respectively, in order to re-align with the end state.

Table 22-3: Notional Briefing Product that Bins Risks and Opportunities with Recommendations.

"How will you know if you are on-plan? Through Indicators	Intelligence	Fires	Movement and Maneuver	Sustainment	Protection	Mission Command
Staff Section	G2	G3 Fires, G39, G9, Air Force, G3 Air	G3, subordinate commands, G35, G5	G1, G4	Protection, EN	G6
Focus	Enemy, Collection, Terrain, Weather	Lethal, nonlethal, Civilians, PSYOPS, Cyber, Space, Irregular targeting	Current operations, future operations, Plans	Logistics, Personnel, Religious, Medical, Legal	Protection plan, mobility plan	Communication and command and control
Deep						
Close						
Security						

Success / Opportunity	On Plan	Plan refinement – FRAGORD	Recommend Branch Plan	Recommend Sequel	Not Yet Assessed.

Another critical contribution to the assessments process was the formation of an Assessment Board, chaired by the G5 or Chief of Staff (CoS). Before the Commander receives recommendations for action, the recommendations must be vetted and validated through the senior officers on the primary staff. The board met and reviewed the work that the AWG developed, which helped to refine the risks and opportunities recommendations. This step permits and requires colonel-level primary staff members to have ownership of recommendations that go before the CG.

22.6 ASSESSMENT OBSERVATIONS

Considering an Army focus on decisive action operations, rather than stability only, may require expanding assessment best practices to the larger community. Based on historical best practices as well as experience in decisive action WFX exercises, the author recommends organizing assessment efforts for two types of operation as shown in Table 22-4. This breakdown could aid Corps and Division assessors into developing a holistic assessment plan based on the type of operation. In a real world scenario, as the phases of military operations transitions, the staff will adjust the assessment scheme accordingly.

Table 22-4: Most Likely Attributes Associated by Type of Military Operation.

	Decisive Action	Stability Operations
AWG Frequency	Daily.	Weekly / 2 x month.
Assessment Framework	Deep, Close, Support.	LOEs with leads.
Assessment Boards (COL and General Officer)	~Weekly, as required.	~Monthly.
Available Information	Limited, unfiltered data.	Able to develop historical database.
Focus	Next phase risks (gaps, seams, shortfalls) and opportunities.	Deliberate research to discover trends and recommend adjustments.

Decisive action fights for Division and Corps staffs would likely see the AWG occurring daily with a deep, close, and support operations framework that focuses to the next one and two tactical phases' risks and opportunities. Assessment Boards would exist as required by the Chief of Staff or Senior Leader. The feedback, adjustment, or approval from a colonel-level assessment board then becomes a product for the CG. The assessment conversation between the commander and staff becomes the primary benefit of the assessment, as opposed to having a product or report as a stand-alone document. As the operation transitions to stability operations, the assessment framework would likely transform into a model that is more familiar to assessors who conducted assessments in stability operations. However, this shift is gradual, as decisive action and stability operations can occur within one battlespace at the same time.

22.7 LESSONS LEARNED

Through the WFX assessment learning experience, the assessments team developed observations that may enable future assessors that plan assessments for decisive action operations. These observations are summarized:

- Operations assessment should be planned for deliberately while the planners are developing the OPORD. Assessors should work with planners to have definable objectives and end states by phase. If the planners do not develop LOEs, objectives, or effects, consider decomposing the end state into measurable components within a deep, close, and support operations framework. For decisive action, look beyond the 24-48 hour fight in order inform the commander and staff of risks and opportunities that will materialize.
- Every staff section has a role in creating their own indicators with success criteria for each. Assessment is everyone's business, not just the assessment team [6]. Indicators are important, but they assess the past, and only get to the equals sign of a hypothetical equation. The actualized assessment is the recommendation of what to do next in order to be more effective, based on evidence that the MOEs and MOPs provide.
- The AWG offered an avenue for staff synchronization and collaboration. The assessment lead, frequently an ORSA, does not have the expertise to have insight into every staff section and WFF. Therefore, each staff section should provide one attendee to assess the current validity of the plan from that section's perspective, along with risks (from gaps, seams, shortfalls) and opportunities and mitigating recommendations.
- The goal of the assessment process is not a final presentation or product, but rather the continually improved understanding of the operational environment. The assessment is not a stand-alone event,

but is the results of the conversation between the commander and the staff, and the decisions that result from that conversation. The entire staff deserves to have buy-in and validation of risks and opportunities.

A short training exercise allows for a limited set of assessment iterations. For example, a WFX may be less than two weeks and can only fit in a certain number of working groups, staff meetings, and decision briefings. However, each time, tremendous learning can occur between the commander and staff. This chapter should be able to guide Corps and Division staffs in establishing effective operations assessment processes in decisive action exercises and real world engagements as well as the stability operations that follow.

22.8 AUTHOR'S BIOGRAPHY

Lieutenant Colonel Andrew Swedberg is currently serving as the "Chief of Assessments" in U.S. Army I Corps at Joint Base Lewis-McChord, Washington. An Operations Research/Systems officer, he holds a Master's of Science degree in Applied Mathematics from the Naval Postgraduate School and a Bachelor of Science degree in civil engineering from the United States Military Academy at West Point, New York.

22.9 REFERENCES

[1] U.S. Joint Publication. 2017. Joint Publication 5-0: Joint Planning. VI-1, June.

[2] Air Land and Sea Application Centre. 2015. Operation Assessment: Multi-Service Tactics, Techniques, and Procedures for Operation Assessment: ATP 5-0.3, MCRP 5-1C, NTTP 5-01.3, AFTP 3-2.87.

[3] FORSCOM website. 2018. https://www.forscom.army.mil [Last Modified: 10 January 2018].

[4] Army Doctrine Reference Publication (ADRP). 2016. 3-0, Operations. November.

[5] Kramlich, G. 2013. Assessments vs. Decision Support. In the author's possession.

[6] Shilling, A. 2016. Centre for Army Analysis. A Quick Reference Guide to the New Paradigm of Operation Assessment. May. In the author's possession.

Chapter 23 – SIERRA LEONE: ASSESSMENT OF DEFENCE (AND SECURITY) SECTOR REFORM ACTIVITIES

Karsten Engelmann
US Army
UNITED STATES

Jim North
US Army
UNITED STATES

ABSTRACT

A significant effort at US Africa Command in the 2008 –2010 timeframe was to better understand how US efforts could assist African countries with weak security and defence sectors to successfully strengthen and reform those sectors. A research team visited Sierra Leone in 2010. Its purpose was to help the US Africa Command execute future security assistance activities more effectively. This chapter provides an approach for evaluating Security Sector Reform (SSR) and Defence Sector Reform (DSR) that was reflected in observations by a US Africa Command team evaluating SSR and DSR conditions in Sierra Leone as of 2010. The chapter takes information on the state of the security and defence sector as in Sierra Leone and evaluates progress toward achieving reform with the assistance of the Security/Defence Sector Reform Concept Model that US Africa Command developed (described below). The US Africa Command SSR/DSR Concept Model breaks Security Sector Reform into its sub-components: Reform of the non-military Security Sectors (Justice, Police, Border Security, etc.) and Defence Sector Reform.

23.1 BACKGROUND

A significant effort at US Africa Command was to better understand how to enhance weak security and defence sector institutions in some African countries so they might be capable of defending their countries and providing internal security. Liberia was among the first countries to receive US assistance. But by 2010, the United States had seen mixed results in their security reform efforts there. Liberia and its neighbour, Sierra Leone, had similar historical experiences involving the decline of government institutions, including defence and security. However, while US focused its security assistance on Liberia, the United Kingdom (UK) focused its assistance on Sierra Leone, and the UK effort was perceived as more effective. Therefore, it seemed reasonable to compare progress in Liberia with that in Sierra Leone to see what could be learned to make future reform efforts more effective. So, the authors were asked by the Deputy, Strategy, Plans, and Programs US Africa Command to evaluate the Defence Sector Reform in Sierra Leone.

The consensus amongst US Africa Command leadership was that the DSR activities conducted in Sierra Leone had been more effective than those conducted in Liberia. One piece of evidence used to support the "success" of SSR/DSR in Sierra Leone is the fact that in 2010, the Sierra Leoneans were contributing a unit to the African Union peacekeeping mission in Darfur; whilst the Liberians were still (as of August 2010) training below the company-level. This chapter presents an assessment, using a model developed by Dr. Jim North and Dr. Karsten Engelmann (both then at US Africa Command) of the Sierra Leone SSR/DSR efforts as of August 2010.

As the US Africa Command team learned through interviews, there had been progress in Sierra Leone's SSR/DSR by 2010 but there had also been many unsuccessful attempts within the SSR/DSR efforts. The success that had been achieved to date by the UK-led International Community resulted from five best practices:

- A lead donor (UK);
- Flexibility by that donor;
- Unity of action by the International Community;
- Persistent commitment; and
- A whole-of-security (whole of governance) approach.

However, issues regarding SSR/DSR in Sierra Leone remained. These included an occasional lack of unity of action, and constraints placed on the lead donors' activities by the Sierra Leone government. Perhaps the most significant issue regarding SSR/DSR according to one individual in the International Community was the attitude by the Sierra Leoneans of a desire for "sovereignty and independence of action with donor dependence monetarily". In other words, the Sierra Leoneans wanted to do as they wish with regard to reform but to have others pay for it. This lack of desire for true ownership by the Sierra Leonean government (ownership includes not only decision making – but the responsibility to pay for those decisions) threatened all the SSR/DSR achievements in Sierra Leone to date.

Although the two countries are neighbours and had similar recent experiences in the 1990s and 2000s, Liberia and Sierra Leone had had very different histories. Liberia was founded in the early 19th century by the United States as a repatriation site for slaves from America. Liberia has always been "independent', although with strong US influence. Sierra Leone meanwhile had always been a colony, either of the Portuguese or later of the United Kingdom until independence 1961. Both Liberia and Sierra Leone had significant and damaging civil wars in the 1990s (many argue the civil war in Sierra Leone was a "spill-over" to an extent from the civil war in Liberia). Although they both experienced civil war, the civil wars of Liberia and Sierra Leone were fought differently and had different impacts upon each country.

In recent times Liberia experienced two civil wars: From 1989 through 1996 ending when the warlord Charles Taylor was victorious (and affirmed head of the government through open election in July of 1997); and from 1999 through 2003 when eventually Taylor resigned and handed over power to the National Transitional Government of Liberia which was protected by UN forces.

The US Africa Command SSR/DSR Concept Model breaks Security Sector Reform into its sub-components: Reform of the non-military Security Sectors (Justice, Police, Border Security, etc.), and Defence Sector Reform. In the SSR/DSR Concept Model, the Defence Sector is further broken into two major components; defence institutions and defence capabilities. Defence institutions include the following components: Civilian control of the Military; Relationship of the Military with other Security Organizations; Military Legal Processes; personnel management (recruiting, vetting, selection, and retirement); and defence fiduciary responsibilities. Separate these and one might argue a by-product of strong defence institutions is the development of Defence Capabilities (i.e., combat forces). The US Africa Command SSR/DSR Concept Model places emphasis on the criticality of ownership of reform by the partner state, and the importance of fostering ownership by the International Community.

23.2 ORGANIZATION

The United States Africa Command, is one of nine unified combatant commands of the United States Armed Forces, headquartered at Kelley Barracks, Stuttgart, Germany. It is responsible for US military operations, including fighting regional conflicts, and military relations with 53 African nations. Its area of responsibility covers all of Africa except Egypt, which is within the area of responsibility of the United States Central Command.

Within the Command, the main assessment team is located in the J-1/8 – or the "Resources Directorate". The Sierra Leone assessment was carried out in support of the J-5, Directorate for Strategy, Engagement, and Programs.

23.3 MISSION

US Africa Command has three focus areas: Partnership, Operations, and Readiness. For Partnerships, through multinational exercises and military-to-military engagements, the Command aids African partner nations to build the defence capability and capacity of their security forces.

US Africa Command embraces a holistic approach to security challenges by working closely with US Government interagency counterparts and partner nation militaries from around the world.

23.4 A SECURITY SECTOR REFORM/DEFENCE SECTOR REFORM MODEL

To frame assessment, US Africa Command developed a model for Security Sector Reform and the subordinate Defence Sector Reform. Security Sector Reform, led by the US State Department, forms the basis upon which the Command conducts Defence Sector Reform (DSR). This section of the chapter is a brief overview of the SSR/DSR Concept Model developed by North and Engelmann. The authors adopted a definition of SSR as provided by the Organization for Economic Cooperation and Development (OECD).

23.4.1 Elements of Security Sector Reform

The OECD defines Security Sector Reform as follows: Transformation of the security sector, which includes all actors, their roles, responsibilities and actions, working together to manage and operate the system in a manner that is more consistent with democratic norms and sound principles of governance, and thus contributes to a well-functioning security framework. A graphical example of the various elements of SSR is presented in Figure 23-1.

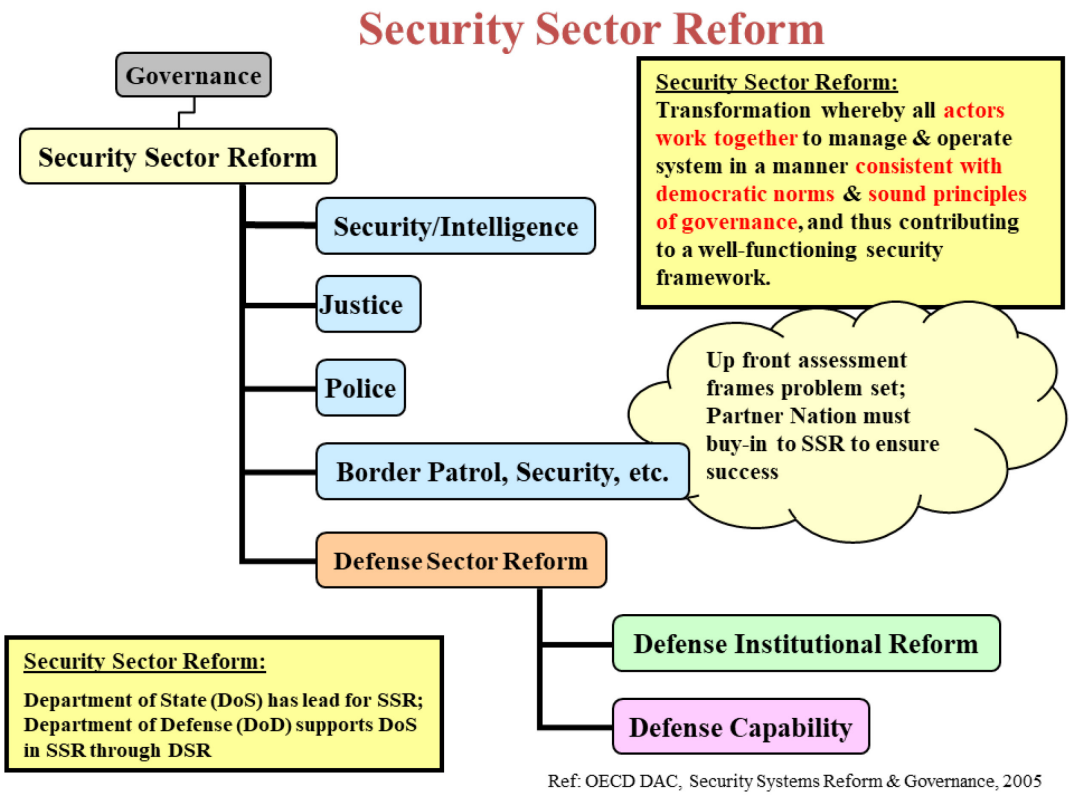


Figure 23-1: Security Sector Reform [1].

23.4.2 Elements of Defence Sector Reform

This section of the chapter is focused on the DSR Concept Model developed by US Africa Command in 2010. The Model consists of two components: reform of defence institutions and defence capability development. The various elements of DSR are presented in Figure 23-2 [2].[1]

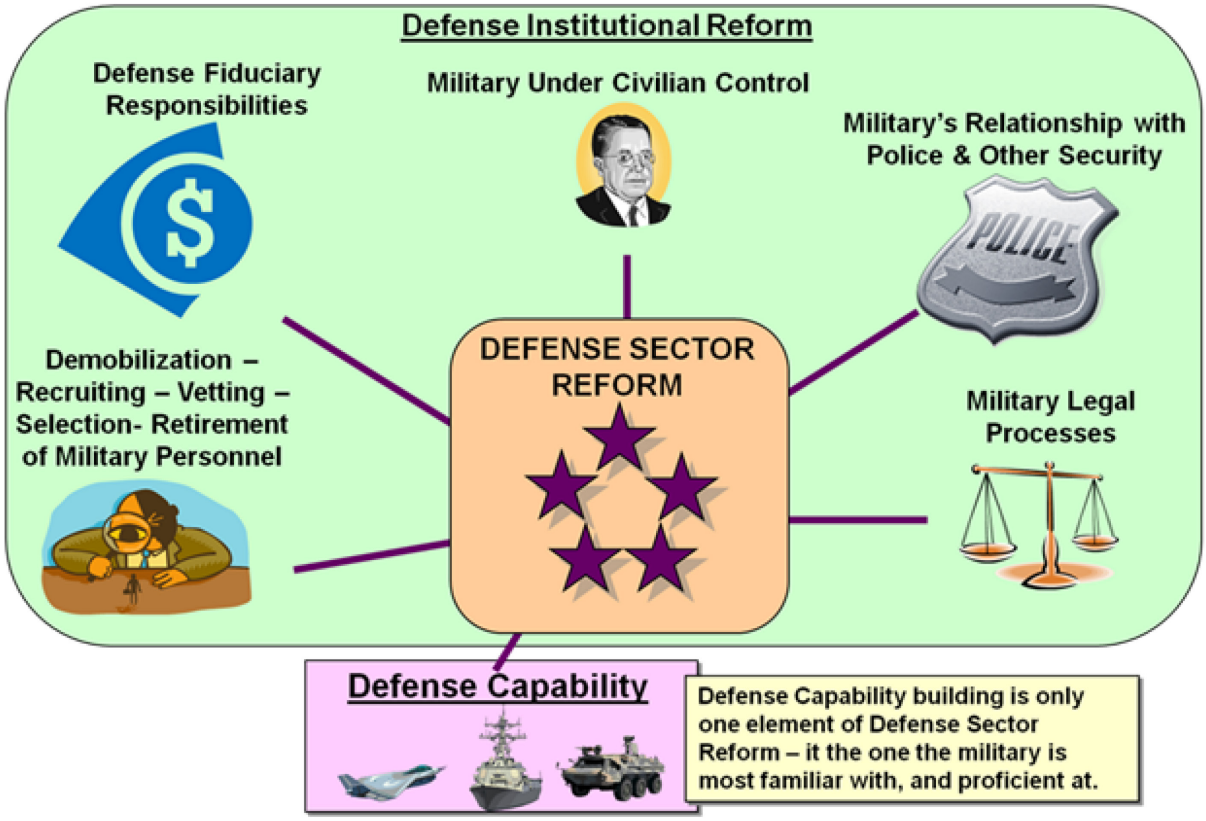


Figure 23-2: Elements of Defence Sector Reform.

Figure 23-2 identifies the main elements of the defence institutional aspect of DSR and highlights the importance of dealing with these elements while engaging in defence capability building. Sustained development of defence capabilities is derived from a strong defence institutional base. There are several areas recognized as important to generate good governance in defence institutions [3]. These areas are defined in detail below, and their application/implementation in Sierra Leone is covered under the "Discussion" section associated with each area.

23.4.2.1 Defence Sector is Under Civilian Control

Definition: A defence sector is under civilian control when:

1) The legislature has legal control over the defence sector and expertise and authority over defence matters in order to set defence policy; and
2) The Ministry of Defence has civilian leadership with expertise on defence matters capable of setting policy.

[1] This framework for DSR is derived from Ref. [2]. It does not include all aspects of the defence institutions, such as the development and implementation of defence strategy and force management. The elements presented are viewed by Ms. Boucher as the most important elements of required DSR.

While not part of a formal definition of civilian control, it is also advantageous for a state to have non-government institutional expertise (e.g., the Media and what is often called "Civil Society" – groups of interested citizens that form the basis of a functioning society) on military matters and their involvement in the debate about employment of the military.

23.4.2.2 Well-Defined Relationship with Police and Other Security Organizations

Definition: A well-defined relationship with police and other security organizations occurs when:

1) A National Security Strategy is in place defining the roles and responsibilities of various security organizations;
2) An understanding by the security organizations of the relationships between security sector organizations is built through exercising the relationships; and
3) The security organizations are mutually capable and able to build a supporting relationship.

23.4.2.3 Force Management and Fiduciary Responsibility

Definition: Force management and fiduciary responsibility occurs when the Ministry of Defence is capable of conducting policy planning and strategy development, force development and integration, acquisition and budgeting processes, and human resource management. At each level below the Ministry, processes are in place to ensure proper management of resources and fiduciary responsibility. This includes an auditing/inspection process. These systems should include necessary checks and balances.

23.4.2.4 Processes of Recruiting, Vetting and Selection of New Recruits and Officers; Demobilization and Retirement

Definition: Vetting and Selection of new recruits and officers occurs when the armed forces engages in the maintenance or expansion of the personnel base. The system that is implemented needs to include vetting for past human rights violations. The state should successfully recruit and select the best qualified recruits regardless of ethnicity, tribal relationships, and gender. Demobilization of military personnel after a conflict is another key element of the selection process. Finally a valid retirement system must be in place to ensure individuals are not working beyond capability, and to open up potential promotions for younger soldiers and officers.

23.4.2.5 Military Legal Processes and Regulations are in Place

Definition: The military needs a code for legal matters with commensurate processes, like the Uniformed Code of Military Justice (UCMJ) and subject matter experts should be assigned to ensure that legal matters are dealt with consistently and fairly.

23.4.3 SSR/DSR Model Summary

The main focus for the US Africa Command is Defence Sector Reform. However, DSR cannot progress without a larger Security Sector Reform in place. Even with SSR in place, traditionally, US Department of Defense (DoD) organizations tend to focus their engagement with partners on building partners' defence capabilities as opposed to defence institutions. This occurs for two reasons: First, legal constraints limit DoD's ability to work with certain elements (e.g., police) of a partner nation's government without official approval (i.e., of the US State Department or US Congress). Second, the DoD organizations prefer to train and equip partners in a way that builds "combat units" and associated combat support units. However, in order to build a professional military, it is critical that the other elements of DSR (the institutions) are also considered and developed because without strong defence institutions, it is difficult to sustain defence capabilities.

As it engages in DSR, the US Africa Command *must be aware* that institutional reforms will often be unpopular with entrenched elements of the partner government (and society). This is in part because DSR will invariably involve a change in the balance of powers and responsibilities with new winners and losers in the partner country.

Through focusing the Command's engagement on these institutional aspects as well as dealing with the potential institutional roadblocks, the Command can increase the likelihood that changes in defence capabilities, as well as institutional reform, will take hold and will be lasting. Regardless of the form that engagement takes, emphasis must be on reform ownership by the partner nation, and the Command must encourage the partner nation in the ownership of the reform of the defence sector.

23.5 SIERRA LEONE DSR/SSR ASSESSMENT AND LESSONS LEARNED

In making an assessment of DSR/SSR effectiveness, we must organize the data for assessment. Data can be organized into three categories: quantitative, qualitative, and perceptive. Quantitative data are of course hard numbers. Actual counting – such as the number of personnel in the Republic of Sierra Leone Armed Forces (RSLAF), or the amount of training ammunition expended in a given year. Qualitative information is about qualities; information that cannot actually be measured. Some examples of qualitative data are the success or not of civilian control of the military. Perceptive data – gathered through questionnaires, interviews, and observations – help researchers understand what individuals think about the security sector reform environment. People act according to what they believe about different topics, so if we want to change a group's perceptions, we have to know about their beliefs. Because the DSR Concept Model examined questions at the highest level (i.e., strategic in nature for Sierra Leone), the majority of the data needed (and available) was qualitative and perceptive data.

The Command learned several key items from the Sierra Leone assessment. First, the criticality of unity of command within the UK government in all efforts to assist with DSR/SSR in Sierra Leone. This unity of command allowed for a more focused effort on the part of the United Kingdom. The second lesson learned is the criticality of reform ownership on the part of the partner nation being assisted. If the partner nation does not have a vested interest in the development of their military forces with institutions capable of managing them, and does not take the lead (especially financial lead), there is little chance of success in the long-run.

23.5.1 Security Sector Reform in Sierra Leone

"I need 400 vehicles before the elections in 2012" stated Mr. Michael von der Schulenburg, Executive Representative of the Secretary General and head of the United Nations Integrated Peacebuilding Office in Sierra Leone (UNIPSIL). As with all individuals the team interviewed in Sierra Leone regarding SSR, continued logistical support for the security sector by donor countries was the paramount request. Even more than logistical support, the development of the economy is critical for the future of Sierra Leone. In addition, the team learned that increased support/coordination between the police and military is critical in Sierra Leone.

As the US Chargé stated, "Sierra Leone not being in civil war does not define success…this country could slide back [to civil war]." According to Ibrahim Abdullahone, of the major reasons for the last civil war was the lack of opportunity for many Sierra Leoneans [4]. The implication was clear – more than SSR is required. There also needs to be increased economic opportunities or "economic security" in Sierra Leone.

Because of how DSR had been conducted in Sierra Leone, and the availability of SSR/DSR officials, the US Africa Command team interviewed individuals mainly engaged in the larger SSR-related activities as well as DSR activities. Individuals interviewed included (in order interviewed):

- The US Defense Attaché (DATT) to Sierra Leone and the reserve officer filling in for the DATT while she was out of country;
- The Deputy Chief of Mission to the US Embassy (acting as Chargé as the Ambassador post was not filled);
- Lead US (UN) mentor to the Sierra Leone Police (SLP);
- Members of the Office of the Presidential Strategy and Policy Unit;
- Members of the International Military Assistance Training Team (IMATT);
- Chief of Defence Staff and the General Officers of the RSLAF;
- The Inspector General of the Sierra Leone Police;
- The Security Sector Coordinator for the UK Department for International Development (DfID);
- The Commander of the RSLAF 4th Infantry Brigade and his staff; and
- The head of UNIPSL.

The US Africa Command team interviewed such a wide range of individuals because just as DSR is only a small portion of SSR and governance, SSR is only a portion of the development efforts that were being conducted in Sierra Leone. The emphasis by the International Community (and the Sierra Leone government) was on SSR (and other elements of governance) – with DSR efforts supporting the larger SSR plan. The main DSR effort was being conducted by the British-led International Military Advisory and Training Team (IMATT).

23.5.2 Defence Sector Reform in Sierra Leone

Albrecht and Jackson, two UK researchers that examined the Security System Transformation in Sierra Leone 1997 – 2007, chose to characterise Sierra Leone's security reform process as "transformation", rather than simply the reform of one government sector [5]. From the late 1990s onwards, the UK was involved in attempting to rebuild/reform the Sierra Leone Military – as part of a much larger (in fact all-encompassing) SSR and governance reform. The early efforts by the UK (involving the military, police, and other security sector elements) had mixed results until a significant policy change occurred in 2002 which made the total effort much more effective.

In 2002 the United Kingdom and the Republic of Sierra Leone agreed to a ten-year memorandum of understanding that provided sweeping powers to the UK to support the rebuilding of Sierra Leone, and committed the UK to providing resources and expertise to achieve the rebuilding of Sierra Leone's Security Sector. Upon agreement, the UK poured resources, money, and personnel into Sierra Leone with an understanding that the effort was to be sustained for ten years.

In the beginning (2002), the United Kingdom approached DSR by reforming the then standing military. The reform included: integrating elements of all warring factions from the civil war, "correct-sizing" the Republic of Sierra Leone Armed Forces (RSLAF), the rebuilding of facilities, and training and equipping of the RSLAF.

Throughout the process, UK personnel were embedded in the Sierra Leone Security Sector elements. As a result, the reform process was "owned" by the Sierra Leoneans with expertise being supplied by the UK This is in dramatic difference to those efforts where a Western power will either enter at periodic intervals to "grade the papers" of the partner nation…or where the Western power conducts the security sector efforts by itself, preventing the partner nation from gaining expertise or taking ownership of the Security Sector Reform process.

The UK Department for International Development (DfID) served as lead organization for the UK in Sierra Leone's Security Sector Reform effort. This already empowered organization was given special authorities by Prime Minister Blair, allowing DfID to have a unity of command for the UK as well as the ability to employ all elements the UK's national power.

It is of significant value to the US Africa Command to evaluate the progress to date in Sierra Leone's Defence Sector Reform by applying the Command's DSR Concept Model. The evaluation, subjective in nature, rates each DSR element on an ordinal scale including:

- **FAILURE** (no progress has been made);
- **LIMITED SUCCESS** (some minor progress has been made);
- **MODERATELY SUCCESSFUL** (limited success achieved);
- **SUCCESSFUL** (DSR element is meeting basic goals); and
- **EXCEEDS SUCCESS** (DSR element is exceeding the basic goals).

The DSR Concept Model evaluated the following elements of DSR: Civilian Control; well-defined relationships with police and other security organizations; force management and fiduciary responsibility; recruiting, vetting, selection, and retirement; legal processes; and defence capabilities. For each of the elements of DSR examined, an evaluation was made of conditions in 2010.

23.5.2.1 Civilian Control

The overall assessment of Sierra Leone's Civilian Control element is **EXCEEDS SUCCESS**. The Republic of Sierra Leone has worked to develop a strong Civilian Control of the Security Sector (and Defence Sector) forces. The long-term commitment of the UK and flexibility and adaptability DfID enjoyed ensured that Civilian Control over the military (and other security forces) was strong. This is an amazing accomplishment given that at the end of the civil war in 2002, the elected president had no control over or support of the state's security forces.

When DfID began work in 2002, the Sierra Leone President also acted as the Minister of Defence. Whilst this was an appropriate structure during wartime, with peace, these two roles needed to be separated. The Parliament also has a strong interest in approving the RSLAF budgets. The Parliament had not funded security forces adequately. In 2009, the RSLAF budget approved by the Parliament was only 20 billion Leones ($5 million), out of a request of 70 billion Leones ($17.5 million)…and of this budget, 99% was spent on recurrent costs, like rice, fuel, and ammo [6]. However, the Parliament was actively helping to develop the security sector forces. The most significant example of creative civilian efforts to enhance the security sector was the Parliament's creation of MACP – the Military Assistance to Civilian Powers. This wide-ranging statute enabled civil authorities to mobilize elements of the military to support police – for example when a particularly violent series of armed robberies occurred in 2008, the authorities enacted MACP to provide the police with logistical and fire-power support. MACP was one of several agreements that helped to establish a well-defined relationship between the military and Sierra Leone's other security organizations.

23.5.2.2 Well-Defined Relationship with Police and Other Security Organizations

The overall assessment of Sierra Leone's military Relationship with Police and Other Security Organizations element is **SUCCESSFUL**. The Sierra Leone government, supported by the International Community, worked hard on two significant initiatives and several minor ones to develop and define relationships between the military and other security organizations were undertaken by Sierra Leone. Significant initiatives included the creation of the Office of National Security and the passing of the Military Aid to Civil Power act. Minor initiatives included the training and employment of military police elements.

The Office of National Security (ONS) believed that without security there could be no sustainable development. The ONS was created in 2002 by an act of the National Parliament. Starting from scratch, the ONS had a good chance of maintaining standards of avoiding corruption and maintaining apolitical behaviour in its staff [5].

The MACP act was created to reduce the "misconception of possible acrimony" between the two security organizations – the RSLAF and the SLP [7]. It was very successful – and was used multiple times by leadership. In 2010, President Ernest Bai Koroma invoked MACP for the Local Government Bi-Election that took place in the Lugbu Chiefdom, Bo District [8]. Although successful, the US Africa Command interviewers were told by some leaders that MACP was used too often. The RSLAF 4th BDE commander, who, although he supports MACP, was worried that the frequent common-place use of RSLAF alongside the SLP could cause the populace to associate the RSLAF with the petty bribery the SLP was frequently accused of. This worry was also echoed by the IMAAT who stated that MACP provides opportunities to showcase the RSLAF as a capable, competent agent of the state (while highlighting SLP deficiencies); but also presents opportunities for corrupt interaction with the public and private businesses [9].

23.5.2.3 Force Management and Fiduciary Responsibility

The overall assessment of Sierra Leone's Force Management and Fiduciary Responsibility element is **FAILURE**. This is the area the Sierra Leoneans are most lacking. The IMATT briefs that domestic tax revenues for Sierra Leone raise only 11% of GDP (comparatively, the UK taxes raised 36% of GDP and the US taxes raised 25% of GDP). Additional monies required to pay for government expenditures came from international donations. Corruption concerning fiduciary activities at all levels of the Security Sector continued to be a concern. The Sierra Leoneans established an Anti-Corruption Commission (ACC). In 2010, the ACC brought to light numerous individuals involved in corruption, including one senior-level official, the Minister of Fisheries. Corruption was so endemic in Sierra Leone that the country ranks 146 out of 180 in the world for corruption according to Transparency International [10]. Force management and fiduciary challenges for the RSLAF were significant. As IMATT asked, is a force of 8,500 for a population of about 6 million too large? According to IMATT, the RSLAF had difficulties managing the force when 90+ % of the budget is spent on recurrent costs (rice, fuel, ammo). The RSLAF faced deteriorating infrastructure and equipment with no provision for development. As a result, there is little opportunity to train unless the effort was IMATT-supported. As stated by many officials interviewed including IMATT, the government of Sierra Leone needed to concentrate on income generation opportunities. However, as described by one of the interviewees, the SL leadership desired full sovereignty and independence with donor dependence; or put another way, the SL leadership wanted the authority to spend without the responsibility to raise the funds.

23.5.2.4 Demobilization, Recruiting, Vetting, Selection, and Retirement

The overall assessment of Sierra Leone's Demobilization, Recruiting, Vetting, Selection, and Retirement element was **MODERATELY SUCCESSFUL**. Unlike Liberia, the Republic of Sierra Leone Armed Forces maintained a structure after the civil war. Upon conclusion of the civil war, the RSLAF underwent a Military Reintegration Plan (MRP). Demobilization of combatants was a major goal of the International Community. At Abuja in early 2001, an agreement was signed (Abuja 2) that called for disarmament of the rebel forces. It was not until late 2001 that the Disarmament, Demobilization, and Reintegration (DDR) process was considered to be making good progress [5] p. 62. In 2001, the RSLAF had 12,700 on its rolls. After the MRP process was developed, the RSLAF consisted of over 15,000 – including individuals from the official army as well as vetted members of the Revolutionary United Front (RUF) and the Civilian Defence Forces. Since 2002, the RSLAF decreased in numbers to 8,500 in November of 2009.

The interview team did not learn how the vetting process at the end of the civil war was accomplished. As stated above, over three thousand former rebels and local militias were vetted and added onto the roles of the

RSLAF. Because so many of the records were lost during the conflict (the RUF specifically focused attacks on government buildings in an attempt to create anarchy), proper vetting was impossible to conduct.

Selection in the RSLAF appeared to still be a difficult process. According to IMATT, there was significant corruption and patronage in the RSLAF. Additionally, there was a significant divide between generations – the older soldiers/officers had an attitude of "what's in it for me" and had too little sense of service. This was in contrast to the more junior officers and enthusiastic young soldiers that represent fertile ground for training and a more positive future. However, as IMATT states, the enthusiasm of these soldiers is tenuous – if issues were not resolved soon, their idealism could be lost [9].

Another significant issue regarding selection was that of retirement. The authors were told by interviewees that individuals up to 80 years old were still on the rolls of the Sierra Leone Military. This was caused by the lack of a retirement system. So elderly soldiers continued to "serve" in the RSLAF. This, in turn, provided few advancement slots for new soldiers because the more senior positions were filled with individuals who should have retired.

23.5.2.5 Legal Processes

The overall assessment of Legal Processes was **NOT COMPLETED.** Due to time constraints, this element of Defence Sector Reform was not examined.

Definition: The military needs a code for legal matters with commensurate processes, like the US Uniformed Code of Military Justice (UCMJ) and subject matter experts should be assigned to ensure that legal matters are dealt with consistently and fairly.

In addition to a state's defence institutions, a state's military capabilities are likely in need of development. To be effective as a military, soldiers and units must receive training and be capable of conducting required tasks identified in the national military strategy. This includes the development of forces capable of conducting combat. It also requires the development of combat support units capable of performing specialized functions, such as fire support and engineering. Combat service support, capable of sustaining operating forces, must be developed and capable of managing facilities and logistics. Finally, complementary direct level, organizational level, and strategic leadership expertise must be developed.

Each of these aspects of the defence sector must be addressed in a DSR effort. Some of these areas can be addressed without coordination with other Security Sector elements while others require coordination.

23.5.2.6 Defence Capabilities

The overall assessment of Sierra Leone's Defence Capabilities element was **MODERATELY SUCCESSFUL**. The Republic of Sierra Leone Armed Forces (RSLAF) was comprised of 8,500 personnel in nine battalions deployed in three brigades around the country (Figure 23-3 and Figure 23-4).

Sierra Leone's military is dominated by the army. The vast majority of the 8,500 personnel are in the army – which also receives the majority of the resources. Due in part to equipment shortages, the IMATT focused on army training. Sierra Leone had a small maritime wing – which was increasing in size and capabilities (although the IMATT noted that vessels gifted to the maritime wing of Sierra Leone were of dubious quality/serviceability [11]). The air wing was sold off to raise cash [6].

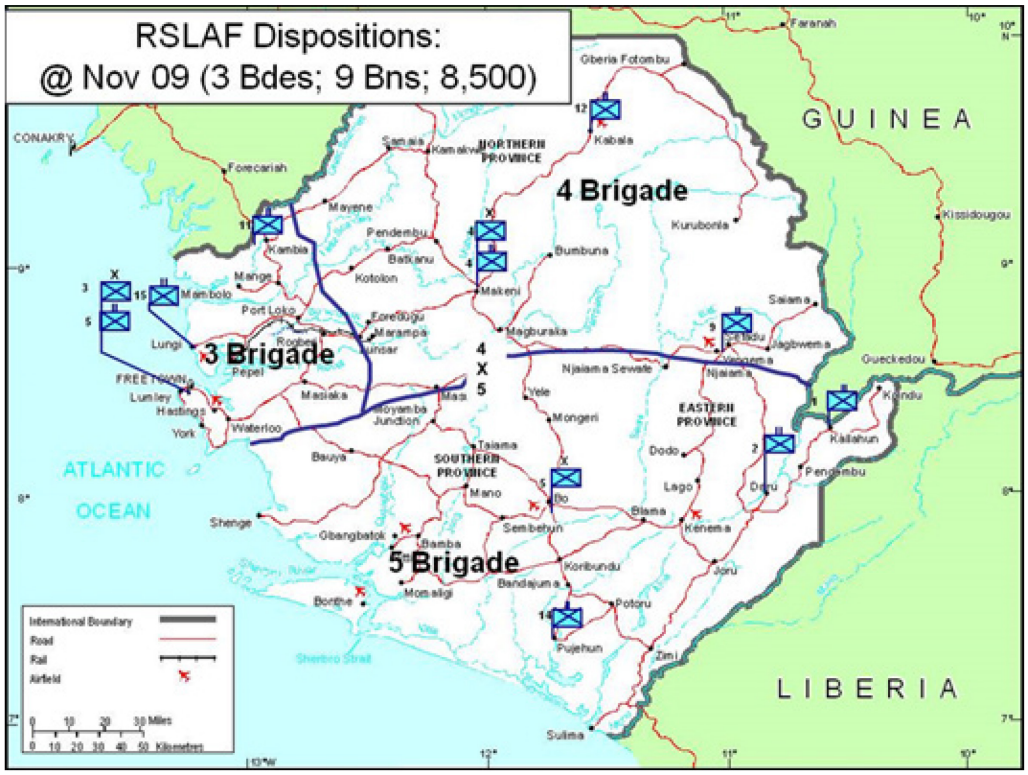


Figure 23-3: RSLAF Dispositions as of November 200 Shows Where the Brigades and Battalions were Based.

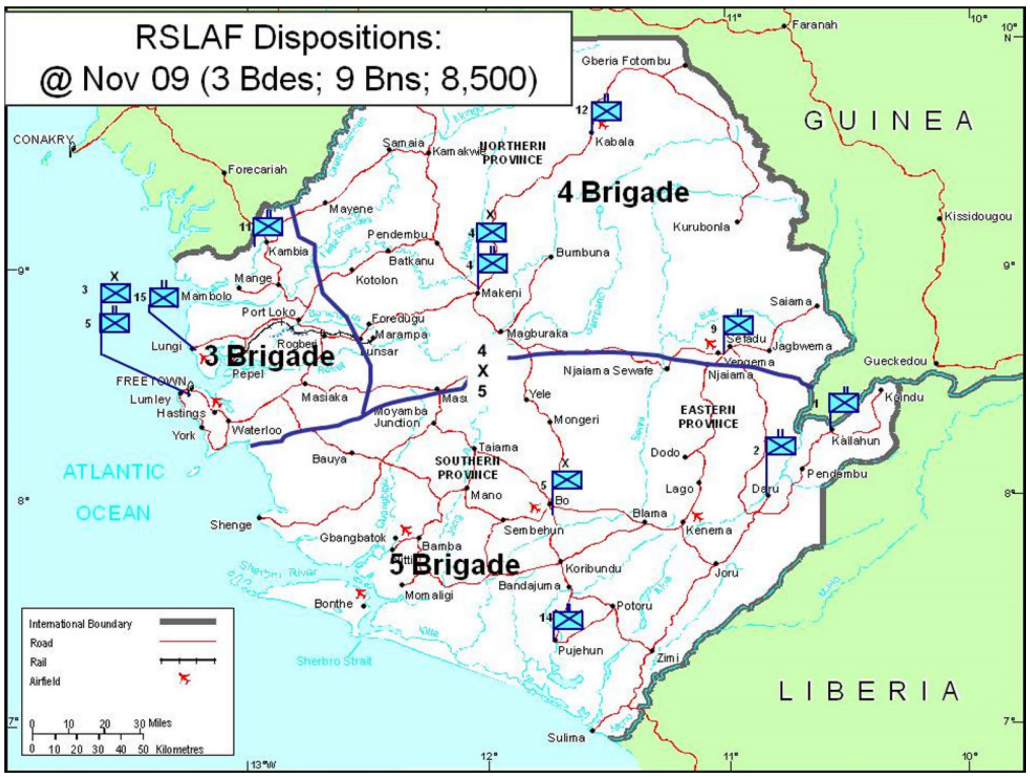


Figure 23-4: RSLAF Dispositions as of November 2009 [9].

23.6 CONCLUSIONS

Sierra Leone has since before the conclusion of the civil war worked in tandem with the International Community, especially the United Kingdom, to achieve security sector reform and defence sector reform. The purpose of the research team's visit to Sierra Leone was to help the US Africa Command execute future security assistance activities more effectively. This paper evaluates the progress in Sierra Leone's Defence Sector Reform along the elements of the DSR Concept Model and included: Civilian Control, well-defined relationships with police and other security organizations, force management and fiduciary responsibility, recruiting vetting selection and retirement, legal processes, and defence capabilities. The assessment for each of the model elements is outlined in Table 23-1.

Table 23-1: Evaluation of the Progress in Sierra Leone's Defence Sector Reform Along the Elements of the DSR Concept Model.

Element of Defence Sector Reform	Assessment / Evaluation
Civilian Control	EXCEEDS SUCCESS (DSR element is exceeding the basic goals)
Well-Defined Relationships with Police and Other Security Organizations	SUCCESSFUL (DSR element is meeting basic goals)
Force Management and Fiduciary Responsibility	FAILURE (no progress has been made)
Demobilization, Recruiting, Vetting, Selection, and Retirement	MODERATELY SUCCESSFUL (limited success achieved)
Military Legal Processes	NOT EVALUATED
Defence Capabilities	MODERATELY SUCCESSFUL (limited success achieved)

To achieve the successes Sierra Leone had with respect to SSR/DSR, three main conditions were required: Ownership, long-term commitment by a unified International community, and flexibility in the SSR/DSR process.

Ownership of the process by Sierra Leone as well as strong and enduring commitment by the United Kingdom (and other elements of the International Community) ensured success of SSR/DSR. This was also emphasised by authors who stated [5] p. 200:

> *The commitment of a core team of Sierra Leoneans' leading personalities, at both political and senior civil service levels, was absolutely critical in driving the reform process and exercising national ownership.*

The importance of International Community was also emphasised, but without ownership of the process by the Sierra Leonean leaders, far less would be achieved.

When impediments to SSR/DSR arose, or when leadership realized the wrong path had been taken, Sierra Leone and her partners showed flexibility and adaptability. As a result, many of the SSR/DSR efforts had been successful and resulted in the Sierra Leoneans being able to deploy a company-sized unit to Darfur. The flexibility and adaptability, as well as other practices mentioned in this paper, will help increase the effectiveness of future security assistance activities conducted by the US Africa Command.

23.7 AUTHORS' BIOGRAPHIES

Dr. Karsten Engelmann was liaison from the Centre for Army Analysis to the US Africa Command from 2008 – 2010. In that capacity he assisted the Command in developing assessment models and conducting assessments in order to increase the effectiveness of US Africa Command's activities in Africa. His main focus was country-level assessments. From 2010 – 2013, Dr. Engelmann supported US Army Africa (USARAF) and again worked assessments. While at USARAF he developed, with the rest of the USARAF Staff, a methodology for country-level assessment. These assessments evaluated the status of USARAF's African Partners' military capabilities in order to increase the effectiveness of USARAF activities on the African Continent.

Dr. James North's most recent prior assignment was as the CNA Representative to US Pacific Command in Hawaii. Previously, he was the CNA Representative to III MEF and Seventh Fleet in Japan and had an earlier assignment in Hawaii to work at Marine Forces Pacific. He also served as the CNA Representative to U. S. Africa Command in Stuttgart Germany, and to Carrier Strike Group Nine on the USS Abraham Lincoln. In addition, he spent eight years working at CNA headquarters in Northern Virginia. The focus of his analysis has ranged from future operational concepts to posture and logistics analysis. He also has been involved in analysis of operations assessment, security cooperation and building partner capacity. In 2001/2002, he spent time in Afghanistan providing analytical support to the Marine Corps for Operation Enduring Freedom-Afghanistan. Dr. North received a Bachelor of Arts in History from DePauw University and a Ph.D. in Economics from the University of Oregon.

23.8 REFERENCES

[1] Organization for Economic Co-Operation and Development, Development Assistance Committee. 2005. Security System Reform and Governance. p. 20.

[2] Boucher, A. 2009. Defense Sector Reform. Stimson Center, July.

[3] Bucur-Marcu, H. (editor). 2009. Essentials of Defence Institution Building. Geneva Centre for Democratic Control of the Armed Forces (DCAF), May.

[4] Abdullah, I. 2004. Between Democracy and Terror. Council for the Development of Social Science Research in Africa.

[5] Albrecht, P. and Jackson, P. 2010. Security System Transformation in Sierra Leone, 1997 – 2007. Geneva Centre for the Democratic Control of Armed Forces (DCAF).

[6] International Military Assistance Training Team (IMATT). 2009. Briefing to General Ward, September.

[7] Williams, A.C. 2010. Nelson, Major General, Personal Interview. Interview by Karsten Engelmann.

[8] Author unknown. 2010. Awareness Times Newspaper. 2010. Freetown, SL: Feb 24, 2010.

[9] International Military Assistance Training Team (IMATT). 2010. Briefing to General Ward, July.

[10] Transparency International. 2009. Corruption Perception Index.

[11] International Military Assistance Training Team (IMATT). 2010. Briefing to General Ward, May.

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's References	3. Further Reference	4. Security Classification of Document
	STO-TR-SAS-110 AC/323(SAS-110)TP/899	ISBN 978-92-837-2224-3	PUBLIC RELEASE

5. Originator

Science and Technology Organization
North Atlantic Treaty Organization
BP 25, F-92201 Neuilly-sur-Seine Cedex, France

6. Title

Operations Assessment in Complex Environments: Theory and Practice

7. Presented at/Sponsored by

Final Report of RTG SAS-110.

8. Author(s)/Editor(s)	9. Date
Multiple	December 2019
10. Author's/Editor's Address	**11. Pages**
Multiple	290

12. Distribution Statement

There are no restrictions on the distribution of this document. Information about the availability of this and other STO unclassified publications is given on the back cover.

13. Keywords/Descriptors

Assessment
Complex operations
Indicators
Measures of effectiveness
Measures of performance
Monitoring and Evaluation
Operation Assessment
Operations Assessment
Operations Research
Operations Research and Analysis

14. Abstract

The complexity of the environments in which NATO and its member nations conduct military operations is increasing. The rise of powerful non-state actors and the increasing likelihood of conflict taking place among civilian populations makes operations more difficult and the progress of operations toward their objectives more difficult to perceive. Hence, the increased attention we are paying to operations assessment, which is our attempt to determine the effectiveness of operations in these complex environments.

The key finding of this activity is that a clear, explicit focus on improving the effectiveness of the organisation performing the assessment of its activities, vice a focus on accounting for resources, improves the quality of the assessment findings and products. These, in turn, lead to more effective operations.

This report contains a variety of works on various theoretical topics: the assessment of the protection of civilians, containment of nuclear materials, strategic communication, and command and control systems; the potential of red teaming and surveys for assessment; the selection of indicators; and several other models or techniques for making sense of complex environments. Practical examples include real world assessment processes from the Pacific region, Afghanistan, the Horn of Africa, Western Africa, and major combat operations exercises.

NORTH ATLANTIC TREATY ORGANIZATION

BP 25
F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE
Télécopie 0(1)55.61.22.99 • E-mail mailbox@cso.nato.int

SCIENCE AND TECHNOLOGY ORGANIZATION

DIFFUSION DES PUBLICATIONS
STO NON CLASSIFIEES

Les publications de l'AGARD, de la RTO et de la STO peuvent parfois être obtenues auprès des centres nationaux de distribution indiqués ci-dessous. Si vous souhaitez recevoir toutes les publications de la STO, ou simplement celles qui concernent certains Panels, vous pouvez demander d'être inclus soit à titre personnel, soit au nom de votre organisation, sur la liste d'envoi.

Les publications de la STO, de la RTO et de l'AGARD sont également en vente auprès des agences de vente indiquées ci-dessous.

Les demandes de documents STO, RTO ou AGARD doivent comporter la dénomination « STO », « RTO » ou « AGARD » selon le cas, suivi du numéro de série. Des informations analogues, telles que le titre est la date de publication sont souhaitables.

Si vous souhaitez recevoir une notification électronique de la disponibilité des rapports de la STO au fur et à mesure de leur publication, vous pouvez consulter notre site Web (http://www.sto.nato.int/) et vous abonner à ce service.

CENTRES DE DIFFUSION NATIONAUX

ALLEMAGNE
Streitkräfteamt / Abteilung III
Fachinformationszentrum der Bundeswehr (FIZBw)
Gorch-Fock-Straße 7, D-53229 Bonn

BELGIQUE
Royal High Institute for Defence – KHID/IRSD/RHID
Management of Scientific & Technological Research
for Defence, National STO Coordinator
Royal Military Academy – Campus Renaissance
Renaissancelaan 30, 1000 Bruxelles

BULGARIE
Ministry of Defence
Defence Institute "Prof. Tsvetan Lazarov"
"Tsvetan Lazarov" bul no.2
1592 Sofia

CANADA
DGSIST 2
Recherche et développement pour la défense Canada
60 Moodie Drive (7N-1-F20)
Ottawa, Ontario K1A 0K2

DANEMARK
Danish Acquisition and Logistics Organization
(DALO)
Lautrupbjerg 1-5
2750 Ballerup

ESPAGNE
Área de Cooperación Internacional en I+D
SDGPLATIN (DGAM)
C/ Arturo Soria 289
28033 Madrid

ESTONIE
Estonian National Defence College
Centre for Applied Research
Riia str 12
Tartu 51013

ETATS-UNIS
Defense Technical Information Center
8725 John J. Kingman Road
Fort Belvoir, VA 22060-6218

FRANCE
O.N.E.R.A. (ISP)
29, Avenue de la Division Leclerc
BP 72
92322 Châtillon Cedex

GRECE (Correspondant)
Defence Industry & Research General
Directorate, Research Directorate
Fakinos Base Camp, S.T.G. 1020
Holargos, Athens

HONGRIE
Hungarian Ministry of Defence
Development and Logistics Agency
P.O.B. 25
H-1885 Budapest

ITALIE
Ten Col Renato NARO
Capo servizio Gestione della Conoscenza
F. Baracca Military Airport "Comparto A"
Via di Centocelle, 301
00175, Rome

LUXEMBOURG
Voir Belgique

NORVEGE
Norwegian Defence Research
Establishment
Attn: Biblioteket
P.O. Box 25
NO-2007 Kjeller

PAYS-BAS
Royal Netherlands Military
Academy Library
P.O. Box 90.002
4800 PA Breda

POLOGNE
Centralna Biblioteka Wojskowa
ul. Ostrobramska 109
04-041 Warszawa

PORTUGAL
Estado Maior da Força Aérea
SDFA – Centro de Documentação
Alfragide
P-2720 Amadora

REPUBLIQUE TCHEQUE
Vojenský technický ústav s.p.
CZ Distribution Information Centre
Mladoboleslavská 944
PO Box 18
197 06 Praha 9

ROUMANIE
Romanian National Distribution
Centre
Armaments Department
9-11, Drumul Taberei Street
Sector 6
061353 Bucharest

ROYAUME-UNI
Dstl Records Centre
Rm G02, ISAT F, Building 5
Dstl Porton Down
Salisbury SP4 0JQ

SLOVAQUIE
Akadémia ozbrojených síl gen.
M.R. Štefánika, Distribučné a
informačné stredisko STO
Demänová 393
031 01 Liptovský Mikuláš 1

SLOVENIE
Ministry of Defence
Central Registry for EU & NATO
Vojkova 55
1000 Ljubljana

TURQUIE
Milli Savunma Bakanlığı (MSB)
ARGE ve Teknoloji Dairesi
Başkanlığı
06650 Bakanliklar – Ankara

AGENCES DE VENTE

The British Library Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
ROYAUME-UNI

Canada Institute for Scientific and Technical Information (CISTI)
National Research Council Acquisitions
Montreal Road, Building M-55
Ottawa, Ontario K1A 0S2
CANADA

Les demandes de documents STO, RTO ou AGARD doivent comporter la dénomination « STO », « RTO » ou « AGARD » selon le cas, suivie du numéro de série (par exemple AGARD-AG-315). Des informations analogues, telles que le titre et la date de publication sont souhaitables. Des références bibliographiques complètes ainsi que des résumés des publications STO, RTO et AGARD figurent dans le « NTIS Publications Database » (http://www.ntis.gov).

NORTH ATLANTIC TREATY ORGANIZATION

BP 25
F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE
Télécopie 0(1)55.61.22.99 • E-mail mailbox@cso.nato.int

SCIENCE AND TECHNOLOGY ORGANIZATION

DISTRIBUTION OF UNCLASSIFIED STO PUBLICATIONS

AGARD, RTO & STO publications are sometimes available from the National Distribution Centres listed below. If you wish to receive all STO reports, or just those relating to one or more specific STO Panels, they may be willing to include you (or your Organisation) in their distribution.

STO, RTO and AGARD reports may also be purchased from the Sales Agencies listed below.

Requests for STO, RTO or AGARD documents should include the word 'STO', 'RTO' or 'AGARD', as appropriate, followed by the serial number. Collateral information such as title and publication date is desirable.

If you wish to receive electronic notification of STO reports as they are published, please visit our website (http://www.sto.nato.int/) from where you can register for this service.

NATIONAL DISTRIBUTION CENTRES

BELGIUM
Royal High Institute for Defence – KHID/IRSD/RHID
Management of Scientific & Technological Research for Defence, National STO Coordinator
Royal Military Academy – Campus Renaissance
Renaissancelaan 30
1000 Brussels

BULGARIA
Ministry of Defence
Defence Institute "Prof. Tsvetan Lazarov"
"Tsvetan Lazarov" bul no.2
1592 Sofia

CANADA
DSTKIM 2
Defence Research and Development Canada
60 Moodie Drive (7N-1-F20)
Ottawa, Ontario K1A 0K2

CZECH REPUBLIC
Vojenský technický ústav s.p.
CZ Distribution Information Centre
Mladoboleslavská 944
PO Box 18
197 06 Praha 9

DENMARK
Danish Acquisition and Logistics Organization (DALO)
Lautrupbjerg 1-5
2750 Ballerup

ESTONIA
Estonian National Defence College
Centre for Applied Research
Riia str 12
Tartu 51013

FRANCE
O.N.E.R.A. (ISP)
29, Avenue de la Division Leclerc – BP 72
92322 Châtillon Cedex

GERMANY
Streitkräfteamt / Abteilung III
Fachinformationszentrum der Bundeswehr (FIZBw)
Gorch-Fock-Straße 7
D-53229 Bonn

GREECE (Point of Contact)
Defence Industry & Research General Directorate, Research Directorate
Fakinos Base Camp, S.T.G. 1020
Holargos, Athens

HUNGARY
Hungarian Ministry of Defence
Development and Logistics Agency
P.O.B. 25
H-1885 Budapest

ITALY
Ten Col Renato NARO
Capo servizio Gestione della Conoscenza
F. Baracca Military Airport "Comparto A"
Via di Centocelle, 301
00175, Rome

LUXEMBOURG
See Belgium

NETHERLANDS
Royal Netherlands Military Academy Library
P.O. Box 90.002
4800 PA Breda

NORWAY
Norwegian Defence Research Establishment, Attn: Biblioteket
P.O. Box 25
NO-2007 Kjeller

POLAND
Centralna Biblioteka Wojskowa
ul. Ostrobramska 109
04-041 Warszawa

PORTUGAL
Estado Maior da Força Aérea
SDFA – Centro de Documentação
Alfragide
P-2720 Amadora

ROMANIA
Romanian National Distribution Centre
Armaments Department
9-11, Drumul Taberei Street
Sector 6
061353 Bucharest

SLOVAKIA
Akadémia ozbrojených síl gen M.R. Štefánika, Distribučné a informačné stredisko STO
Demänová 393
031 01 Liptovský Mikuláš 1

SLOVENIA
Ministry of Defence
Central Registry for EU & NATO
Vojkova 55
1000 Ljubljana

SPAIN
Área de Cooperación Internacional en I+D
SDGPLATIN (DGAM)
C/ Arturo Soria 289
28033 Madrid

TURKEY
Milli Savunma Bakanlığı (MSB)
ARGE ve Teknoloji Dairesi Başkanlığı
06650 Bakanliklar – Ankara

UNITED KINGDOM
Dstl Records Centre
Rm G02, ISAT F, Building 5
Dstl Porton Down, Salisbury SP4 0JQ

UNITED STATES
Defense Technical Information Center
8725 John J. Kingman Road
Fort Belvoir, VA 22060-6218

SALES AGENCIES

The British Library Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
UNITED KINGDOM

Canada Institute for Scientific and Technical Information (CISTI)
National Research Council Acquisitions
Montreal Road, Building M-55
Ottawa, Ontario K1A 0S2
CANADA

Requests for STO, RTO or AGARD documents should include the word 'STO', 'RTO' or 'AGARD', as appropriate, followed by the serial number (for example AGARD-AG-315). Collateral information such as title and publication date is desirable. Full bibliographical references and abstracts of STO, RTO and AGARD publications are given in "NTIS Publications Database" (http://www.ntis.gov).

ISBN 978-92-837-2224-3